EXPERIENCING INTERACTION DESIGN: A PRAGMATIC THEORY

Volume 2

by

RONALD LENGKONG WAKKARY

A thesis submitted to the University of Plymouth
in partial fulfillment for the degree of

DOCTOR OF PHILOSOPHY

School of Computing, Communications & Electronics
Faculty of Technology

November 2009
COPYRIGHT STATEMENT

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without the author’s prior consent.

© Ronald Lengkong Wakkary, November 18, 2009
ABSTRACT

Experiencing Interaction Design: A Pragmatic Theory

Ronald Lengkong Wakkary

This thesis contributes a theory for the field of interaction design based on philosophical pragmatism. The theory frames interaction design as a pragmatic experience shaped by the inquiries of designers. The contributions of the theory are that it positions the designer at the centre of a theory, describes interaction design practice to be more than a collection of methods and strategies, and provides a sound basis for generating and verifying new knowledge through design. The thesis describes and analyzes two interaction design research projects through self-reflexive accounts that illustrate the proposed theory. The projects are a tangible museum guide and a responsive environment for physical play.

The thesis examines the value of understanding interaction design through pragmatism and how interaction design when viewed as experience opens the field up to a new theoretical framework. The two interaction design research projects are described as design inquiries constituted by a design inquirer, designer intentions, and design rationales. Further descriptions of the projects show interaction design to be comprised of design actions based on judgment and interpretation. Interaction design can be assessed by the degree to which there is integrity between the design inquiry and design actions, as well as by the transferability and discursiveness of the design inquiry findings that are relevant to the wider field of interaction design and related disciplines like human-computer interaction. The implications of the theory lead to new ways of mobilizing interaction design research and interaction design education. The pragmatic theory shows capacity for clear descriptions and analysis of interaction design inquiries in ways that extract and communicate new knowledge from interaction design practice and research. The theory shows interaction design to be a distinct and independent field of inquiry that generates knowledge through design.
LIST OF CONTENTS – VOL. 2

Copyright Statement................................................................................................................................. 1
Abstract.................................................................................................................................................. 3
List of contents – Vol. 2 ............................................................................................................................ 4
Appendix 4 - Related publications: journal articles................................................................. 5
Appendix 5 - Related publications: conference proceedings.................................................. 128
Appendix 6 - Related publications: book chapters................................................................. 243
Appendix 7: Abstracts of other publications.................................................................................. 266
APPENDIX 4 - RELATED PUBLICATIONS: JOURNAL ARTICLES


Situating Approaches to Interactive Museum Guides

Ron Wakkary*, Kevin Muise, Karen Tanenbaum, Marek Hatala & Leora Kornfeld

School of Interactive Arts and Technology, Simon Fraser University, British Columbia, Canada

(Received 1 May 2008; final version received 29 July 2008)

This paper examines the current state of museum guide technologies and applications in order to develop an analytical foundation for future research on an adaptive museum guide for families. The analysis focuses on three critical areas of interest in considering group and social interaction in museums: tangibility — the role of tangible user interfaces; interaction — visit types and visit flows; and adaptivity — user modeling approaches. It concludes with a discussion of four interrelated trajectories for interactive museum guide research including embodied interaction, gameplay, transparent and opaque interaction and the role of personal digital assistants.

Keywords: museum guides; adaptive museum guides; tangible user interfaces; museum user models; group interaction; visitor experience; digital heritage

Introduction

Museum technologies have increasingly become the focus of research in such areas as ubiquitous computing, tangible computing and user modeling. Since the advent of audio-based museum guides, much research and development has been placed on increasing the technological capacity for augmenting the museum visit experience. Early prototypes such as Bederson (Bederson 1995) provided evidence that it is possible to support visitor-driven interaction through wireless communication, thus allowing visitors to explore the museum environment at their own pace. More recent prototypes and fully functional systems are much more complex, supporting a variety of media, adaptive models and interaction modalities. However, as Bell notes in her paper on museums as cultural ecologies: 'The challenge here is to design information technologies that help make new connections for museum visitors' (Bell 2002). Bell further argues for the importance of sociality in museums, where visits are as equally social and entertaining as they are educational. Addressing the social qualities of a museum visit remains a clear challenge to designing with museum guide technologies. While families are by far the most common visitor type to science, history and natural history museums, few systems are designed for families.

In this paper, we describe our theoretical analysis of the current state of museum technologies and adaptive museum applications in order to provide a detailed understanding in support of our current research goals. We have narrowed our focus to three critical trends in adaptive museum guides that form the theoretical anchors of our future research. The areas include: tangibility, interactivity and adaptivity. We conclude with a discussion of research trajectories and their benefits. We believe this paper provides a

*Corresponding author. Email: rwakkary@sfu.ca
critical summary of electronic museum guides of the last decade and outlines key areas of concerns for academic researchers.

Background
At this stage in our research in adaptive museum guides we are exploring how best to address issues of social engagement, play and learning for groups of visitors such as families. Our early research in the sociality of mobile devices (Wakkary et al. 2001) and technical systems in support of virtual and adaptive learning (Richards and Hatala 2005) led us to the design and technical research of museum guides. Previously, we researched a prototype tangible and adaptive museum guide system known as ec(h)o that we discuss at several points in this paper. The aims of the research included evaluation of the role of tangibility and play in user experience (Wakkary and Hatala 2006, 2007) and evaluation of user modeling (Hatala and Wakkary 2005; Hatala, Wakkary, and Kalanlari 2005). In exploring tangibility, we researched how we could integrate our tangible user interface with visitors’ playful and curious engagement with physical interactives in a museum. Our research in user models extended work on museum visitor types in which different patterns of museum visits by different visitors can be described in user profiles. We used ambient and stereoscopic audio to display information to users of our system. This created a novel interface but, because users were required to wear binaural headphones, it had the serious drawback of limiting social engagement.

In response to the inherent antisocial aspects of ec(h)o, we researched an embodied interaction prototype known as socio-ec(h)o. Embodied interaction systems treat the physical world as a medium rather than a metaphor as in virtual three-dimensional (3-D) environments. Embodied interaction systems are both tangible and social as they ‘are manifest in our environment and are incorporated in our everyday activities’ (Dourish 2001). Socio-ec(h)o comprised a prototype environment for group play. The research goals were to explore the design of an embodied interaction system utilizing ambient lights and audio, a method for composing group user models and group interaction utilizing a game structure. Relevant research to date has focused on the technical platform (Wakkary et al. 2005), interaction and gameplay (Wakkary et al. 2005, 2007), group interaction (Wakkary et al. 2008) and user group models (Jiang 2008). While this project did not take place in a museum setting, it enabled us to extend research issues from ec(h)o into issues of sociality and group collaboration.

Our most current research, a project named Kurio, motivates this present discussion on tangibility, interactivity and adaptivity. Kurio is a museum guide system aimed at families and friends visiting the museum. In Kurio a family imagines themselves as time travelers lost in the present because their time map is broken. In order to repair the time map, family members complete small tasks and collect information from the museum reconstructing the time map each time. In addition to the game-like interaction, the interaction design and underlying user group model are based on a constructivist-learning model. The user group model relies on the constructivist schema to adapt to the different learning behaviors of each family member and assign appropriate tasks to each member. The interactive museum guide itself is distributed over a number of different components, tangibles, a tabletop computer and a personal digital assistant (PDA). The main component is a set of tangible computing devices we designed. Overall we have five tangibles including a pointer for selecting museum artifacts; a finder for finding different locations in the exhibition space; a listener for hearing audio files in different locations in the exhibition space; a reader for collecting text from didactic displays; and a gesturer for mimicking and collecting gestures.
relevant to different artifacts on display (see Figure 1). At this stage it is too early in the research to report on the design, assessment or findings. Rather, we describe the project to discuss what motivated the theoretical discussion in this paper. The analysis reported here is a result of our theoretical discussion and rationale that preceded the design of Kurio.

**Situating tangibility in museums**

Tangibility, or tangible user interfaces (TUI), imbue physical artifacts with computational capabilities. Hiroshi Ishii and Brygg Ullmer introduced the notion with the salient phrase of 'coupling of bits and atoms' (Ishii and Ullmer 1997). By their own account, scientific instruments from a museum collection at Harvard University inspired them. They experienced a quality of aesthetics in the oak and brass instruments on display that, in their minds, have been lost with the advent of computing. It became their aim to 'rejoin the richness of physical world with HCI [human-computer interaction]' (Ishii and Ullmer 1997). Ishii and Ullmer's idea of tangible computing built on earlier work on grasappable interfaces (Fitzmaurice et al. 1995) and real-world interface props (Hinckley et al. 1994). They describe TUIs as the 'seamless coupling of everyday grasappable objects (e.g. cards, books, models) with digital information that pertains to them' (Ishii and Ullmer 1997). Later they refined and expanded the definition of TUIs to include:

1. Physical representations are computationally coupled to underlying digital information;
2. Physical representations embody mechanisms for interactive control;
3. Physical representations are perceptually coupled to actively mediate digital representations;
4. Physical state of tangibles embodies key aspects of the digital state of a system.
   (Ullmer and Ishii 2001)

In 1992, Durrell Bishop's Marble Answering Machine (Crampton-Smith 1995) was an early embodiment of the immediate and playful qualities of tangible user interfaces. The prototype uses marbles to represent messages on an answering machine. A person replays the message by picking up the marble and placing it in an indentation in the machine. Natalie Jerimijenko's Live Wire is a strikingly minimal and whimsically simple demonstration of digital bits transformed into physical atoms. Jerimijenko dangled a plastic wire from a motor attached to the ceiling: the motor accelerates or decelerates based on the amount of traffic across the computer network.

In our previously mentioned research we developed a prototype system known as ec(h)o. In this project, simple physical interactives and wooden puzzles at the natural

![Figure 1. Components of the Kurio adaptive museum guide.](image-url)
history museum inspired us. As a result, the user interface for *ee(h)o* was a TUI that coupled a wooden cube with digital navigation and information. In *ee(h)o*, museum visitors held a light wooden cube and were immersed in a soundscape of natural sounds of and information on the artifacts on display (see Figure 2). Visitors navigated the audio options presented to them by rotating the cube in their palm in a direction that corresponded to the spatial location of the audio they were hearing [we will return to this project below but for a detailed account see (Wakkary and Hatala 2007)].

In our experience with TUIs, they bridge the virtuality of the museum guide system and the physical surroundings of the exhibition. As such, the adaptive museum guide first becomes more integral to the physical ecology of the exhibition including artifacts, display systems and architecture rather than being a separate technology. It is important to note that our social interactions are mediated, in large part, by objects and environments as much as they are by direct contact with others (Kaptelinin and Nardi 2006). Awareness of context is critical to sociality. Secondly, with TUIs, our understanding of interacting with technology is informed by our rich experience with physical artifacts and surroundings – because we can leverage existing embodied and cognitive understanding. Analytically, Kenneth P. Fishkin (2004) provides a taxonomy that allowed us to explore these factors further. Fishkin’s taxonomy is a two-dimensional space across the axes of embodiment and metaphor. Embodiment characterizes the degree to which ‘the state of computation’ is perceived to be in or near the tangible object. As we discussed, it expresses the integration of computation with the physical artifact and environment. Fishkin details four levels of embodiment:

1. **Distant** – representing the computer effect is distant to the tangible object;
2. **Environmental** – representing the computer effect is in the environment surrounding the user;
3. **Nearby** – representing the computer effect as being proximate to the object; and
4. **Full** – representing the computer effect is within the object.

Along the second axis, Fishkin uses metaphor to depict the degree to which the system response to user’s action is analogous to the real-world response of similar actions – the existing embodied and cognitive mappings. Fishkin divides metaphor into noun metaphors,

![Figure 2. The tangible-user interface of ee(h)o.](image-url)
referring to the shape of the object, and verb metaphors, referring to the motion of an object. Metaphor has five levels:

1. None – representing an abstract relation between the device and response;
2. Noun – representing morphological likeness to a real-world response;
3. Verb – representing an analogous action to a real-world response;
4. Noun + verb – representing the combination of the two previous levels; and
5. Full – representing an intrinsic connection between real-world response and the object which requires no metaphorical relationship.

In Figure 3, we have applied Fishkin's taxonomy to ec(h)o. Embodiment would be considered 'environmental' because the computational state would be perceived as surrounding the visitor, given the spatialized audio display output. With regard to metaphor, the ec(h)o TUI would be a 'noun and verb' as the wooden cube is reminiscent of the wooden building blocks and the motion of the cube determine the spatiality of the audio – as turning left, in the real world, would allow the person to hear on the left. If we consider the visitor’s movement, the embodiment factor would still be environmental and we would have to consider the visitor’s body as being ‘full’ in Fishkin’s use of metaphor. So in representing the entire system, we plotted ec(h)o between ‘noun + verb’ and ‘full’ on the metaphor axis.

It is natural to think that optimal TUIs would be ‘full’ in both embodiment and metaphor dimensions, and Fishkin suggests as much. However, we caution that in the case of museum guides and, in particular, when considering the sociality of museum visits, embodiment is between people, technology and the environment. Accordingly, TUIs in museums are optimal when computation is perceived to be environmental and nearby rather than solely within the device.

In summary, we believe tangibility is critical in considering the social dimensions of museum visits in museum guides. TUIs can integrate the computational and physical affordances of the museum visit experience, and visitors can leverage embodied and cognitive models of interaction in incorporating new technology. Visitors benefit from augmentation of computation yet maintain awareness and embodied connection to their surroundings, which ultimately supports social interaction.
Situating interactivity in museums

Interactivity can be understood in many ways. We focused on individual or group experiences in museum visits and models of content delivery. In this section, we review a range of systems from the last decade with a particular eye on how these past approaches provide insights into designing for families in which sociality and group activity are paramount. In reviewing the various systems, we have developed a matrix that compares the visit type (individual/group) to visit flow (information delivery/game-interaction) which aims to uncover the major factors that affected the design of the various systems.

In visit types, we refer to **individual** as systems that target interaction with single individuals, whereas **group** refers to interaction models aimed at groups of individuals. **Information delivery** refers to the approach of an information corpus or repository that is presented to the visitor, whether adaptively or by user selection. Providing the visitor with the ability to interact with the exhibit content in a playful manner we have referred to as **game-interaction**. Often the **game-interaction** approach will use games as a means to educate the visitor, as opposed to providing information about an artifact. By employing a cross-matrix of both types of categorizations, it becomes possible to understand how both past and current museum guides mediate the museum experience.

Many of the systems we reviewed were **information delivery** in style, while also falling under the **individual** category. Typically they involved a PDA and an audio/visual interface such as the Blanton iTour (Manning and Sims 2004). Among the first of such guides was the **HyperAudio** project (Petrelli and Not 2005), which used an adaptive model that we will discuss again in the next section. In the **HyperAudio** system, individual visitors were encouraged to walk about the exhibition with the handheld device and headphones, stopping at various exhibits to learn more about specific artifacts in their surrounding. As noted by Petrelli and Not, 'the presentation (audio message and hypermedia page) would be adapted to each individual user, taking into account not only their interaction with the system, but also the broad interaction context, including the physical space, the visit so far, the interaction history, and the presented narration' (Petrelli and Not 2005). The presentation that is displayed also provides a link that the participant can click on with a palmtop pen to gather further information about the artifact of interest. By so doing, the visitor is provided a map of the museum on which the location of the new artifact is displayed, to allow the visitor to see the artifact in person if he desires. The authors also describe their interest in keeping the graphical user interface to a minimum so as not to distract the visitor. It is for this reason that audio is chosen as the main channel of information delivery.

The **PEACH** project (Stock et al. 2007) provides the visitor with a digital character on their PDA that provides information on various artifacts within the exhibition. The **PEACH** guide also contains rich media in the form of video close-ups of frescoes and detailed descriptions of paintings. Unique to this system is the availability of a printout, which provides the visitor with an overview of the exhibits he encountered while at the museum.

We have observed, in a commercial system developed by a member of our team as part of Ubiquity Interactive known as VUEguide, that rich media images can at times become more engaging than the authentic object on display. We believe the virtual image can be integrated easily into a narrative world of information delivery. For example, interactive images of Bill Reid’s ‘Raven and the First Men’ sculpture engaged visitors deeply with the screen and encouraged a 'back and forth' exploration between the actual and the virtual (see Figure 4).
Figure 4. Interactive image of a Haida sculpture that visitors found engaging in the VUEguide system.

Berkovich et al. (2003) developed the Discovery Point prototype, a four-button device that delivers audio to the museum visitor. The device allows the visitor to listen to short audio clips about artifacts that are controlled by the visitor. The visitor can also create a virtual souvenir by pressing the ‘MailHome’ button, which adds the artifact in question to their personal website created upon entering the museum. The device functions without the use of headphones. Instead, audio is delivered through specialized speakers which direct audio to a precise area around the artifact, so as not to disturb other museum visitors.

The Sotto Voce system, produced by Aoki et al. (2002) provided the first instance where researchers focused on group activity. The system contained an audio sharing application called eavesdropping that allowed paired visitors to share audio information with each other while on an information delivery tour. In designing the application with three main factors in mind (the information source, the visitor’s companion and the museum space), the authors remarked that the system showed that the visitors used the system in creative ways, and with social purpose, while presenting the opportunity for co-present interaction.

In recent years, research has continued on group-based museum tours as an area of interest. Additionally, orientation has shifted from information delivery tours to game-interaction activities. This is evident in the CoCicero project implemented in the Marble Museum in Carrera, Italy (Laurillau and Paternò 2004). The CoCicero prototype focused on four types of group activities; (i) shared listening – similar to the Sotto Voce system; (ii) independent use – to allow individuals to choose not to engage in group-based tours; (iii) following – to allow an individual to lead other members of a group; and (iv) checking in – which allows members in a group to know how others are doing through voice communication while not being physically present. The authors state that communication, localization, orientation and mutual observation are four elements that are key to a collaborative visit. The guide functions by providing museum groups a series of games, such as a puzzle and multiple-choice questions, which require the visitors to gather clues through viewing the exhibits within the museum.

Similarly, the ARCHIE project (Loon et al. 2007) has developed a learning game for school children that allows visitors to trade museum-specific information to gain points in order to win a game. With each player having a specific role that he plays, the visitor must understand various levels of information gained from exploring the museum in order to improve his score. Although the ARCHIE project is only in its prototype stage, it is clear from their initial findings that the game-interaction approach does foster user interest in
museum content. Our own system, \textit{ec(h)o}, is unique in relation to other systems. It is exclusively an individual type system (which we now see as a significant drawback) yet it used a \textit{game-interaction} approach. Our interface aimed for an interaction based on an open-ended game qualities or what we referred to as play. The play took on two forms: (1) \textit{content play} in the delivery of information in the form of puns and riddles; and (2) \textit{physical play} that consisted of holding, touching and moving through a space: simple playful action along the lines of toying with a wooden cube.

In summary, it is evident that the majority of research and development has been in the area of \textit{information delivery} and \textit{individual} visit types; see Figure 5, where we plotted the different systems we reviewed. \textit{Sotto Voce} is among the first to explore co-visiting or group interaction with a museum guide within an information delivery context. Group interaction is a new area and chronologically represents a trend, especially when combined with a \textit{game-interaction} approach.

\textbf{Situation adaptivity in museums}

This section surveys those projects that deployed user modeling. User modeling is the use of artificial intelligence software techniques to construct conceptual models for users that enable predictions and responsiveness to user interactions. The projects we review include: \textit{Hippie}, the \textit{Museum Wearable}, \textit{ec(h)o} and \textit{HyperAudio}. We have also analyzed a system that is currently web-based, known as \textit{CHIP}. Each system offers the museum visitor some form of personalized content such as audio or video clips, text and images, and each uses aspects of visitor interaction with the system to achieve customization. The details of each project are discussed first, where we provided a detailed account of each project's approach to adaptivity. The overview will be followed by an analytical discussion that includes a cross-matrix used to frame the current research in the field.

Earlier we introduced the \textit{HyperAudio} system from 1999 (Petrilli and Not 2005). Audio content generation was based on a \textit{visit model} that looked at the grouping each user was in, whether it was a first or repeat visit, how long they planned to stay and what kind of interaction type they preferred. Underlying the model was a simple Boolean variable indicating interest in each item, based on whether the user remained in front of an object after the relevant presentation finished or whether they turned off the presentation.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.png}
\caption{Cross-matrix of visit flow and visit type.}
\end{figure}
Another early attempt to develop a context-sensitive, adaptive museum guide was the *Hippie* project (Oppermann and Specht 1999, 2000). *Hippie* contained models for three distinct components: (1) a static domain model (objects to be presented and processed about), (2) a static space model (physical space) and (3) a dynamic user model. In constructing the user model, the authors assumed that visitors had a ‘stable interest trait structure’, but that environmental and contextual factors played a role in the actual activation of that structure at any given moment. The goal of their user model was ‘to predict the information needs of a user in a given episode of a visit’; thus the model made inferences regarding the next exhibit to visit and the next piece of information to present. Interest in particular exhibits was inferred by time spent there as well as the number of information items selected. They also suggest using ‘differentiated navigation behavior’ to indicate interest, i.e. viewing artworks from multiple locations rather than only one indicates more interest. Content was classified according to type, and used to infer what kind of information the visitor was interested in. The HIPS project also applied Veron and Levasseur’s 1999 ethnography work in museums, which identified four physical movement patterns of visitors: ant, butterfly, fish, grasshopper (Marti et al. 1999).

1. Ants proceed methodically through the entire museum space, looking at everything;
2. Butterflies fly around from work to work based on interest level;
3. Fish swim through the space quickly, glancing at things in a cursory fashion; and
4. Grasshoppers also bounce around, but with less of a defined idea of what they are interested in.

Visitor type and information preference was proposed as a way to assemble appropriate length audio clips for each individual.

While the *Hippie* system asked the users what areas they wish to explore, other approaches did not give this level of control to the user, such as the *Museum Wearable*, which assembles and delivers personalized content to the user without explicit user interaction (Sparacino 2002). A Bayesian network is used for user classification and decisions about content assembly. Sparacino uses a ternary classification of visitors as busy (cruiise through everything quickly), selective (skip some things, spend long on others) and greedy (spend a long time on everything). Data were collected and used to set prior probabilities of a model based on path and pause duration patterns. Upon approaching a new exhibit, a sequence of content clips was assembled dynamically using user type to determine how long it should be.

The *ec(h)o* project provided visitors with a choice of audio clips throughout their museum visit (Hatala and Wakkary 2005). The *ec(h)o* user model captured two basic kinds of information about the museum visitor using the guide: interaction history and user behavior information. Interaction history recorded movement through the museum space along with the sound objects selected for listening, whereas user behavior information kept track of the user’s interests. Interest was set by the user’s explicit choice from a number of different concepts at the beginning of the museum tour and then updated based on the exhibits they paused at during their exploration of the museum space and the choices they made regarding which audio content to listen to. If the user showed special interest in a specific concept or two, those concepts would become highly weighted in the model and content related to them would be offered frequently. However, to avoid the complete stereotyping of a visitor, a ‘variety’ element was also included in the recommendation algorithm, so there was the possibility of sound objects
being offered that would allow the user to move away from their heavily weighted interests.

The most recent project to use adaptability in museums is the CHIP project, which is currently web-based only. This project allows a user to generate a personal profile via an online rating system for artwork (Aroyo 2007). Images in the system are annotated semantically for artist, period, style, visual content and themes. As users rate them with one to five stars, the user model is refined to reflect their preferences more clearly. The classification ratings (e.g., four stars for Impressionism) are inferred from the explicit work ratings, and can be viewed at any time. If the user disagrees with the system's calculation of the classification rating, they can adjust it manually. The user can also ask to see 'recommended works' which pulls out unrated artwork that the model believes the user would rate highly (four to five stars). Users can query as to why the system recommended this work, and will be provided with the individual classification ratings that underlie the suggestion. The next step in the project is to integrate this with technology in the physical museum, so that users can access their profile in the exhibit.

The PEACH project also allows the user to participate actively in the construction of her user-model through a widget (Stock et al. 2007). The 'like-o-meter' widget allows the visitor of the system to state whether she likes or dislikes a given museum artifact, which affects the amount of information the system provides for subsequent related artifacts. The authors suggest that their widget was clearly understood and enjoyed by visitors.

Our aim here is to determine what is common practice in user-modeling, what is more uncommon and experimental and what has not yet been attempted. One of the first aspects to consider is what kind of data the systems gather and use as input to the model. There are two basic categories of input commonly used: physical data and content-based data. Physical data include such things as knowledge of the visitor's current location (Hippie, Museum Wearable, ech(o) and HyperAudio), of their overall path through the museum space (Hippie, Museum Wearable), and of their stop duration at each specific exhibit (Hippie, Museum Wearable, HyperAudio). Content-based information includes knowledge of what content the visitors have listened to or selected (Hippie, Museum Wearable, ech(o), CHIP, HyperAudio) and how they rate that content (CHIP, PEACH). From this basic input, all the systems extrapolate and make inferences about categories into which the user falls or characteristics the user might have, based on the observed behavior. Every system infers user interest in an artifact/exhibit, based usually on their movement towards it, presence near it or selection of it in some way. The systems all have the capacity to detect user interest in broader themes or topics, as signaled by the selection of multiple items, which have been annotated similarly. Some of the systems also include the ability to detect interest in a specific information type, as signaled by selection of specific kinds of content (the Museum Wearable, Hippie and HyperAudio). For example, a user might ask repeatedly for artist biography content, or for details on how an artifact was constructed. The final type of inference found in these systems is the classification of visitors as a certain 'user type'. This includes the fish/ant/butterfly/grasshopper path classification from the Hippie project and the greedy/busy/selective user type from the Museum Wearable, both of which used movement patterns to sort users.

One important distinction among the ways the projects handle system input is the degree of control that the user has over the cues that the system is picking up on. Another way to think about this would be in terms of the opacity level present in the interaction. With some of the systems, it is easy for users to tell what kinds of information the system is picking up on, such as their presence next to an object or their deliberate selections within
the system; we describe this approach as *transparent interaction*. Other systems have much more opaque interactions that are being picked up on, such as the path through the space or their patterns of stop duration and movement. In these cases, the user of the system may never guess that such elements are being used as input for the user model, and thus it is beyond their control to affect what the system is doing in response to them; we describe this approach as *opaque interaction*. It is possible that this opacity will yield a more fluid, intuitive interaction with the system, but it is equally possible that it could result in a frustrating experience where the user does not understand why the system is responding the way it is.

With the collected and inferred data as input, the next concern is what the system presents back to the visitor as output. Several of the guides offer a set of further audio and/or video content that the user can select from (Hippie, ec(h)o). The *Museum Wearable* also presents audio/video content, but does not allow the user to select what they view; instead, the model automatically assembles a tailored presentation for the user. *HyperAudio* combines these approaches by assembling a tailor-made audio presentation and also allowing for individual exploration of topics via the handheld device. A third option for output is to generate recommendations of other pieces the visitor might like based on the model's understanding of the visitor (*CHIP*). There is a distinction here between systems that have *static content* and those which have *personalized content*. The *Museum Wearable* and *HyperAudio* actually tailor the content itself to the individual user. In the other systems, the content remains the same, but the order in which it is presented or the options available at any moment are tailored to the individual user.

In summary, Figure 6 shows a cross-matrix of this static-personalized content dimension with the opaque-transparent interaction dimension. As can be seen from the diagram, the majority of research thus far with user modeling in museum applications has involved static content and transparent interaction. There have been a couple of different ways in which the input/inference data affects the output. All the guides have the simple correlation of high user interest leading to more content related to that user interest. *Hippie* and the *Museum Wearable* also use user-type classifications to affect the duration of the content offered to the user. Room still exists for a range of creative thinking in this area, especially with regard to how collected data can influence the system interaction and output.

![Figure 6. Cross-matrix of interaction and content types.](image)

---

16
Discussion

Through gathering research in this study, it is possible to understand the variety of approaches taken to augment the museum space. When faced with group-based activities, the PDA may further distract the visitor from his/her companions. The group-based applications discussed in this study revealed a shift towards a game-interaction approach—while all used PDA devices to guide the visitors through the museum. Again, we believe TUIs may have a greater role and affect in this regard. Research within group-based tour applications is fairly limited, and has become the focus of two newer studies (CoCicero and ARCHIE). The game-interaction approach seems not only to affect the way in which individuals browse through the museum, but may also affect the manner in which learning occurs. The game-interaction approaches discussed here encourage the visitors to find artifacts which match specific criteria in a quest-like fashion. This type of activity might be useful in teaching visitors what singular artifacts are, but may inhibit the communication of the contextual significance that often surrounds artifacts in museums. The earliest of systems to adopt a group-based approach, Sotto Voce, note that co-present interactions should be supported. The CoCicero project takes the concept further by introducing other group-based behaviors to be supported, such as following and checking-in. As our focus is on group-based interaction, using the findings of such projects may prove helpful in designing an application that fosters group visits by families to museums, such as our own Kurio project (see Background section).

With regard to a user-modeling approach, the majority of the research lies within supporting transparent interaction and static content, although it is difficult to state whether one method is superior to another at this time. It could be argued that a push towards even more transparent input is desirable. The CHIP project is the most transparent of all the interactions, allowing the visitor to adjust the user model manually and gain feedback from the system as to why certain recommendations have been made for them. No such hybrid adaptive and adaptable system has been implemented within the user space itself. Alternatively, it could be fruitful to explore the area on the extreme other side of the transparent-opaque interaction continuum, by creating a system where users are unable or unlikely to guess what aspects of their behavior are being used within the system. If conducted properly this could yield a very immersive, fluid-feeling interaction. Projects that have attempted to classify visiting patterns have done so within the individual-based tour context, and there is a lack of research on museum visitor patterns for groups—a potential area of interest for us to address.

There is also little research on adapting to the group conditions of the visitor, e.g. taking into consideration the fact that person A is a mother visiting with three children, while person B is a senior citizen visiting with friends. Modeling users on the level of group dynamics is a new and growing area within user modeling, one that we have explored (Wakkary et al. 2005) and hope to bring to the museum domain. In terms of content, there is definitely space to explore in the realm of adaptive content—taking the personalization of the museum beyond simple recommendations and on-demand information access and moving it into the realm of individualized presentations.

We began the paper with a discussion on tangibility as a mediating bridge between computation and context that mitigates the lack of support for social interaction. Tangibility may preclude some of the directions in user modeling, such as transparent interaction of systems such as CHIP, which is web-based after all. However, tangibility could be seen as an enabler in game-interaction approaches and group interaction, given its
relative playfulness in comparison to graphical interfaces, the ease of sharing tangible devices and the leveraging of past patterns of social interactions mediated by objects.

**Future trajectories in the research of interactive museum guides**

It is inherently risky to predict what may develop in the future. Given the analysis in this paper, however, there is evidence for the following future trajectories resulting from a balance of assessed needs of a museum visit and the capacity to design and implement technology in the case of interactive museum guides.

**Embodied interaction**

As discussed previously, embodied interaction is a computing paradigm in which tangible computing and social computing become intertwined and incorporated into our everyday environment and activities (Dourish 2001). This notion fits the majority of museum settings very well, given the sociality of museums and the reliance on curious and playful interaction with physical objects and interactives. An additional contributing factor to the fit with most museums may be a heightened awareness on the part of visitors of the physicality of things due to the presence of remarkable artifacts on display. We discussed how, in *ec(h)o*, tangibility afforded a greater awareness of the surroundings of the museum, as visitors were not distracted by a handheld visual display and visitors' actions were perceived to be part of typical interaction with the environment through actions such as walking and gesturing. A goal within this trajectory of research will be to bring tangibility qualities together with the social computing aspects of works such as *Sotto Voce* and *CoCicero*, in which displays, communication and activities are shared through PDAs. Wireless technology underlies the advent of interactive museum guides and, while the networking capabilities were utilized initially for distribution of information, there is the needed shift to social communication and coordination, not unlike the current online social networking.

**Gameplay**

Gameplay refers to the quality of the mechanics or play in a game. We observed that the idea of gameplay can be applied to interaction with museum guides, often in the service of learning. Games and play are used in the museum as an interaction convention that is a general set of expectations that people accept for a museum visit as a means of exploring and learning about the exhibition. Gameplay, or game interaction, can be used as a more entertaining way to deliver information but we observed in *ec(h)o*, Co-Cicero and *ARCHIE* that gameplay supports learning in ways that are equally about constructivist exploration and discovery. Additionally, the latter two utilized games as a coordinating and collaboration function for families and groups.

**Transparent or opaque interaction**

In our discussion of user models and adaptivity, we made the important distinction between transparent and opaque interaction. Transparent interaction occurs if it is understandable to a visitor how their actions and interface interactions influence the user model, as in the case of user ratings, selection of content or navigation options or explicit request of user preferences. Opaque interactions occur when visitor actions inform the user
model in ways that are not apparent to the user, such as movement through the exhibit, duration of pauses and gaze tracking. The degrees of transparency in collection of user model data directly affect the quality and nature of the visitor experience. In our view, there are almost equal trade-offs in either direction, as we discussed, and research will probably continue on both fronts. A probable practical outcome, and current practice in some cases, is the combination of both approaches within a single system.

**Personal digital assistants, graphical user interfaces and hybrid systems**

Past interactive museum guides adopted a mobile computing approach utilizing graphical user interfaces (GUI) on hand-held PDAs as the device and interface modality. Traditionally, the PDA/GUI method has served the information rich experiences well. We believe, however, that the trends we have been discussing may alter this approach. First, the emergence of tangibility offers an approach different from mobile computing. The user interface demands can be distributed across different tangible devices that are each simpler (without a GUI), can be designed to work together and are customized to the particular setting and activity. We believe tangibility to be a group-centric approach with more reliance on context than mobile computing which, to date, has been individual-centric and virtual over contextual. We believe that this emergent tangibility model will support group and family interaction more effectively in museums. Having said that, the trends of group and game interaction require coordination and collaboration. Here, a PDA or shared display with a GUI can serve a key role in coordinating different members of the group and family during a museum visit. Additionally, a GUI provides a virtual collaborative space that in game interaction can represent the shared state of the game as it and provide a collaborative space to communicate and help each other.\(^1\) Hence, we see a shift in the use of PDA/GUIs from information delivery to providing shared virtual space for coordination and collaboration. Considering all this, we came to the conclusion that hybrid systems best support the type of embodied interaction of tangible and social computing that we have been discussing. A hybrid system might include tangible devices and shared displays in the form of a PDA and or tabletop computer. This is exactly the approach we have taken in Kurio, which we discussed briefly earlier (see Background section and Figure 1).

**Benefits of re-situating interactive museum guides**

Many research and technology areas evolve and change in waves and we believe that interactive museum guides are no different. Further, for this paper we aimed to illustrate fundamental shifts in re-thinking interactive museum guides to address tangibility, interactivity and adaptivity. This re-situating of interactive museum guides holds benefits for museum administrators considering interactive technologies for their museum, or researchers and developers looking anew or again at the interactive museum guide domain:

- **Closing the social gap.** In our analysis there is a growing interest in families and groups, which begin to address the gap that in reality families and groups make up the majority of museum visitors while in the past interactive museum guides have been designed for individuals.

- **Naturalizing technology.** Technology in the museum adapts to the everyday environment of the museum, such as physical interactives, exhibit displays and artifacts in the proposed shift to tangibility. As Ishii suggests, it is time for technology to 'rejoin the richness of the physical world' (Ishii and Ullmer 1997),
which allows us to consider the larger design problem of the whole museum experience rather than focusing on the novelty of new technology.

- **Shift to exploration and discovery.** Interactive museum guides in the past focused on information access and richness. The move towards embodied interaction and game interactions creates the opportunity to design learning activities with interactive technology that are based on personal exploration and discovery, rather than information retrieval and retention.

**Conclusion**

As shown in this paper, there are currently several approaches for adaptive museum guides under exploration. Through analyzing current contributions through the three key areas of interest – tangibility, interactivity, adaptivity – we were able to develop a comprehensive outline that provides a theoretical grounding for our future research in a family and social-based museum guide. The majority of the current literature focuses on the interactions of a single visitor, but through analyzing trends in the museum guide research, the focus appears to be shifting. We discussed four interrelated trajectories for interactive museum guide research including: embodied interaction, gameplay, transparent and opaque interaction and the new role of PDA/GUIs. We concluded with the benefits of the shifts in research and practice, including addressing the presence of family and group visitors, integrating technology in a museum setting more effectively and opportunities for designing learning activities based on exploration and discovery.

**Note**

1. In discussing PDAs, further consideration should be given to impact of the changes in industry and use of smartphones. The growing prevalence and comfort with smartphones will have an impact on the role PDAs in museums; in particular, visitors will have expectations of using their own mobile devices through downloadable software or online applications.

**References**


Jiang, Y. 2008. Exploring personality type for composition-based group user modeling in an embodied interaction system. Surrey, BC: Simon Fraser University. 102 School of Interactive Arts and Technology.


Situated play in a tangible interface and adaptive audio museum guide

Ron Wakkary · Marek Hatala

Received: 30 August 2005 / Accepted: 24 April 2006 / Published online: 4 November 2006
© Springer-Verlag London Limited 2006

Abstract This paper explores the design issues of situated play within a museum through the study of a museum guide prototype that integrates a tangible interface, audio display, and adaptive modeling. We discuss our use of design ethnography in order to situate our interaction and to investigate the liminal and engagement qualities of a museum visit. The paper provides an overview of our case study and analysis of our user evaluation. We discuss the implications including degrees of balance in the experience design of play in interaction; the challenge in developing a discovery-based information model, and the need for a better understanding of the contextual aspects of tangible user interfaces (TUIs). We conclude that learning effectiveness and functionality can be balanced productively with playful interaction through an adaptive audio and TUI if designers balance the engagement between play and the environment, and the space between imagination and interpretation that links the audio content to the artifacts.

1 Introduction

In our adult lives play is an experience set apart from our everyday activities: Huizinga referred to play as invoking a magic circle, a liminal space for games [1]; Carse describes deep play as a profound level of ritualized engagement causing reflection on everyday experiences [2]; and psychologist Csikszentmihalyi has described flow as a high level of engagement, risk and challenge found in play and ritualized in sport [3]. Do we play in museums? Art historian Carol Duncan sees the museum as a “stage” that encourages visitors to perform rituals that are not part of their daily life [4]. Anthropologist Genevieve Bell extends this notion of extraordinary ritualized play together with learning. She describes museums as different cultural ecologies in which the museum visit has the qualities of liminality (a space and time set apart from everyday life) and engagement (where visitors interact to both learn and play) [5].

Guided by the notion of play in a museum experience we have considered playfulness equally with functionality and learning in the design of an adaptive museum guide. Our approach includes a tangible user interface (TUI) for its inherent playfulness and poetic simplicity, spatial audio display for the diversity of human voice and its imaginative qualities, and an integrated user modeling technique combined with semantic technologies that support exploration and discovery. We understood our interface as playful action along the lines of aesthetic interaction. By this we do not mean the type of structured play that is found in a software game on a mobile device, rather we refer to the less structured and open play that is always possible and often can be subtle and implicit like toying with a ball.

Furthermore, we aimed for our design to be situated within the setting we were designing that is to design an interface and interaction that felt a part of the museum.
Toward this end we adopted the idea of museums as ecology informed by Bell’s cultural ecologies and Nardi’s and O’Day’s information ecology [8, 42]. Bell sees museum visits as determined by the ecological interplay of space, people and design. Nardi and O’Day view organizations as organic relationships among people, practices, technology, values and locale. We utilized ecologies to situate our design and frame it ethnographically and theoretically. This approach led to us being inspired by simple physical displays and puzzles we observed in our ethnographic sessions. These observations encouraged the playful tangible object and use of puzzles in our audio content. We were also motivated by the storytelling of the museum staff and researchers that was often humorous as well as informative. We found the ecologies analytically essential in understanding how we were situating play, our interaction, and technology within the museum.

We provide here an account of the reasons and rationale of our design concept and the approach of our case-study known as ec(h)o. In the paper, we discuss related research to our case study followed by a discussion of our design motivations, our ecology informed design ethnography and resulting design implications. We then describe the case study, which we installed and tested at Canadian Museum of Nature in Ottawa, and analyse the TUI and aesthetic interaction aspects of our interface. We provide an overview and analysis of our evaluation and a discussion of lessons learned including several issues relevant to ubiquitous computing: the experience design of play in interaction; the balance in developing information models; and the need for a better understanding of the contextual aspects of TUIs. We conclude that based on our results of our pilot study, learning effectiveness and functionality can be balanced productively with playful interaction through an adaptive audio and TUI if designers balance the engagement between play and the environment, and the space between imagination and interpretation that links the audio content to the artifacts.

2 Relevant research

Bedersen [6] was among the first to develop an electronic museum guide prototype supporting visitor-driven interaction by utilizing portable mini-disc players and an infra-red system to allow museum visitors to explore at their own pace and sequence. Today interactive museum guides have reached significantly higher level of functionality including visitor-driven interaction, media rich delivery, context-awareness and adaptivity. We aim in our prototype system, ec(h)o to maintain a standard level of functionality with the exception of media rich delivery. While we sacrificed the ability to deliver diverse types of media we gained the opportunity to move away from a graphical user interface (GUI) and the personal digital assistant (PDA) in the hopes of creating a more playful and aesthetic interaction through a physical and embodied interface. We were also able to simplify our content approach and focus on the potential of audio to create imaginative and ludic possibilities. However, we do see possible future implementations that include images, video and dynamic text information within our TUI approach through the use of distributed visual displays within the exhibition spaces.

Previous work most relevant to our case study includes museum guide systems that utilize an adaptive approach, GUI and PDA interfaces in museum guides, and a discussion of work outside of the museum domain that utilizes audio interfaces in ubiquitous and mobile computing contexts. Equally important to our discussion are the ludic qualities of TUIs, and related ideas of aesthetics and play in interaction.

2.1 Adaptive museum guide systems and audio display

Adaptation and personalization approaches have been successfully applied to museums in the context of the World Wide Web [7, 8] and in handheld museum guides. ec(h)o shares many adaptive characteristics with the systems of HyperAudio, HIPS and Hippie [9–11]. Similar to ec(h)o, the systems respond to user’s location and explicit user actions through the interface. HyperAudio uses a static user model set by a questionnaire completed by the visitor at start-up time and HIPS and Hippie can infer the user model dynamically from the interaction but they treat user interests as static. All systems adapt content to the user model, location and interaction history. Among the main differences with ec(h)o is that these systems depend on a PDA GUI, ec(h)o uses audio display as the only delivery channel and a tangible object as an input device. Another difference lies in how the system generates response: ec(h)o uses inference at the level of semantic descriptions of independent audio objects and exhibits ec(h)o extends the work of the Alfaro et al. [12] by building a rich model of the concepts represented by the audio objects while HyperAudio and HIPS use partly pre-configured annotated multimedia data [13], and Hippie uses a simpler domain model. The last main difference is that ec(h)o treats user interests as dynamic, we look to evolving interests as a...
measure of sustainable interaction and go one step further by ensuring a high degree of diversity of interests is available. These differences exist in order to create an experience of discovery in which visitor's are given the latitude to explore new and previously unconsidered related topics of interests.

Prior to the evolution of adaptive and user modeling approaches in museum guide systems, there had been a strong trajectory of use of the PDA GUI. Typically, hypertext is combined with images, video and audio [14-17]. A good example of this is the MEG system [18]. It was created for the Experience Music Project in Seattle. It allows visitors 20 h of audio and video on demand. Visitors make their selections either by use of the keyboard within the PDA device or by pointing the device at transmitters located adjacent to artifacts. For further interaction with the information, visitors are dependent on the GUI, which is a typical browser and hierarchical menu format. There are clear functionality advantages in the PDA GUI approach including the organization and accessibility of large amounts of data, a user interface that is familiar since it resembles a personal computer (PC), multimodal input from pointing to text to voice, and multimedia delivery. Yet researchers such as Hans Tap have identified a tension in relationships between computer systems that rely on desktop computers as the basis for interaction and the artifacts, physical environment and everyday activities of most people [19]. He uses the term desktop gravitation to describe how desktop computers force people to move to the desk to carry out their work. We ask the question whether we should carry around our desks in order to experience such things as museums—in what might be described as a world-behind-a-desk approach to mobile computing? Furthermore, a PDA is essentially a productivity tool for business, not a device that lends itself easily to playful interaction.

Aoki and Woodruff [14] have argued that in interactive guidebooks, designers are challenged to find the balance between burdening the visitor with the functions of selection, information management and contextualization. The PDA GUI approach comes at a cognitive and experiential cost. It requires the full visual attention of the visitor such that it becomes a competing element with the physical environment rather than a valued and integrated addition to that environment. Museum systems have mostly maintained the PDA GUI approach despite the shifts in other domains to other approaches that better address the experience design issues most prominent in social, cultural and leisure activities. The play constraints of these devices are too great for the level of interaction that goes beyond playing a software game on a mobile device. For example, in the area of games and ubiquitous computing, Björk and his colleagues have identified the need to develop past end-user devices such as mobile phones, PDAs and game consoles [20]. They argue that we need to better understand how “computational services” augment games situated in real environments. The same can be said for museum visits.

Non-visual interfaces, particularly audio display interfaces have been shown to be effective in improving interaction and integration within existing physical contexts. For example, Brewster and Pirhonen [21, 22] have explored the combination of gesture and audio display that allows for complicated interaction with mobile devices while people are in motion. The Audio Aura project [23] explores how to better connect human activity in the physical world with virtual information through use of audio display. Audio is seen as an immersive display that can enrich the physical world and human activity while being more integrated with the surrounding environment. In addition, audio tends to create interpretive space or room for imagination as many have claimed radio affords over television. In the HIPS project, different voices and delivery styles were used to create an “empathetic effect” between the user and the artifacts they engaged [24]. We have adopted a similar approach to our use of audio content. Audio augmented reality systems combined with TUIs often create very playful and resonant interaction experiences [25]. In fact, the distinction between augmented reality and TUIs can be blurry indeed [26].

2.2 The poetics and play of TUIs

Tangible user interfaces like no other user interface concept is inherently playful, imaginative and poetic. In addition, the concept has immediacy due to its physicality. Ishii’s and Ullmer’s notion of coupling bits and atoms was informed by earlier work in graspable interfaces [27] and real-world interface props [28]. cc(h)’o’s TUI draws on this notion by coupling an everyday and graspable object, a wooden cube with digital navigation and information. Ishii was inspired by the aesthetics and rich affordances of scientific instruments [26] and the transparency of a well-worn ping-pong paddle [29]. Simple physical display devices and wooden puzzles at the natural history museum where we conducted ethnography sessions inspired us as well.

In 1992, Bishop’s Marble Answering Machine [30] was an early embodiment of the immediate and playful qualities of TUIs. The prototype uses marbles to represent messages on the machine. A person replays the
message by picking up the marble and placing it in an indentation in the machine. Jeremijenko’s Live Wire is a strikingly minimal and whimsically simple demonstration of digital bits transformed into physical atoms [31]. Jeremijenko dangled a plastic wire from a motor attached to the ceiling. The motor accelerates or decelerates based on traffic across the Ethernet networking. Ishii’s PingPongPlus [29] explores the intertwining of athletic play with imaginative play. The ping-pong table becomes an interactive surface, the ball movement is tracked and projections on the table of water ripples, moving spots, and schools of fish among other images react to where the ball hits the table. ambientROOM [26] is a collection of tangible interfaces integrated in an office environment in order to enhance and exploit the user’s peripheral awareness, for example a physical icon moves and rotates on a desk mirroring the actions of a nearby hamster. More recent work, such as Andersen’s Clownsparkles [32], engage children explicitly in exploratory play and emergent learning through sensor-augmented everyday objects (dresses, hats, costumes, and purses) and audio display. The work explores the role of TUIs in an open-ended game of children’s dress-up. Andersen’s work reveals how theatrical settings provide an emotional framework that scaffolds the qualitative experience of the interaction. While ec(h)o is more constrained in its play, the everyday wooden cube provides such a scaffold to a physically playful experience of interaction.

2.3 Aesthetics of interaction

Researchers in human–computer interaction (HCI) have recently explored beyond the goals of usefulness and usability to include enjoyment [33], emotions [34, 35], ambiguity [36], and ludic design [37]. Nowhere is this need more evident than in the richly interpretive and social environments of museums [24, 38]. Our emphasis is on the qualities of interaction that result in play that facilitates discovery. While we address this on an informational level in regard to our use of audio content and information retrieval, we aimed to equally explore the embodied and situated aspects of interaction or aesthetic interaction as expressed by Djajadiningrat [39] and Petersen [40]. Djajadiningrat argues for a “perceptual-motor-centered” approach to tangible interfaces [39]. He is less sympathetic toward the cognitive view of interaction in what he terms the “semantic approach” where objects communicate action through metaphor. Rather, he argues for a “direct approach” for its “sensory richness and action-potential” of the objects to carry meaning through interaction. He describes this notion of meaning in interaction as aesthetics of interaction whereby the “beauty of interaction” as opposed to the beauty of the artifact or interface, tempt the user to engage as well as “persevere” in their engagement [39]. He describes three factors as having a role in aesthetic interaction: the interaction pattern of timing, rhythm, and flow between the user and the object; the richness of motor actions found in the potential space of actions and skill development; and freedom of interaction in which a myriad of interaction paths coexist.

Petersen et al. [40] description of aesthetic interaction shares the embodied aspects described above as well as the sense of aesthetic potential that is realized through the action or engagement. They bring to the concept the philosophical view of Pragmatism that aims to situate aesthetic interaction within everyday experiences and the surrounding environment. For example, Petersen developed a playful interaction approach as part of the WorkSPACE project [41] utilizing a ball that is thrown against a floor projection of documents and work materials as a way of manipulating and exploring the information. Inherent to the ball are kinesthetic challenges, affordances and the situated relationship with the environment. These aspects are realized in action with the object. The aim of the interaction approach is to create new views of the work material through the playful actions of aiming, throwing and bouncing.

3 Design motivations

We were strongly influenced by the awareness of museums as complex and dynamic spaces. Vom Lehn et al. [42] describe museum experiences as multivariate that is they cannot be assessed by a single factor such as exhibit design, signage, or time spent in front of an artifact. Instead, the museum experience is subject to multiple influences and results in multiple outcomes. Given this understanding, we endeavored to consider how our design both intervenes in and integrates with the complex museum experience. The ecological models of cultural ecologies and information ecologies provided us with frameworks for contextual analysis. This approach allowed us to look further into the design process past the interface for guidance into how our design decisions were integral to the ecology or ecology inhabitants, thus supporting us in developing more appropriate design responses. We provide here a summary of the ecological concepts and a discussion of their use in our ethnographic sessions. For further
discussion of the role of ecologies in museums we refer readers to [43].

3.1 Museums as ecologies

Bell sees the museum visit as a ritual determined by space, people and design [5]. She decomposes the visiting ritual into three observational categories: space, visitors, and interactions and rituals. Different types of museums have different ecologies, for example Bell describes different attributes in each of the observational categories between art museums and science museums. These ecologies are seen to be distinct and supportive of very different kinds of museum visits. Bell also describes interaction concepts that are common to all museum ecologies. We have drawn on two of these concepts in developing our approach, liminality and engagement:

• **Liminality** defines museums as places that embody an experience apart from everyday life. Positive museum experiences are transformative, spiritual, and even moving. A museum visitor should be inclined to pause and reflect, thus liminality can be seen to permit a deeper engagement.

• **Engagement** is a key concept for museums as people go to museums to learn, however this engagement is often packaged in an entertaining way; museums are a balance between learning and entertainment spaces.

Nardi and O’Day draw on activity theory [44, 45] and field studies to develop their concept of information ecologies. The concept they describe strives for a more systematic view of organizations based on the relationships among people, practices, technology, values and locale. For example, a library is an ecology for accessing information. It is a space with books, magazines, tapes, films, computers, databases and librarians organically organized to find information. Nardi and O’Day utilize the concept of ecology in order to depict the complex relationship among elements and influences of which technology is only one part. Constituent elements of information ecologies include a system, diversity, co-evolution, locality, and keystone species. Two of these elements were essential in supporting our design:

• **Locality** can be described as participants within the ecology giving identity and a place for things. For example, the habitation of technology provides us with a set of relationships within the ecology, to whom a machine belongs determines the family of relationships connected to the technology. In addition, we all have special knowledge about our own local ecologies that is inaccessible to anyone outside thus giving us local influence on change.

• **Keystone species** are present in healthy ecologies; their presence is critical to the survival of the ecology itself. Often such species take the role of mediators who bridge institutional boundaries and translate across disciplines. For example, introduction of new technologies in an ecology is often reliant on mediators who shape tools to fit local circumstances.

3.2 Design implications of our design ethnography

Our observations that fall within Bell’s categorization of interaction and ritual emphasized that our system should be open to multiple forms of input such as movement and physical interaction with the displays, and responsive to different learning styles. In many respects, our prototype became a virtual extension of the exhibition space and acted as an augmentation to the physical interactives and other learning materials.

The displays and installations revealed diverse forms of interaction: microscopes with adjustable slide wheels that could be turned to explore different specimens; wooden puzzles which, once completed, would fall apart at the pull of a handle, creating a loud crashing sound that captured the attention of others (see Fig. 1); a collecting game called The Rat Pack Challenge which tasked visitors to search the room and discern collectable artifacts from non-collectable ones; discovery drawers filled with objects like fossils, fur pelts, and minerals which visitors could touch and inspect at close range (see Fig. 2); push button audio and video installations; scale models and artist recreations of

![Fig. 1 A wooden puzzle interactive in the Finders Keepers exhibition](image-url)
A discovery (Jr;iwcr in ihc
"lttdcr Kvepen
exhibition
dinosaurs that people could walk up to and touch;
terrariums and aquariums filled with living specimens;
magazines, coloring books, and a small library of nat­
ural history artifacts that were lent to students.

Bell notes that an attribute of science museum ecol­
gogies is to support the fact that people learn in a variety
of ways. Alternative approaches to learning turned up
throughout our obscrvatit>ons. such as the interactive
puzzles, quizzes, and games that require visitors to ex­
plore and think about the artifacts being displayed.

The design implication here is that the observed
activities support a highly tactile approach that in­
cludes holding, manipulating and being highly inter¬
active with your hands. A TUI would situate itself well
among these puzzles, games and physical displays.
Another design implication is the use of puzzles and
riddles as modes of interaction and content delivery.
Visitors are not spoon-fed factual information in the
form of didactics, rather they engage in play and dis­
covery to learn about the artifacts and the broader
concepts that tie the artifacts together thematically.

Nardi's and O'day's information ecology also guided
design decisions. For example, the stories and infor­
mation we heard in our interactions with staff and
researchers at the museum were examples of the eco­
logy concepts, locality and keystone species. This led to a
novel approach to content design and development we
have described in detail in another paper [46]. We ob­
served numerous informal yet engaging delivery of
specialized knowledge on behalf of the museum
researchers. The majority of these types of exchanges
happened as we toured the collections and storage
facility. Stories connected to artifacts ranged from
anecdotes on where the artifact was found and how cold
it was at the time or how difficult the terrain was, stories
of the difficulties of mold-making on site or humorous
tales of transportation and objects temporarily getting
lost, to what the objects tell us or how their meaning has
changed. Often these were first hand accounts and
discussed in the most informal and wide-ranging man­
ner. Factual or thesis driven accounts of artifacts were
mixed with anecdotal and humorous tales related to the
discovery, processing or research of the actual artifact.
This experience deeply struck us since our shared per­
ception of the public exhibition display space was quite
the opposite. Not unlike many exhibitions, the artifacts
and contextualizing information appeared static and
lifeless, the puzzles and games notwithstanding. In
locality terms, it was evident to us that once the artifacts
were connected to people, the understanding of these
artifacts became deeply connected to all aspects of the
ecology and came out in the form of storytelling that
covered activities related to the artifact, conservation,
storage, research and display technologies, meaning
and values associated with the artifacts.

A resulting design outcome was to bring this degree
of liveliness to the artifacts on display. We aimed to
model our information delivery and audio experience
on the informal storytelling we had experienced. We
aimed to create a virtual cocktail party of natural his­
tory scientists that accompanied the visitor through the
museum.

For our purposes, both ecological frameworks served
our goals despite their strong differences. Bell's cultural
ecologies formally linked different actions and attri­
butes of the museum visitor into a coherent description.
As a descriptive tool it validated our assumptions and
provided a clearer link between what we observed and
the design implications. It was therefore generative
much like Nardi and O'Day's information ecologies
framework. Both guided us in specific design decisions,
mainly the high degree of physical interaction that
suggested a TUI; the wide use of puzzles, riddles and
games as modes of learning which led to our use of a
riddle-like approach to our audio content; and the
localized and informal storytelling on behalf of the
museum staff and researchers that inspired us to
structure our audio experience like a virtual cocktail
party. As we set out to approach an adaptive museum
guide from an experience design perspective, we ex­
plored situated play in the museum and uncovered
specific qualities of liminality and engagement rooted in
the museum within which we were designing.

4 Case study

The design motivations and ethnography findings led
us to a design that was minimal, playful and supported
exploration. Our approach includes a TUI for its inherent playfulness and poetic simplicity, spatial audio display for the potential diversity of human voice and its imaginative qualities, and an integrated user modeling technique combined with semantic technologies that supported exploration. Our aim is to improve the visitor engagement by considering playfulness equally with functionality and learning. We adopted what can be described as a rich and discovery-based approach to interaction. While arguably other interface approaches could have been utilized in conjunction with the integrated modeling technique, such as a simple pushbutton device for input or a mobile text display device for output, such a strategy would be incongruent with our experience design goals.

4.1 Visitor scenario

In order to better understand the system we developed, we describe below a typical visitor scenario. The scenario refers to an exhibition about the history and practice of collecting natural history artifacts:

Visitors to the Finders Keepers exhibition can use the ec(h)o system as an interactive guide to the exhibition. Visitors using ec(h)o begin by choosing three cards from a set of cards displayed on a table. Each card describes a concept of interest related to the exhibition. The cards include topics such as “aesthetics”, “parasities”, “scientific technique” and “diversity”. A visitor chooses the cards “collecting things”, “bigness”, and “fauna biology.” She gives the cards to an attendant who then gives the visitor a wooden cube that has three colored sides, a rounded bottom for resting on her palm and a wrist leash so the cube can hang from her wrist without her holding it. She is also given a pair of headphones connected to a small, light pouch to be slung over her shoulder. The pouch contains a wireless receiver for audio and a digital tag for position tracking.

Our visitor moves through the exhibition space. Her movement creates her own dynamic soundscape of ambient sounds. As she passes a collection of animal bones she hears sounds that suggest the animal’s habitat. The immersive ambient sounds provide an audio context for the collection of objects nearby.

As she comes closer to a display exhibiting several artifacts from an archaeological site of the Siglit people, the soundscape fades quietly and the visitor is presented with three audio prefaces in sequence. The first is heard on her left side in a female voice that is jokingly chastising: “Don’t chew on that bone!” This is followed by a brief pause and then a second preface is heard in the center in a young male voice that excitedly exclaims: “Talk about a varied diet!” Lastly, a third preface is heard on her right side in a matter-of-fact young female voice: “First dump...then organize.” The audio prefaces are like teasers that correspond to audio objects of greater informational depth.

The visitor chooses the audio preface on the left by holding up the wooden cube in her hand and rotating it to the left. This gesture selects and activates an audio object and she hears a chime confirming the selection. The audio object is linked to the audio preface of the scolding voice warning against chewing on a bone. The corresponding audio object delivered in the same female voice yet in a relaxed tone, is about the degree of tool making on the part of the Siglit people: “Artifact #13 speaks to the active tool making. Here you can actually see the marks from the knives where the bone has been cut. Other indicators include chew marks...experts are generally able to distinguish between rodent chew marks and carnivore chew marks.”

After listening to the audio object, the visitor is presented with a new and related audio preface on her left, and the same prefaces are heard again in the center and to her right. The audio prefaces and objects presented are selected by the system based on the visitor’s movements in the exhibition space, previous audio objects selected, and her current topic preferences.

4.2 Interaction design

Our interaction model relies on a turn-taking approach based on the metaphorical structure of a conversation. \(^1\) Turn taking allows us to structure the listening and selection actions of the visitors. Prefaces and telling let us design the audio object in two parts: prefaces act as multiple-choice indices for the more detailed telling of the audio object. Responses and disengagement provided a selection and silent function for the system. The TUI provided input for a response — our equivalent of a nod. No response from the visitor was interpreted as disengagement.

The audio objects are semantically tagged to a range of topics. At the beginning of each interaction cycle, three audio objects are selected based on ranking using several criteria such as current levels of user interest, location, interaction history, etc. The topics of objects are not explicit to the visitor; rather the content logic is kept in the background.

In regard to the design process, many of the design choices were made through a series of participatory

\(^1\) The idea of using conversation analysis concepts as a structural metaphor for non-speech interfaces is not unique in HCI, see for example: Norman M.A., and Thoma P.J., “Informing HCI design through conversational analysis.” International Journal, Man-Machine Studies (35) 1991, 235–250.
4.2.1 Tangible object

The tangible interface object is an asymmetrically shaped wooden cube with three adjacent colored sides. The visitor holds the cube out in front of them in order to make a selection. The visitor makes a selection by rotating the cube so that the selected colored side faces directly upward (see Fig. 3).

The cube was carefully designed to ensure proper orientation and ease of use. The “bottom” of the cube has a convex curve to fit comfortably in the palm of the hand and a wrist leash is attached to an adjacent side to the curved bottom suggesting the default position of being upright in the palm and at a specified orientation to the body (see Fig. 4). The leash allows visitors to dangle the cube, freeing the hand, when not in use. The opposite side of the bottom of the cube is colored and shows an icon denoting a pair of headphones with both channels active. The sides to the left and right are each uniquely colored and display icons showing active left and right channels of the headphones, respectively. The cube is made of balsa wood. It is therefore very light (approximately 100 g or 3.5 ounces) mitigating tiredness from carrying the object.

The input of the selection is done through video sensing. The ergonomic design of the cube and biomechanics of arm and wrist movement form a physical constraint that ensures that the selected cube face is almost always held up parallel to the camera lens above and so highly readable. We experienced no difficulties with this approach.

4.2.2 Audio display

The audio display has two components, a soundscape and paired prefaces and audio objects. The soundscape is discussed along with navigation in Sect. 4.2.3. In the latter component, we used a simple spatial audio structure in order to cognitively differentiate between objects. Switching between the stereo channels created localization: we used the left channel audio for the left, right channel audio for the right, and both channels for the center. It is an egocentric [22] spatial structure that allowed the three prefaces to be distinguishable and an underlying content categorization structure to exist. The spatialization was mapped to the tangible interface for selection. In addition, we provided simple chimes to confirm that a selection had been made.

The prefaces were written to create a sense of surprise, discovery and above all play, especially in contrast to the informational audio objects. In order to create this sense we utilized diverse forms of puns, riddles and word play, for example:

- **Ambiguous word play**: “Sea urchins for sand dollars” (preface); “Other then the morphology, the sea urchin and the sand dollar are very similar species” (abridged audio object);
- **Simple pun**: “It’s like putting your foot in your mouth” (preface); “The word gastropod comes from two different roots; gastro for stomach, and pod for foot” (audio object);
- **Literary pun**: “Dung beetles play ball!” (preface); “Dung beetles turn dung into balls and are equipped with their forehead and legs to push these balls for some distance” (abridged audio object);
- **Turn of phrase**: “An inch or two give or take a foot” (preface); “Dung beetle nests are usually underground, and can range from a few inches to a few feet deep” (audio object);
- **Definition pun**: “There’s a cat in the garden!” (preface); “Specimen #129 is a John Macoun sample, it is known as a pussy toe because the plant flower and fruit represent a cat’s foot” (audio object);
Fig. 4 A plan drawing of the tangible object revealing the curved bottom that suggest resting in the palm of the visitor’s hand

- **Riddles**: “What is always naked and thinks on its feet?” (preface); “Where gastropods are shelled critters with stomachs that sit on a primary foot, cephalopods are bare critters with heads that sit on a primary foot” (audio object);
- **Understatement**: “Longer than you would want to know” (preface); “Tapeworms come in varying lengths and sizes. Interestingly, the longest recorded tapeworms have been those that live in humans” (audio object);
- **Contradiction**: “Ice age dentistry” (preface); “This deformed tooth is a very interesting case. It was the first recognized pathological problem in an ice age animal” (audio object).

The audio recordings of the prefaces and audio objects used a diverse set of voices that were informal in tonality and style. This added to the conversational feel and created an imaginary scene of a virtual cocktail party of natural historians and scientists that followed you through the museum. As we discussed in Sect. 3.1, we identified the natural history scientists as our keystone species. We organized sessions of recorded walkthroughs of the exhibition asking each scientist to provide commentary [46]. These sessions became the basis for the discrete audio objects that were categorized by topics and relationship to artifacts on display.

4.2.3 Navigation

We structured navigation at a macro level, where visitors move throughout the exhibition space in between artifact displays, and a micro level, where visitors are within a specified interactive zone in close proximity to an artifact display.

On the macro level the input is the visitor’s movement, which creates an ambient soundscape through the audio display related to artifacts nearby. We divided the exhibition space into interactive zones and mapped concepts of interest to each zone and display (in regard to the user model we distinguish between concepts represented in the artifacts and concepts that can be associated with the artifacts based on user’s interests, we refer to the former as visual concepts, see [48]). The concepts are translated into environmental sounds such as the sound of an animal habitat, and sound of animals such as the flapping of crane’s wings. The visitor navigates the exhibit exploring it on a thematic level through the ambient sounds that are dynamically created. If a set of visual concepts strongly matches the visitor’s interest, the related audio is acoustically more prominent. Figures 5 and 6 depict how the visitor’s movement in the exhibition space creates the soundscape. Darkened areas within the superimposed map of the exhibition space represent different visual concepts translated into sound triggered by the proximity of the visitor. In Fig. 5, two dark areas are highlighted. The slightly darker area represents nearer proximity of the visitor to one set of concepts over another signaling that while the audio is composed from both zones, the nearer zone is more prominent. In Fig. 6, the highlighted zone (red in a color version of the figure) represents a strong match between the visitor’s current concepts of interests and the nearby visual concepts and would therefore be acoustically prominent.

Fig. 5 Still frame depicting the prominence of sounds in a soundscape reflecting what’s on display based in the visitor’s proximity
4.3 User model

The adaptive and user model approach in echo is not the focus of this paper, we refer readers to another paper that discusses our approach in considerable depth [48]. Our approach is characterized by the use of an integrated modeling technique, supported by an ontologies and rule-based system for information retrieval. We believe that this unique approach supports a TUI that relies on limited explicit input and substantial implicit input, while at the same time the semantic web approach allows for rich and coherent information output within an audio display that is adaptive to the interactor’s dynamic exploration and discovery within the museum environment. The user model dynamically integrates movement interaction and visitor content selection into initial pre-selected preferences. Based on this dynamic model we could infer potential interests and offer a corresponding range of content choices. In addition, the use of semantic technologies allowed for coherent and context responsive information retrieval.

5 Analysis of the interface and interaction

In order to understand the situated nature of the interface we provide an analysis utilizing the TUI frameworks of Shaer’s TAC paradigm [49] and Fishkin’s taxonomy [50]. Over the years various frameworks have been proposed to better define TUIs. Holmqquist et al. [51] proposed defining concepts of containers, tools, and tokens. Ullmer and Ishii [52] proposed a framework known as the MCRit and later the Token + Constraint System [53] that highlighted the integration of representation and control in TUIs. Shaer and others have extended MCRit to propose their token and constraints (TAC) paradigm [49].

The TAC paradigm defines TUIs across three concepts: token, constraint and variable. A token represents digital information or a computational function, a

<table>
<thead>
<tr>
<th>Turn</th>
<th>Prefaces played</th>
<th>Preface/audio object selected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>left</td>
<td>center</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Fig. 7 "1-2-4" navigation model
constraint limits the token’s behavior, and a variable is digital information that is either statically or dynamically represented by tokens. Shaer defines several categories within TAC in which among other things, TACs can be composed together. We have specified ec(h)o using the TAC paradigm in Table 1. For example in the first TAC, the cube is a token, and the constraint is the cube together with hand dexterity. The variable is the preface, and the behaviors of the cube are specified as well. ec(h)o’s TUI would be in the token + constraint category since the wooden cube is a token and physically sets its own constraints on its behavior.\textsuperscript{2} Further in Table 1, we have added two additional TACs, 2 and 3 that include the visitor’s body as a token and two aspects of the architectural space as constraints. While we have specified ec(h)o within the TAC paradigm it seems to have strayed well beyond a purely TUI when considering the visitor as a token and the architecture as a constraint.

Fishkin’s taxonomy is a two-dimensional space across the axes of embodiment and metaphor \cite{50}. Embodiment characterizes the degree to which “the state of computation” is perceived to be in or near the tangible object. Fishkin provides us with four levels of embodiment: distant representing the computer effect is distant to the tangible object; environmental representing the computer effect is in the environment surrounding the user; nearby representing the computer effect as being proximate to the object; and full representing the computer effect is within the object. Fishkin uses metaphor to depict the degree to which the system response to user’s action is analogous to the real-world response of similar actions. Further, Fishkin divides metaphor into noun metaphors, referring to the shape of the object, and verb metaphors, referring to the motion of an object. Metaphor has five levels: none representing an abstract relation between the device and response; noun representing morphological likeness to a real-world response; verb representing an analogous action to a real-world response; noun + verb representing the combination of the two previous levels; full representing an intrinsic connection between real-world response and the object which requires no metaphorical relationship.

In Fig. 8, we have applied Fishkin’s taxonomy to ec(h)o. Embodiment would be considered “environmental” since the computational state would be perceived as surrounding the visitor given the three-dimensional audio display output. In regard to metaphor, the ec(h)o TUI would be a “noun and verb” since the wooden cube is reminiscent of the wooden puzzle games in the museum and the motion of the cube determines the spatiality of the audio as turning left in the real-world would allow the person to hear on the left. If we consider the visitor’s movement the embodiment factor would still be environmental and we’d have to consider the visitor’s body as being “full” in Fishkin’s use of metaphor. In regard to understanding the entire system we’d have to plot ec(h)o

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{figure8.png}
\caption{ec(h)o plotted in Fishkin’s tangible user interface (TUI) taxonomy \cite{50}}
\end{figure}

\textsuperscript{2} It is worthwhile to note that the TAC paradigm does not account for very minimal tangibles such as ec(h)o and Live Wire in which tokens and constraints are not related components but are integrated into one component alone such as a cube or wire.

---

**Table 1** ec(h)o specifications using the TAC paradigm \cite{49}

<table>
<thead>
<tr>
<th>TAC</th>
<th>Representation</th>
<th>Behavior</th>
<th>Variable</th>
<th>Action</th>
<th>Observed feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cube</td>
<td>Cube and Hand</td>
<td>Preface</td>
<td>Hold up</td>
<td>Audio object heard in the center</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rotate left</td>
<td>Audio object heard on the left</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rotate right</td>
<td>Audio object heard on the right</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Keep down</td>
<td>System is silent</td>
</tr>
<tr>
<td>2</td>
<td>Body</td>
<td>Interactive zone (display area)</td>
<td>Preface</td>
<td>Enter</td>
<td>Soundscape fades and prefaces are heard on the left, right and in the center</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Exit</td>
<td>Soundscape is heard</td>
</tr>
<tr>
<td>3</td>
<td>Body</td>
<td>Exhibition space</td>
<td>Soundscape</td>
<td>Movement</td>
<td>Soundscape changes</td>
</tr>
</tbody>
</table>
between "noun + verb" and "full" on the metaphor axis. While Fishkin's taxonomy addresses context beyond the tangible object rather well, again the inclusion of people themselves as a TUI seems beyond the scope of the taxonomy despite its application here.

The interaction with the tangible object in ec(h)o is characterized as a verb under Fishkin's taxonomy and an action in the case of Shaer yet movements have complex non-linear qualities that elude simple categorization. In Sect. 4.2.2 we discussed examples of the types of word play, puns and riddles we used in our audio to encourage play and discovery. The tangible interface aimed for a complementary physical play, which as we discussed is open and often can be subtle and implicit like toying with a ball in your hand. We designed the tangible object such that it had suggested actions like resting in a palm or pivoting on a wrist yet we knew we could not design the actions directly rather only suggest possibilities, what Djajadiningrat refers to as the action-potential of physical objects [39]. Further, the physicality of the objects meets our bodies in often unique or wide ranging kinesthetic combinations in which optimal efficiency gives way to play and experimentation.

In what are simple actions of holding and rotating the cube we observed a diverse set of interaction techniques when selecting prefaces. We identified at least five basic techniques:

- **Hold and rotate**, one hand holds the cube resting on the palm while the other hand rotates it in place (see Fig. 9a, b);
- **Hold, rotate and cover**, one hand holds the cube resting on the palm while the other hand or both hands rotate the cube. The topside is uncovered until the selection is made and then the topside is covered again until its time to make another selection\(^3\) (see Fig. 9c, d);
- **Cradle and hide**, two hands rotate and cradle the cube, after selection is made the colored side is rotated and hidden against the visitor's body (see Fig. 9e);
- **Rotate wrist**, one hand holds the cube between fingers and thumb, and rotates the wrists to make a selection (see Fig. 9f, g);
- **Rotate with fingers**, one hand holds the cube and rotates it by rolling with the fingers and thumb (see Fig. 9h).

It is important to note that we observed combinations and variations of these techniques, as well as individual experimentation with the different approaches. As one might expect we also observed a range of methods for holding the cube when not selecting prefaces or walking through the exhibition such as cradling it in hands, holding it at one's side or behind one's back, dangling it from the wrist, or holding its leash to gently sway it from side to side. This sense of play extended to participant's movements through the exhibition space. In the interviews, participants commented on how they returned to zones to see if the system would indeed not repeat audio objects already heard. In addition to moving from zone to zone participants appeared to experiment with their movements entering and exiting zones altering the soundscape (for example, see the number of location changes in a short period of participants 3 and 6 in Table 2 in Sect. 7).

We provide these details of interaction to describe the degree of play and variety afforded by the interface as opposed to a single path of interaction—all of these approaches worked equally well. Djajadiningrat points out that aesthetic interaction is where "there is room" for a myriad of types, combinations and sequences of actions [39]. This experiential space is created in the embodied action between physical objects and our bodies. In Sect. 2.3 we discussed the example of the ball as a form of pragmatic aesthetics in Petersen et al [40]. A wooden cube, like a ball is a very familiar object that has a history of use in games and play that can be open-ended and exploratory. As Petersen observed, the ball promotes playfulness and promises a different type of potential than a tool. Rather then the promise of efficiency and accuracy, the ball and in our case the cube promises discovery and exploration.

6 Implementation

Our prototype for testing consisted of four main components: position tracking, vision sensing, audio engine, and reasoning engine. Two main types of events trigger the communication between the components: visitor's movement through the exhibition space and selection of audio objects. The high level architecture is shown in Fig. 10. The knowledge models and ontologies refer to the semantic web approach to information retrieval which is not pertinent to the discussion here [54].

The prototype was installed and tested in the Finders Keepers exhibition at the Canadian Museum of Nature. The exhibition theme was collecting natural
Fig. 9  a–h Different interaction techniques for selecting *prefaces*: a, b Hold and rotate; c, d Hold, rotate and cover; e Cradle and hide; f, g Rotate wrist; h Rotate with fingers
Table 2 Test session characteristics

<table>
<thead>
<tr>
<th>Participant</th>
<th>Length</th>
<th>No. of cycles</th>
<th>No. of selections</th>
<th>No. of locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>10:36</td>
<td>27</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>Participant 2</td>
<td>6:19</td>
<td>11</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Participant 3</td>
<td>8:56</td>
<td>22</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Participant 4</td>
<td>9:53</td>
<td>21</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>Participant 5</td>
<td>9:18</td>
<td>22</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>Participant 6</td>
<td>5:01</td>
<td>16</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>Expert 1</td>
<td>15:03</td>
<td>32</td>
<td>23</td>
<td>9</td>
</tr>
<tr>
<td>Expert 2</td>
<td>17:58</td>
<td>36</td>
<td>29</td>
<td>7</td>
</tr>
</tbody>
</table>

history artifacts in Canada. While theoretically we could have installed the system throughout the exhibition we created only three zones of interaction due to our limited installation times between the open hours of the museum. We produced over 6(K) reusable audio objects and annotated them with the ontological information. The average length of an audio object is approximately 15 seconds. The shortest is 5 s and the longest 31 s. The prefaces typically are 3 s in duration.

Position tracking We used a combined radio frequency identification (RFID) and optical position tracking system developed by Precision Systems (http://www.precision-sys.com). Optical tags were attached to the tops of the headphones. Visitors carried an active RFID tag in a pouch. We installed cameras over the interactive zones and one in the central area of the space. This was adequate for tracking the visitor location throughout the sessions.

Audio engine We developed a multi-channel editor, mixer and server in the Max/MSP™ environment to function as the audio engine. This engine created dynamic soundscapes and delivered unique channels of stereo audio to individual users. The audio was delivered wirelessly over FM transmitters that provided a stereo signal. Each visitor carried a small inexpensive digital receiver in a pouch.

Vision sensing A vision sensing system supported the selection of audio objects via the tangible interface. We developed a system in Max/MSP based on the "eyes" system (http://www.squishedyeballs.com). Cameras were installed over each interactive area.

Reasoning engine The reasoning engine receives all the input and directs output based on inferences based on a rule system and user model. Information retrieval actually employed a semantic web approach that allowed us to select the audio objects based on their semantic properties and how they relate to the museum artifacts, exhibits, individual user interests and user’s interaction history. The system was implemented using the JESS inference engine with the DAMLJessKB extension that converted DAML + OIL ontologies to Jess facts. The reasoning module was connected with other modules through the user datagram protocol (UDP) socket connections [55].

7 Evaluation

The exhibition, 'Finders and Keepers' contains seven exhibits, five of which are booth-type exhibits, each with several dozens of artifacts organized around topics (see Fig. 11). Two exhibits are open exhibits with larger artifacts such as a mastodon skeleton (see Fig. 11 An example of a “booth-type” display in the exhibition Finders Keepers
For the exhibition we created three interactive zones: two in booth-type exhibits and one in an open space exhibit.

The formal user evaluation effort involved sessions with six participants and two expert reviewers. The participants had previous experience with interactive museum systems such as docent tours (three participants), interactive kiosks (3), audiotape systems (4), film and video (5), seated and ride-based systems (2) and PDA systems (2). The test group included two men and four women, from 25 to 53 years old. The experts included a senior researcher and senior interaction designer from the museum. Both were familiar with the exhibit and its underlying concepts. In addition to an extended discussion with the expert reviewers they provided us a written evaluation of the system.

Table 2 shows the characteristics of each user session: the total length of the interaction, number of interaction cycles, number of selected and listened to audio objects, and number of location changes.

Our evaluation is based on Miller and Funk's [56] use of traditional 'validation' and 'verification' approaches in evaluating ubiquitous computing systems. Our verification efforts focused on user experience and the perception of the system. Our validation efforts focused on the user model and system response components. Since user experience is more relevant to our discussion in this paper we provide here a short summary of the validation results that have been discussed in detail in [48].

7.1 Summary of user model and performance evaluation

The validation of the ec(h)o components, namely user model and object selection, showed that these performed at the required level of accuracy and flexibility. In regard to the experience design goals of play and discovery, our integrated modeling approach implemented two techniques to facilitate wider exploration and the discovery of new topics of interests and the ability to make new connections among topics and artifacts. The first being the aim of keeping interests balanced such that a given topic or set of topics does not dominate and prevent exploration of new topics, for this we used a spring model to proportionately moderate levels of interest. We felt it was important that the user model learns to "forget older interests" so that newer ones can be invoked. The second technique is to maintain a high level of variability of primary and secondary interests among the objects presented. This affords greater opportunity for the user to evolve his or her interest through a reflection on content as discussed above (see Sect. 6.3). The results of a separate laboratory tests showed that these techniques contribute to the goal of establishing dynamics in the user model that support exploration and discovery of new interests through moderating evolution in the user interests, maintaining significant influence of changing context (when a visitor moves to another exhibit), and protecting against the domination of a few concepts that would choke off exploration.

We introduced the evaluation of system response or in our case, object selection based on interaction criteria of variety, the richness of choices for further interaction at each interaction step; sustained focus, ability of the system to sustain the focus on particular interests; and evolution, ability of the system to follow shifting user interests during interaction with the system. We can conclude that the system offers the highly variable objects when user changes the location and the variety increases as the user continues the interaction in a particular location. The high variety during the object selection steps is supported while the system maintains the focus on the concepts of interest as expressed in the user model. The low value of evolution during the object selection stage indicates the continual change in topics offered corresponding to the modest changes in the user model. This behavior matches our expectations. Several ranking criteria are combined to select audio objects offered in the next step. It is the weight with which these criteria contribute to the object ranking that determines the combination of the concepts of interest in the objects offered. To achieve different behavior from the system the relative weight of contributing criteria would have to be altered.

Fig. 12 An example of an "open exhibit" display in the exhibition Finders Keepers
7.2 Evaluation of user experience

We evaluated user experience through observation, a questionnaire, and a semi-structured interview. The questionnaire included 63 questions that assessed user experience related to the overall reaction to the system, the user interface, learning how to use the system, perceptions of the system's performance, the experience of the content, and degree of navigation and control. Majority of the questions in the questionnaire were on a Likert scale yet it provided for open-ended written comments. Throughout the questionnaire, and especially during the semi-structured interviews we looked for an overall qualitative assessment of the experience based on Bell’s ecological components of liminality and engagement [5]. For a summary of the questionnaire results see Fig. 13.

Overall, participants found the system enjoyable and stimulating, perhaps in part due to its novelty. The general sense of satisfaction was split between those participants who liked the playful approach and those who did not. While our sample was small we noted a clear age difference in that the “younger” participants rated satisfaction higher based on their liking of the playful approach (this was confirmed in the semi-structured interviews).

Among the factors that stood out as most positive for the participants was that the cube and audio

---

**Fig. 13** Summary of the questionnaire results on user experience (n = 6; 63 questions on Likert scale of 1-5 (five being best))

- **Overall reaction:** 5 ratings including "terrible-wonderful; difficult-easy" averaged for each participant. Overall average rating is 3.6 with a standard deviation of 0.40.
- **Tangible user interface:** 7 ratings including "uncomfortable-comfortable; difficult-easy to manipulate; annoying-enjoyable" averaged for each participant. Overall average rating is 4.24 with a standard deviation of 0.45.
- **Headset:** 2 ratings including "uncomfortable-comfortable" averaged for each participant. Overall average rating is 2.92 with a standard deviation of 0.59.
- **Learning Curve for the system:** 8 ratings including "difficult-easy to get started; risky-safe to explore features; unclear-clear feedback" averaged for each participant. Overall average rating is 4.07 with a standard deviation of 0.31.
- **Perception of system performance:** 8 ratings including "slow-fast system response; never-always reliable" averaged for each participant. Overall average rating is 3.83 with a standard deviation of 0.56.
- **Quality of the content:** 15 ratings including "uninformative-informative; generalized-customized for me; rigid-playful; predictable-surprising" averaged for each participant. Overall average rating is 3.78 with a standard deviation of 0.47.
- **Quality of the audio experience:** 9 ratings including "confusing-clear; mechanical-human; wasteful-valuable" averaged for each participant. Overall average rating is 3.67 with a standard deviation of 0.60.
- **Navigation and control:** 8 ratings including "never-always able to navigate efficiently; always-never found myself lost in the system; always-never found myself uncertain of the system state" averaged for each participant. Overall average rating is 3.23 with a standard deviation of 0.95.
delivery were seen as playful. The open-ended written comments and semi-structured interviews made this point clear as well. The TUI was well received especially in terms of ergonomics and ease of use. This was not a surprise to us since our early testing and participatory design sessions provided us with considerable feedback, especially on ease of use and enjoyment. We went through several iterations and form factors of the wooden cube and tested it against different hand sizes. This may have also resulted in the fact that learning to use the interface and navigation were rated highly and participants felt the system had a low learning curve and that it was easy to get started:

Umm, I found it was really easy. Sometimes I got so engaged in listening to what they were saying that I forgot in which orientation I was holding the cube. And I found that I would have to occasionally look down. But the way it was designed with the round part to go in your palm... it was really easy to quickly reorient myself to how I was holding that cube. (Participant 5)

It should be stated that we provided a short tutorial on the system at the beginning of each evaluation but nevertheless this feedback is encouraging. Interestingly, the audio content was perceived to be both accurate and clear. The issue of trust and delivery style is an area to further investigate. Since we collected the information directly from scientists and staff at the museum rather than a more generic source we wonder if this contributed in part to this result [46]. These results lead us to believe that the system meets or satisfies many of the current advances of museum guide systems.

The questionnaire did point out challenges and areas for further research. Some things we expected such as the headphones were uncomfortable, yet to such a degree that we are currently rethinking the tradeoff between personalized spatial audio and use of headphones. Other results point to a threshold in the balance between levels of abstraction and local information. Since visitors had difficulties at time connecting what they were listening to and what was in front of them (in part this was an inherent challenge in the exhibition since the display cases had dozens to over a hundred artifacts, see Fig. 10a, b). In many respects this contributes to our finding that the ontological approach did not provide a clear enough contextual link between the artifacts and the audio information. In addition, we see both a threshold point in play versus focused attention on the exhibit in that the question relating to the content asking if it was "distractive-synergistic" scored 2.83. This raises the issue of balance in play and the possibility to shift attention away from the environment rather than play as a means of further exploring the environment.

In an open-ended question in the questionnaire and through the interviews we explored the issues of liminal play and engagement. The results here are quite clear that play was a critical experiential factor in using the system. It was often remarked how the experience was similar to a game:

The whole system to me felt a lot like a game. I mean I got lost in it, I found myself spending a lot of time in a particular area then I normally would. And just the challenge of waiting to hear what was next, what the little choice of three was going to be. Yeah... So I found it over all engaging, it was fun, and it was very game-like. (Participant 4)

The playfulness did in most instances suggest a quality of engagement that led to learning even through diverse types of museum visits from the visitor who browses through quickly but is still looking to be engaged to the repeat visitor who experiences the audio information differently each time:

I learned a lot and well you know I am a scientist here, and I think anybody going through, even people who are in a real rush, are going to pick up some interesting facts going through. And... I mean, that was good, the text was great and was short enough that somebody in a rush is still going to catch the whole thing. (Participant 1)

As mentioned earlier, there is a threshold between play in support of the exhibit on display and play with the system that can be an end in itself and even a distraction. For example, one user's enthusiasm for the game-like quality led her to at times pay more attention to the interaction with the system than the exhibition. In addition, people respond to play differently and can be argued to belong to different types of players [57]. One participant would have preferred a more serious and "non-playful" approach.

The prefaces were playful, but the text was not at all, you know, that contrast between them.... but I find it was too playful and I think maybe, either you, or maybe you could give people the choice between you know choosing a playful or a non-playful version." (Participant 2)

In addition, participants' observations on the liminality of the experience manifested in comments suggesting that play was more natural for children rather than themselves, however as expressed below, they soon overcame this issue:
At first it felt a little bit strange, especially holding this cube that looked like a children's toy, and I felt a little bit awkward about doing that, but I got over that pretty quickly. (Participant 5)

It was quite chatty, which was kind of fun. I kind of felt like 'Oh, I bet like a twelve year old would really like this'. (Participant 3)

8 Discussion

In this paper we've explored situated play in a tangible and adaptive audio museum guide. Our approach in ec(h)o was to create a coherent space for play and discovery across all components of the design including reasoning, audio delivery and interface. The space suggests actions and meaning but maintains an openness and interpretation that requires playful interaction on the part of the user in order to realize the action-potential or relevancy of the information. While we see that the results of our pilot study support the notion that learning effectiveness and functionality can be balanced productively with playful interaction, we see further research and some caution when dealing with the space of playful and interpretive interaction. With the practicalities of design in mind we see issues of balance in between play and the environment, and the space between interpretation and information that links the audio content to the artifacts. Theoretically we have questions on the degree to which we best understand the contextual and situated aspects of TUIs.

8.1 Design issues

The balance of playful intervention When is a good thing too much? In our case, playfulness does not directly lead to satisfaction. In our results, playfulness was identified positively in all aspects of the interface yet overall satisfaction was split between those participants who enjoyed playing and those who did not. As we reported, one participant explicitly asked for a non-playful version. However, we did not expect our approach or any approach to museum interaction to be universally accepted. We actually find the question of too much play to be of more interest. There is a need to find the balance between play in support of the exhibit and play with the system that can be a distraction and even an end in itself. Otherwise, designers run the risk of users engrossed in playing with the system at the expense of interacting with their surroundings, as one participant commented happened to her periodically. This is not the same issue as the one we raised about PDAs demanding full attention for that is an inherent design and cognitive relationship given the GUI nature of the device. Playful interaction lends itself well to integrating with the context and in many cases depends on it, as in bouncing a ball off the floor or wall. While we achieved a reasonable balance and are generally on the right track with our approach, we feel more is required for a better understanding of how to design situated TUIs in regard to play.

Balancing the richness of ambiguity & the richness of information When is a good thing too little? At times participants had difficulties connecting what they were listening to and what was in front of them. It is possible that the system did not always provide a coherent story, a resulting tradeoff of our aims of open discovery. Nevertheless, a much richer model of discourse and storytelling could be an option to pursue, for example a richer world model for location as Goßmann and Specht describe [58]. Visitors in museums clearly invest in connections with concrete artifacts while ec(h)o experimented with the idea of connections between artifacts and audio objects at the higher ontological level. The results indicate that a much richer model is needed or the hypothesis of linking objects at higher abstract ontological levels is not the best approach for ubiquitous context-aware applications or it has to be combined with other approaches.

Puns, riddles or icons What is ten pixels square, black and white all over and not funny? We discussed the range of puns, riddles and word play we used for the prefaces that served as indices for navigation choices. In comparison, we performed preliminary testing with other approaches like earcons [59] and the more traditional question and answer structure. The earcon design was perceived as too confusing and abstract. It was simply too difficult to encode the range of concepts of interests and themes into communicable earcons that could be remembered by the user. The question and answer design was viewed as static and unrelenting after only a few turns. We feel the early efforts of our word play approach are promising. The use of word challenges as either indices or user instructions has interesting potential in interaction design.

8.2 Situating TUIs

The concept of TUIs is deceptively simple. We manipulate the world through physical atoms with overwhelming ubiquity. This includes manipulating the world of digital bits since Fishkin argues a keyboard
can be considered a tangible interface. A possible criticism of his taxonomy is that it may be too broad and inclusive to be useful yet in our view this approach widens the concept to expose boundaries [50]. We found this approach very useful for the fact that it considers the contextualization of TUIs. As we encountered in our analysis, it also opens interesting questions such as the nature of the human interactor and the role of embodied interaction in a tangible interface. At the moment however, we are most interested in the contextualization issues of TUIs.

In Fig. 14 we plotted TUIs that we cited and described in our discussion of related works (see Sects. 2.1, 2.2, 2.3).

We consider these projects to be contextual in that the environment beyond the immediate interface elements affects the interaction or meaning of the interaction, or as in ec(h)o or Live Wire [31], the works are situated in an identifiable setting. Live Wire mirrors the connection between network activity within the immediate office space and network traffic originating on the network outside of the office such as email. Despite the differences in ambientROOM [26] and WorkSPACE [41] they are both office environments and thus context specific. ambientROOM is the most complex of the projects here and in fact represents a number of different TUIs connected only by their shared context. In respect to contextual TUIs, a state of "full" embodiment is not a desirable quality. Utilizing Ishii's notion of "foreground" and "background" activity, the "foreground" activity comes at the cost of awareness of "background" bits or activity. The contextualized realm is the awareness of the activity and setting around you. The state of "full" metaphor is interesting in that without "full" embodiment, the fullness seems to come from the active presence of the human body. In PingPongPlus [26] it is difficult to consider the ping-pong paddle as active without an arm and body attached to it moving it to hit the ball. According to Ishii, the paddle "can co-evolve with a user by changing its physical form and being united with the human hand," and paddles are a "transparent physical extensions of our body" [26]. The traces of the body presence are left on a well-used paddle in the form of thumb and finger marks. We have already discussed the notion of the museum visitor in ec(h)o as a "full" metaphor tangible interface (see Sect. 5).

This is important since Fishkin concludes his discussion of his taxonomy by identifying that the domain of TUIs is evolving toward TUIs converging on "full" embodiment and "full" metaphor. He cites "Sketchpad" [60] as an example of a "full" metaphor and "Illuminating Clay" [61] as an example of "full" embodiment. We strongly feel this overlooks the situational value of the taxonomy and risks overlooking developments in situated TUIs.

9 Conclusion

ec(h)o is an augmented audio reality system for museum visitors that utilizes a tangible interface. We developed and tested the prototype at the Canadian Museum of Nature in Ottawa. In ec(h)o we tested the feasibility of audio display and a TUI for ubiquitous computing systems — one that encourages an experience of play and engagement. In this paper we have presented relevant work in the domains of adaptive museum guides and audio displays, ludic approaches to TUIs, and aesthetic interaction. We provided an overview of our design motivations rooted in ethnography and concepts of ecologies that together led to our approaches in audio delivery and tangible interface. We described the components of our prototype and gave an analysis of our interface utilizing TUI frameworks that revealed the embodied and contextual nature of our design. We also analyzed the interaction revealing the aesthetic qualities of the interaction pattern between the object and the visitor, and the myriad of interaction paths. We also described our implementation and evaluation design.

The findings of this project are positive while also calling for more research in several areas. We conclude that based on our results from our pilot study learning effectiveness and functionality can be balanced productively with playful interaction through an adaptive audio and TUI if designers balance the engagement
between play and awareness of the environment, and balance the richness of ambiguity with the richness of information that links the audio content to the artifacts. We see further research in the role of puns, riddles and word play in interaction design, and we especially see the need to further develop theoretical frameworks for TUIs that reveal and explain the situated nature of the many projects that adopt a tangible and aesthetic interaction approach.

Acknowledgments Work presented in this paper is supported by a Canarie Inc. grant under the E-Content program. Authors would especially like to thank Dr. Mark Graham and his colleagues at the Canadian Museum of Nature in Ottawa for their enthusiastic support of this project. We would also like to thank our colleagues and students, Kenneth Newby, Dale Evernden, Leila Kalantari, Doreen Leo, Gilly Mah, Robb Lovell, Mark Brady, Jordan Willms and Phil Thomson for their contributions.

References

30. Crampton-Smith G (1995) The hand that rocks the cradle. ID magazine, pp 60-65


Ontology-Based User Modeling in an Augmented Audio Reality System for Museums

MAREK HATALA and RON WAKKARY
School of Interactive Arts and Technology, Simon Fraser University, Surrey, BC, Canada
V3T 2W1. e-mail: {mhatala, rwakkary}@sfu.ca

(Received: 16 November 2004; accepted in revised form 13 August 2005)

Abstract. Ubiquitous computing is a challenging area that allows us to further our understanding and techniques of context-aware and adaptive systems. Among the challenges is the general problem of capturing the larger context in interaction from the perspective of user modeling and human-computer interaction (HCI). The imperative to address this issue is great considering the emergence of ubiquitous and mobile computing environments. This paper provides an account of our addressing the specific problem of supporting functionality as well as the experience design issues related to museum visits through user modeling in combination with an audio augmented reality and tangible user interface system. This paper details our deployment and evaluation of etc(o) - an augmented audio reality system for museums. We explore the possibility of supporting a context-aware adaptive system by linking environment, interaction objects and users at an abstract semantic level instead of at the content level. From the user modeling perspective etc(o) is a knowledge-based recommender system. In this paper we present our findings from user testing and how our approach works well with an audio and tangible user interface within a ubiquitous computing system. We conclude by showing where further research is needed.

Key words. audio augmented reality, context-aware, museum guide, ontologies, semantic technologies, tangible user interface, testing, ubiquitous computing, user evaluations, user modeling

1. Introduction

Fundamental to human-computer interaction (HCI) is the design of interactive systems that support people’s goals and respond to individual backgrounds. In ubiquitous computing it is equally important to consider the influence of context on people’s interactions and experiences. The intent is, as Fischer argues “to say the ‘right’ thing at the ‘right’ time in the ‘right’ way” (Fischer, 2001). A critical factor in ubiquitous computing is that what is perceived as “right” is largely mediated by the context within which the users find themselves.

In the area of user-adapted interaction, user modeling has attempted to address many issues related to HCI. Fischer provides a clear account of the successes and future challenges of user modeling in HCI (Fischer, 2001). Among these challenges is the general problem of capturing the larger context in interaction (see Fischer,
The imperative to address this issue is great considering the emergence of ubiquitous and mobile computing environments. This paper provides an account of addressing the specific problem of supporting functionality as well as the experience design issues related to museum visits through user modeling in combination with an audio augmented reality and tangible user interface system. We developed and tested a museum guide prototype, known as ec(h)o in order to research interaction design, user modeling, and adaptive information retrieval approaches that respond to the richness of a museum visit and the museum context.

Our aim is to support the limited input common to tangible user interfaces while maintaining rich and adaptive information output via a three-dimensional audio display. We believe an integrated modeling technique that is weighted toward modeling of implicit communication works well with a tangible user interface in creating a playful and discovery-rich experience. We believe this approach combined with ontologies and a rule-based system for information retrieval provides a richness of information that is responsive to the context and unique aspects of the museum visitor’s interaction.

Our findings are both encouraging and cautionary. First, we found that it is possible to build a highly flexible and accurate user model and recommender system built on information collected from user interaction. This approach supported a user experience of liminal play and engagement. The ontologies and rule-based approach proved to be a strong combination. However, the ontological approach did not provide a clear enough contextual links between the artifacts and audio information and either more extensive knowledge engineering is needed or our approach has to be combined with stronger narration or discourse models.

In this paper we first review the general problem of context, our intended approach, and provide theoretical and related research as background. Following that we provide an account of our design and rationale for the prototype and its implementation. We give a detailed report of our evaluation and findings. We conclude with a brief analysis of our findings and discussion of future issues and research direction.

2. The Challenge of Capturing the Larger Context

Many HCI theorists and researchers identify issues of “context” as putting a strain on the traditional theories of HCI (Bodker, 1990; Dourish, 2004; Gay and Hembrooke, 2004; Nardi, 1995). As Nardi puts it, “we are beginning to feel a theoretical pinch, however – a sense that cognitive science is too restrictive a paradigm for finding out what we would like to know” (Nardi, 1995, p. 13).

For example, a visit to a museum reveals an everyday yet complex interaction situation. The factors within museum experiences are social, cultural, historical, and psychological. The influences on the experience vary from the actions and
previous knowledge of the visitor, visitor's learning style, and the dynamics of others around them including friends, family and strangers. Naturally, the experience is also affected by the presence of the artifacts and collections, which are products of institutional history, curatorship, exhibition design, and architecture. The time of day, duration of visit, room temperature and so on all have an impact. The experience can be characterized as *multivariate*, that is, it cannot be assessed by a single factor such as exhibit design, signage, or time spent in front of an artifact (vom Lehn, et al., 2001). Instead, the museum experience is subject to multiple influences and results in multiple outcomes (Leinhardt and Crowley, 1998).

Many similar situations have been discussed in design research such as how we work (Ehn, 1989), seek information (Nardi and O'Day, 1999), learn (Gay and Hembrooke, 2004), and live in our homes (Bell and Kaye, 2002; Tolmie et al., 2002).

In response to the issue of context, ethnographic and scenario-driven methods have begun to take hold in HCI practice (Carroll, 2000, 2002; Suchman, 1987). An emerging set of "context-based" theories for HCI has adapted ideas from an even wider spectrum of psychological, social, political and philosophical theories based on understanding human activity. For example, Nardi, Bodker, Gay and others (Bodker, 1990; Gay and Hembrooke, 2004; Nardi, 1995) have advocated on behalf of activity theory. Dourish (2001, 2004) argues in his concept of embodied interaction that activity and context are dynamically linked - or "mutually constituent" (Dourish, 2004, p.14).

Suchman (1987) argues that the nature of interaction between systems and people require the same richly interpretive work required in human interaction, yet with fundamentally different available resources. For example, humans make use of non-verbal and inferential resources that can handle ambiguity and result in intelligible actions. This is not the case for computers. Fischer argues this raises two challenges: "(1) How can we capture the larger (often unarticulated) context of what users are doing (especially beyond the direct interaction with the computer system)? (2) How can we increase the 'richness of resources' available for computer programs attempting user modeling to understand (what they are told about their users) and to infer from what they are observing their users doing (inside the computational environment and outside)" (Fischer, 2001). In addition, Fischer cites Weiser and Bobrow (Bobrow, 1991; Weiser, 1993) in arguing that ubiquitous computing (and ultimately tangible user interfaces) aims to address the context issue by eliminating the separation between computational artifacts and physical objects, thus creating computational environments that require new approaches to interface and display.

---

1A theory developed by psychologists in the early 1920s (Vygotsky, 1925/1982), as a research tool and an alternative framework for understanding human activity as it relates to individual consciousness.
3. Background and Related Research

This research ties together several distinct domains that we will briefly review. These include adaptive museum guides, non-graphical user interfaces, user modeling, and semantic technologies.

3.1. ADAPTIVE MUSEUM GUIDE SYSTEMS

It is difficult to directly compare ec(h)o with other museum systems since our approach employs a unique form of interaction. However, ec(h)o shares many characteristics with the adaptive systems of HyperAudio, HIPS and Hippie (Benelli et al., 1999; Oppermann and Specht, 2000; Petrelli et al., 2001). Similar to ec(h)o, the systems respond to a user's location and explicit user input. HyperAudio uses a static user model set by a questionnaire completed by the visitor at start-up time. HIPS and Hippie infer the user model dynamically from the interaction but they treat user interests as static. All systems adapt content based on the user model, location and interaction history. There are however many key differences between ec(h)o and these systems. HyperAudio, HIPS and Hippie depend on a personal digital assistant (PDA) graphical user interface (GUI), for example Hippie's audio interface is dependant on the GUI in such instances as eurcoms (Oppermann and Specht, 2000). ec(h)o uses an audio display as the only delivery channel, and a tangible user interface for input. Another difference lies in how the system generates response: ec(h)o uses inference at the level of semantic descriptions of independent audio objects and exhibits. ec(h)o extends the work of the Alfaro et al. (2003) by building a rich model of the concepts represented by the audio objects while HyperAudio and HIPS use partly pre-configured annotated multimedia data (Not and Zancanaro, 2000), and Hippie uses a simpler domain model. The last key difference is that ec(h)o treats user interests as dynamic, we look to evolving interests as a measure of sustainable interaction.

A museum guide that is conceptually more closely related to ec(h)o is the LISTEN project (Eckel, 2001), it is the follow-up to the Hippie system (Goßmann and Specht, 2002). It provides a personalized immersive audio environment delivered through wireless headphones. The LISTEN system is driven by the directional location tracking of the museum visitors and delivers “three-dimensional sound emitted from virtual sound sources placed in the environment” (Terrenghi and Zimmermann, 2004). The sound sequences are pre-processed by curators and artists. They are selected for the visitor based on a user-specified type. ec(h)o's user model changes dynamically based on the interaction. Its approach to the style of audio delivery and interaction model are also different. However, it is difficult to thoroughly compare LISTEN with ec(h)o as comprehensive evaluation results have not been reported beyond preliminary findings (Terrenghi and Zimmermann, 2004).
3.2. NON-GRAphICAL USER INTERFACES

Prior to the evolution of adaptive and user modeling approaches in museum guide systems, there has been a strong trajectory of use of the PDA graphical user interface. Typically, hypertext is combined with images, video and audio (Aoki et al., 2002; Aoki and Woodruff, 2000; Proctor and Tellis, 2003; Semper and Spasojevic, 2002). Aoki and Woodruff have argued that in electronic guidebooks, designers are challenged to find the balance between burdening the visitor with the functions of selection, information management and contextualization (Aoki and Woodruff, 2000). The PDA graphical user interface approach comes at a cognitive and experiential cost. It requires the full visual attention of the visitor such that it is a competing element with the physical environment rather than a valued addition to that environment. Aside from projects like LISTEN, museum systems have mostly maintained the PDA graphical user interface approach despite the shifts in other domains to other approaches that better address the experience design issues most prominent in social, cultural and leisure activities.

Non-visual and non-graphical user interfaces, particularly audio display interfaces have been shown to be effective in improving interaction and integration with existing physical contexts. For example, Brewster and Pirhonen (Brewster et al., 2003; Pirhonen et al., 2002) have explored the combination of gesture and audio display that allows for complicated interaction with mobile devices while people are in motion. The Audio Aura project (Mynatt et al., 1998) explores how to better connect human activity in the physical world with virtual information through use of audio display. Audio is seen as an immersive display that can enrich the physical world and human activity while being more integrated with the surrounding environment. In addition, audio tends to create interpretive space or room for imagination as many have claimed radio affords over television. Audio augmented reality systems combined with tangible user interfaces often create very playful and resonant interaction experiences (Hummels and Helm, 2004). In fact, the distinction between augmented reality and tangible user interfaces can be blurry indeed (Ishii and Ullmer, 1997).

Tangible user interfaces like no other user interface concept is inherently playful, imaginative and even poetic. In addition, the concept has immediacy due to its physicality. Ishii and Ullmer’s notion of coupling bits and atoms was informed by earlier work in graspable interfaces (Fitzmaurice et al., 1995) and real-world interface props (Hinckley et al., 1994). ec(h)o’s tangible user interface draws on this notion by coupling an everyday and graspable object, a wooden cube with digital navigation and information (Ishii and Ullmer, 1997). Ishii was inspired by the aesthetics and rich affordances of scientific instruments (Ishii and Ullmer, 1997) and the transparency of a well-worn ping-pong paddle (Ishii et al., 1999). Simple physical display devices and wooden puzzles at the natural history museum where we conducted ethnography sessions inspired us as well.

In 1992, Bishop’s Marble Answering Machine (Crumpton-Smith, 1995) was an early embodiment of the immediate and playful qualities of tangible user
interfaces. The prototype uses marbles to represent messages on the machine. A person replays the message by picking up the marble and placing it in an indentation on the machine. Ishii’s PingPongPlus (Ishii et al., 1999) explores the intertwining of athletic play with imaginative play. The ping-pong table becomes an interactive surface. The ball movement is tracked and projections on the table of water ripples, moving spots, and schools of fish among other images react wherever the ball hits the table. While ec(h)o is more constrained in its play, the everyday wooden cube provides entry to a qualitatively diverse experience of interaction.

Over the years, various frameworks and interaction models have been proposed to better define tangible user interfaces. Holmquist and others (Holmquist et al., 1999) proposed defining concepts of containers, tools, and tokens. Ullmer and Ishii (Ullmer and Ishii, 2001; Ullmer, 2002; Ullmer et al., 2005) proposed a framework known as the MCRit that highlighted the integration of representation and control in tangible user interfaces. Shaer and others have extended MCRit to propose their Token and Constraints (TAC) paradigm (Shaer, 2004). Most relevant to our approach is Fishkin’s proposed taxonomy which is situated and contextual in its thinking (Fishkin, 2004). Fishkin’s taxonomy is a two-dimensional space across the axes of embodiment and metaphor. Embodiment characterizes the degree to which “the state of computation” is perceived to be in or near the tangible object. Metaphor in this sense is the degree to which the system’s response to a user’s action is analogous to a real-world response to a similar action. Further, Fishkin divides metaphor into noun metaphors, referring to the shape of the object, and verb metaphors, referring to the motion of an object. For example, in ec(h)o, according to Fishkin’s taxonomy embodiment would be considered “environmental” since the computational state would be perceived as surrounding the visitor given the three-dimensional audio display. In regard to metaphor, ec(h)o would be a “noun and verb” since the wooden cube is reminiscent of the wooden puzzle games in the museum and the motion of the cube determines the spatiality of the audio as turning left in the real-world would allow the person to hear on the left.

3.3. USER MODELING

‘Knowledge-based HCI’ (Fischer, 2001) explores the possibility of implicit communication channels between a human and a computer. These channels capture the idea of shared knowledge about problem domains, communication processes, and agents involved with communicating parties. This notion is very close to the goals of user modeling (Wahlster and Kobsa, 1989). Several researchers worked on the incorporation of user modeling in order to improve the collaborative nature of human-computer systems (for examples see Fischer, 2001). In our research we expand the role of user modeling into the realms of audio augmented reality and tangible user interfaces.

In the context of our work, the user model performs the function of a recommender system (Resnick and Varian, 1997). “Recommender systems represent user
preferences for the purpose of suggesting items to purchase or examine” (Burke, 2002). Several types of recommendation techniques have been developed: collaborative, content-based, demographic, utility-based, and knowledge-based. Often the researchers combine several techniques to achieve maximum effect. Burke (2002) compares the recommendation techniques from the perspective of their ability to deal with the ‘ramp-up’ problem (Konstan et al., 1998): an introduction of new users and new items. In this regard, knowledge-based recommenders perform favorably. This is an important feature for ubiquitous computing environments that often manifest the ‘walk-up-and-use’ characteristic. Knowledge recommender systems require three types of knowledge (Burke, 2002): catalog knowledge or knowledge about objects to be recommended, functional knowledge of mapping between user needs and objects, and user knowledge. In the case of ubiquitous computing applications the functional knowledge must include the knowledge of the environment since context-awareness is a key requirement of ubiquitous computing systems. The knowledge of the user can be specific to the domain of recommendation; or can expand to general user modeling.

From a user modeling perspective, ec(h)o is a knowledge-based recommender system. Similar to Towle and Quin's (2000) proposal, we build explicit models of users and explicit models of objects. However, in ec(h)o the models are not built around specific content but rather ec(h)o uses ontologies at a higher level of abstraction. Users, objects, and environment are annotated with these ontologies. Another significant feature where ec(h)o differs from other knowledge-based recommender systems (for example Entrée, Burke, 2002), is that it does not solicit user’s feedback about the quality of recommendations.

In addition to user modeling, capturing user interests is a central research focus of several disciplines such as information retrieval and information filtering. Most such systems are based on document retrieval where a document’s content is analyzed and explicit user feedback is solicited in order to learn or infer user interests. In our approach, there is no direct feedback from the user. Our prototype can be categorized as a personalized system, as it observes user’s behavior and makes generalizations and predictions about the user based on their interactions (Fink and Kobsa, 2002; Seo and Zhang, 2000). Our approach to observation of user behavior is unobtrusive, similar to approaches to monitoring user browsing patterns (Lieberman, 1995; Mladenic, 1996) or user mouse movement and scrolling behavior (Goecks and Shavlik, 2000).

3.4. SEMANTIC TECHNOLOGIES

Modeling is an integral part of the user modeling by definition. Several types of models are used ranging from simple categories through statistical models, Bayesian networks to formal knowledge models as known in symbolic artificial intelligence (Wahlster and Kobsa, 1989). It is these latter models that potentially benefit the most from semantic web research.
The semantic web initiative (Berners-Lee et al., 2001) aims to achieve a vision of creating a web of meaning. It argues for a set of technologies and techniques that integrates artificial intelligence into the core of the World Wide Web. The cornerstone of semantic web is ontologies (Chandrasekaran, et al., 1991) that provide a mechanism for modeling domains of interest. The formalization is essential for reasoning (Post and Sage, 1990) about the domain. Ontologies and reasoning are basic semantic web technologies that are useful not only in traditional web application domains such as knowledge management, data integration and exchange, or agent coordination but are extensively used in other domains for representation purposes. For example, Baus and colleagues (2002) use ontologies to model the environment in a mobile navigation system. In the Story Fountain system (Mulholland et al., 2004), ontologies are used to describe stories and the domain in which they relate. In order to determine the appropriate domain, reasoning is employed for the selection and organization of resources from which the stories are built.

A main advantage of ontologies, as the concept has developed within semantic web research is the ability to cross-link different domains (Noy and Hafner, 1997). In the area of user interaction this provides us with a clear formalism to connect knowledge about the user, environment, and user aims.

An obstacle in connecting and sharing data, is that often the knowledge captured within an application is at too low a level of abstraction; it is too domain specific. Ontologies provide a mechanism for building several layers of abstraction into the model (Noy and Hafner, 1997).

The assumption we are testing in our approach is that we can use ontologies and semantic web techniques to build interactive systems that successfully operate at higher levels of abstraction. Such a design can be shared across multiple applications. Furthermore, only low-level application-specific logic has to be developed for a new application. Our approach tests this assumption in the context of an audio augmented reality system with a tangible user interface.

4. Design and Rationale

The aims of our design were to develop a ubiquitous computing museum guide that supports *liminal* and engaging play in its user experience; investigates user modeling limited by implicit input from users’ actions; and delivers a wide breadth of information associated with artifacts on exhibit via audio display that is responsive to users’ changing interests. In short, we aimed to investigate less explored avenues in current museum guide systems research including play, embodied interaction, and highly associative as well as contextualized content delivery.

In the last decade, advances in audio museum guides include visitor-driven interaction, access to large collections of supplementary information for museum artifacts, and the development of adaptive and context-aware systems. Many of these advances have come on the heels of innovations in mobile computing
including computer processing capabilities, data storage, connectivity and size. This has culminated in the growing use of PDA devices combined with sensor systems for use as interactive museum guides (Proctor and Tellis, 2003). Yet, outside the domain of museums, for example in the area of games and ubiquitous computing, Björk and his colleagues have identified the need to develop past end-user devices such as mobile phones, personal digital assistants and game consoles (Björk et al., 2002). They argue that we need to better understand how “computational services” augment games situated in real environments. Our design ethnography observations confirmed that museum interactives such as computer kiosks were less used than physical and play-based interactives (Wakkary and Evernden, 2005). In addition, Proctor (Proctor and Tellis, 2003) has found that in museum use PDAs create expectations of a multimedia experience that lessens the relationship between the visitor and the artifacts. As examples, visitors tend to want more of everything yet they quickly lose interest in audio/visual and interactive clips; the visual screen made the moments in-between interactions problematic since if the screen became blank, visitors thought the devices were broken, yet they did not want the screen on all the time since it distracted them from the exhibition. The main point of these findings is that the focus of the visitor is on the experience of the device rather than the experience of the museum.

The anthropologist Genevieve Bell has described museums in terms of cultural ecologies (Bell, 2002). Bell sees the museum visit as a ritual determined by space, people and design. She decomposes the visiting ritual into three observational categories: space, visitors, and interactions and rituals. Different types of museums have different ecologies, for example Bell describes different attributes in each of the observational categories between art museums and science museums. These ecologies are seen to be distinct and supportive of different kinds of museum visits. Bell also describes concepts that are common to all museum ecologies. We have drawn on and extended two of these concepts in developing our approach, liminality and engagement.

Liminality defines museums as places that embody an experience apart from everyday life. Positive museum experiences are transformative, spiritual, and even moving. A museum visitor should be inclined to pause and reflect, thus liminality can be seen to permit a deeper engagement. Engagement is a key concept for museums as people go to museums to learn, however this engagement is often packaged in an entertaining way; museums are a balance between learning and entertainment spaces. It is easy to see how liminality and engagement include ludic experiences in which play and discovery are encouraged. In our adult lives, play is an experience set apart from our everyday activities: Huizinga refers to play as invoking a “magic circle”, a liminal space for games (Huizinga, 1964); Carse describes “deep play” as a profound level of ritualized engagement causing reflection on everyday experiences (Carse, 1987); and psychologist Csikszentmihalyi has described “flow” as a high level of engagement, risk and challenge found in play (Csikszentmihalyi, 1990).
Our aims led us to a design that was inherently minimal and playful. In order to move past the limitations of device-centered approach we developed a tangible interface supported by an audio display, and a user model and adaptive information retrieval system. The tangible interface creates a playful transition between the physical space and the virtual information space of the audio. The audio display creates a virtual context that allowed us to create new layers of engaging experiential spaces such as ambient sounds and conversational information delivery.

Given the limited input and output of our interface, we chose a user model approach to act as a mediator for the visitor. The user model dynamically integrates movement interaction and visitor content selection into initial pre-selected preferences. Based on this dynamic model we could infer potential interests and offer a corresponding range of content choices even as visitors' interests shifted over time. In addition, the use of semantic technologies allowed for coherent and context responsive information retrieval.

While arguably other interface approaches could have been utilized in conjunction with the integrated modeling technique, such as a simple push-button device for input or a mobile text display device for output, such a strategy would be incongruent with our experience design goals. Nevertheless, we designed our user modeling and semantic technologies technique such that it could be easily modified for other interfaces and applications.

The project was informed by ideas of ecologies, like Bell's cultural ecologies and prominently used audio. This combination led us to the name ec(h)o, which is intended to signify the words eco, an abbreviation for the word "ecology", and echo, denoting the acoustic aspects of the project.

4.1. VISITOR SCENARIO

In order for us to better describe the system we developed, we provide below a typical visitor scenario. It should be noted that the scenario describes aspects of the project that are not the focus of discussion in this paper such as soundscapes. The scenario refers to an exhibition about the history and practice of collecting natural history artifacts in Canada at the Canadian Museum of Nature in Ottawa:

Visitors to the Finders Keepers exhibition can use the ec(h)o system as an interactive guide to the exhibition. Visitors using ec(h)o begin by choosing three cards from a set of cards displayed on a table. Each card describes a concept of interest related to the exhibition. The cards include topics such as "aesthetics", "parasites", "scientific technique" and "diversity". A visitor chooses the cards "collecting things," "bigness," and "fauna biology." She gives the cards to an attendant who then gives the visitor a shaped wooden cube that has three colored sides, a rounded bottom for resting on her palm and a wrist leash so the cube can hang from her wrist without her holding it. She is also given a pair of headphones connected to a small, light pouch to be slung over her shoulder. The pouch contains a wireless receiver for audio and a digital tag for position tracking (see Figure 1).

Our visitor moves through the exhibition space. Her movement creates her own dynamic soundscape of ambient sounds. As she passes a collection of animal bones she
hears sounds that suggest the animal's habitat. The immersive ambient sounds provide an audio context for the collection of objects nearby.

As she comes closer to a display exhibiting several artifacts from an archaeological site of the Sigit people, the soundscape fades quietly and the visitor is presented with three audio prefaces in sequence. The first is heard on her left side in a female voice that is jokingly chastising: "Don't chew on that bone!" This is followed by a brief pause and then a second preface is heard in the center in a young male voice that excitedly exclaims: "Talk about a varied diet!" Lastly, a third preface is heard on her right side in a matter-of-fact young female voice: "First dump ... then organize." The audio prefaces are like teasers that correspond to audio objects of greater informational depth.

The visitor chooses the audio preface on the left by holding up the wooden cube in her hand and rotating it to the left. This gesture selects and activates an audio object that is linked to the audio preface of the scolding voice warning against chewing on a bone. The corresponding audio object delivered in the same female voice yet in a relaxed tone is about the degree of tool making on the part of the Sigit people: "Artifact #13 speaks to the active tool making. Here you can actually see the marks from the knives where the bone has been cut. Other indicators include chew marks ... experts are generally able to distinguish between rodent chew marks and carnivore chew marks."

After listening to the audio object, the visitor is presented with a new and related audio preface on her left, and the same prefaces are heard again in the center and to her right. The audio prefaces and objects presented are selected by the system based on the visitor's movements in the exhibition space, previous audio objects selected, and her current topic preferences.

4.2. INTERACTION MODEL

Our interaction model relies on a turn-taking approach generally based on the structure of a conversation. We designed our audio objects in two parts.

---

We use the term "conversation" in the context of the use of conversation analysis to inform HCI design. The idea of using conversation analysis concepts as a structural metaphor for non-speech interfaces is not unique in HCI, see for example (Norman and Thomas, 1990).
prefaces and audio objects: prefaces act as multiple-choice indices for the more detailed telling of the audio object. The tangible user interface provides input for a response to the delivery of prefaces.

The implementation went as follows: eco(h)o offers the visitor three short audio pieces as prefaces. The system is in effect offering three turn-taking possibilities for the visitor. Switching between the stereo channels created localization: we used the left channel audio for the left, right channel audio for the right, and both channels for the center. It is a simple egocentric (Brewster et al., 2003) spatial structure that allows the three prefaces to be distinguishable and an underlying categorization structure to exist. The spatialization was mapped to the tangible user interface for selection. The visitor responds by rotating the wooden cube in his hand and thus selecting a preface. The system delivers the audio object related to the preface. After the delivery of the object, the system again offers three prefaces. The visitor's response is expressed through the gesture selection with the wooden cube. Additionally, the system may be met by no response, because the visitor does not wish to engage the system. The system will then enter into a silent mode. The visitor may also have moved away and the system will then initiate a soundscape.

The prefaces were written to create a sense of surprise and discovery. The audio recordings used a diverse set of voices that were informal in tonality and style. This added to the conversational feel and created an imaginary scene of a virtual cocktail party of natural historians and scientists that followed you through the museum. The audio objects were developed through interviews with museum staff and researchers (Wakkary et al., 2004).

A topic of interest is conceptually represented by each preface or spatial location. The structure is very simple given the limited choices of three options. The navigation is as follows: a visitor is played three prefaces, one to his left, another to his center and the third to his right. He selects the preface on his right side and listens to the linked audio object. On the subsequent turn the visitor hears the same two prefaces he did not select, and again he hears them to his left and to his center. Since he previously chose the preface to his right he now hears a new preface in that location. If the visitor then selects the center preface, on the subsequent turn only that preface is replaced by a new preface in the center position. If a preface has been replayed three times without being selected, it is replaced by a preface linked to an audio object of a completely new topic.

The audio objects are semantically labeled to a range of topics. At the beginning of each interaction cycle, three audio objects are selected based on ranking using several criteria such as current levels of user interests, location, interaction history, etc (see Section 4.4.2). The topics of each object are not explicit to the visitor; rather the consistency and content logic are kept in the background.

In regard to the design process, many of the design choices were made through a series of participatory design workshops and scenarios, details of which have been written in another paper (Wakkary, 2005, in press). For example, the tangible user interface and its implementation as an asymmetrically shaped wooden
cube resulted from these workshops. We also recreated the exhibition environment in our labs; this aided us in the design of the interactive zones and audio display.

4.3. USER MODEL

At the core of the ec(h)o's reasoning module is a user model (Wahlster and Kobsa, 1989) that is continually updated as the user moves through the exhibition space and selects audio objects.

Figure 2 shows an interaction schema of the user model with other modules. There are two main update sources in the system. First, as the user moves through the exhibition the speed of the movement and/or stops in relation to different artifacts provides updates to the user model. The user type is computed based on the speed and uniformity of the user movement. The slowing down and rest points in front of an artifact are interpreted as an interest in concepts represented by the artifact.

The second source of updates to the user model considers a user’s direct interaction when selecting an audio object. In the model this correlates to an increased interest on behalf of the visitor in concepts presented by the audio object and this is reflected in the user’s interaction history.

4.3.1. USER MODEL COMPONENTS

The interaction history is a record of how the user interacts with the augmented museum environment. Two types of events are stored in the interaction history: the user’s movement and user’s selection of objects. The user path through the museum

![Diagram of user model interaction](image)
is stored as discrete time-space points of locations on the path. A second type of information stored in interaction history is the user's selections of audio objects.

*User type* in the museum context is well studied in museum studies (Dean, 1994) and is used in several systems personalizing the user experience (Serrell, 1996; Sparacino, 2002). In the case of ec(h)o, several categorizations were used, for example one user may review almost every artifact on her path, and another user may be more selective and choose artifacts that have only certain concepts. Our categorization of user types is based on Sparacino's work (Sparacino, 2002). It classifies users into three main categories. These categories were validated by our site studies and interviews with staff at various museums:

- The avaricious visitor wants to know and see as much as possible. He is almost sequential, and does not rush;
- The selective visitor explores artifacts that represent certain concepts and is interested in only those concepts;
- The busy visitor does not want to spend much time on a single artifact preferring to stroll through the museum in order to get a general idea of the exhibition.

In ec(h)o, the user type category is not static but is updated every minute. The rules for the type specification consider the location data accumulated within the longer time interval and concepts of previously selected audio objects. 

*User interests* are represented as a set of weighted concepts from the 'concept ontology' (described in Section 4.4). In ec(h)o, each artifact and exhibition is annotated with a set of concepts from the same ontology. The audio objects present a set of particular concepts as well. In each interaction step the system updates the user interests in response to two update channels described above. The update process is described in detail in Section 5.5.

The interaction of the user with artifacts and audio objects is stored in the interaction history that together with the user types are used to infer the user's interests. Several aspects of the update process are parameterized. We discuss the user model parameters and the user model update process in Section 4.5 after we introduce the model for representing content and context in the next section.

### 4.4. INFERENCE-BASED AUDIO OBJECT RETRIEVAL

The audio object retrieval process is performed by the rules that encode multiple object selection criteria. The rules match semantic descriptions of the objects and the museum environment with user information maintained by the user model.

---

1 We use term 'interest' or 'user's interest' when referring to the user model. We use the term 'concept(s) of interest' when referring to the concepts when used to annotate the objects or before they were used to modify the level of corresponding interests in the user model. The relation of the interest in the user model to the concepts in the concept ontology is crucial as it links user model to the model representing content and context as described in the subsequent section.
The content model is based on the semantic description of all the properties of the audio objects and the museum environment that could help us to select visitor and context relevant audio objects. Our ontological model builds significantly on the standard Conceptual Reference Model (CRM) for heritage content developed by CIDOC (Crofts et al., 2003). The CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in the cultural heritage domain. We have also developed several ontologies specifically for the purpose of ec(h)o.

4.4.1. Ontologies for Describing Content and Context

The content of the audio object is not described directly but annotated with three entities: concepts of interest, topics, and themes. The concepts of interest\(^4\) describe the domains that are expressed by the audio objects such as 'evolution', 'behavior', 'lifestyle', 'diversity', and 'habitat'. We realized that it would be impossible to model the content at the actual descriptive level of objects, science and events, so we opted for higher levels of abstraction that in turn provide a unifying degree of formalization for all audio objects in the collection. The starting point for our concept ontology was a set of concepts used by the museum curators at the time of designing the exhibit. We have further extended this initial ontology with concepts identified through analysis of the content of audio objects used in ec(h)o and through interviews with museum researchers (Wakkary et al., 2004). As a result the concept ontology has a flat structure with 39 identified concepts\(^5\). These concepts are mapped to the Dewey Decimal Classification (represented as an ontology), which indirectly gives our concept ontology a hierarchical structure that can be used for drawing inferences.

The concepts play a significant role in the system in linking audio objects and museum artifacts with user interests. The user model (described in the section before) captures a level of user interest in each concept. The audio object retrieval mechanism uses those levels to determine the most appropriate audio objects for the next interaction turn. Similarly, the exhibits are annotated with the concepts that are visually represented in the exhibit (so called visual concepts). When a visitor slows down or stops in the exhibit those visual concepts are used to update the user model.

A topic is a higher-level category for describing several objects within the same exhibit. Objects annotated with different concepts of interest can still have the same topic. Themes are defined as entities that are represented across several exhibitions and are supported by one or more topics; for example, the theme of 'bigness' can include topics such as 'invertebrates' and 'marine biology'.

\(^4\)The concepts of interests represent interests as used by the user model introduced in Section 4.3.

\(^5\)As a result the concrete user model can contain up to 39 interests. However, this is very unlikely as a result of the implemented user model update process described in Section 4.5.
We have used CRM to describe the museum exhibits and artifacts. CRM provides a comprehensive model for describing physical entities, temporal entities and places. We have used CRM to model events and places related to the objects and narratives captured in the audio objects.

Figure 3 shows an example how audio objects ('IN00327' and 'IN00331') are represented in ec(h)o. Both objects exist as independent entities and are related through several ontological relations. The audio object 'IN00327' is annotated with the concepts of interest 'Anatomy' and 'Genus Info.' 'IN00327' has a topic 'From Head to Toe' and supports the theme 'What Can You Tell Me About That'. The audio object 'IN00331' is annotated with the concepts of interest 'Anatomy' and 'Behavior' but is a 'Guide' object (some relations for 'IN00331' were omitted from the picture). The 'Guide' objects differ from the 'Expert' objects by being directly related or referring directly to the artifacts in the exhibition, while the content of 'Expert' objects describes more general knowledge and is reusable in different contexts.

Both objects 'IN00327' and 'IN00331' describe the same museum artifact 'C3-18' representing a 'common dolphin skull' artifact in the exhibition 'E3'. The 'C3-18' is an instance of a 'Biological object' class in the CRM and has many properties that link it to other artifacts in the exhibition (not shown in the picture). The exhibit instance 'E3,' from the exhibit ontology holds the information about the artifacts in the particular exhibit. In addition, 'E3' is annotated with visual concepts 'Collecting', 'Anatomy', 'Scientific Techniques', 'Diversity' and 'Appearances' that are represented visually in this particular exhibit.
Both topics and themes are common tools used by the curators when designing a museum exhibition. In ec(h)o, we use topics and themes in the audio object selection process to support fluency of the interaction between the user and the system. We use CRM referents of place and time period of the artifacts for the selection of the corresponding background sounds appropriate for the presented audio objects.

4.4.2. The Audio Object Selection Process

The audio object selection is based on the ranking of objects. Multiple criteria contribute to the ranking and the audio object with highest ranking is selected. The ranking criteria reflect the dynamic nature of the interaction that is represented by a level of current user interests, previously listened to audio objects and exhibits visited. The system is not intended to be a guide system but rather to enrich the experience of the exhibits and artifacts.

The ranking criteria are listed in Table I. Criterion 1 contributes to audio objects by further describing previously described artifacts while criterion 2 contributes to the ranking of guide audio objects if a previous audio object was also a guide audio object or the user entered a new exhibit. Criteria 3–5 provide for the continuity in the interaction by contributing to the audio objects that elaborate on the same concepts within the same topic and theme. The contribution of criterion 6 is scaled with the current levels of user interests (which change after each interaction step).

The selection process is parameterized and the contribution of each criterion is weighted by its relative importance. Instead of doing extensive testing for weight values the weights were established in consultation with an expert in interactive narrative and storytelling. Table I shows the relative weight distribution for ranking criteria. The only criterion, which we have tested for a range of values, is the contribution between matching concepts of interest in the user model and matching audio object descriptions (Criterion 6, see Section 7 for evaluation and testing results). The remaining values were kept stable. The ‘From’, ‘ec(h)o’, and ‘To’ columns show the absolute values for the weights and ‘%’ columns show the relative contributions to the overall ranking. The ‘From’ column shows the absolute values for the weights when interests in the user model contributed to the object ranking, at a minimum of 13% and ‘To’ column shows the weight values when interests contributed, up to a maximum of 48%.

Guide objects provide for quick orientation in an exhibit with multiple artifacts by directly referring to those artifacts.

The objects score in all criteria, otherwise the percentage contribution is shifted towards the matched criteria. Also, it should be noted that while criteria 1–5 always contribute their full weight the contribution from the criterion 6 varies. The value of criterion 6 shown in the table is the user level of interest in the audio object represented at the maximum level.
Table 1. Weight distribution for object ranking

<table>
<thead>
<tr>
<th>Criteria</th>
<th>From</th>
<th>%</th>
<th>e(h)o</th>
<th>%</th>
<th>To</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Describing artifact previously referred to by the audio object</td>
<td>10</td>
<td>22</td>
<td>10</td>
<td>16</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>2. Object is a 'guide' type of audio object describing an artifact</td>
<td>6</td>
<td>13</td>
<td>6</td>
<td>10</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>3. Continuing in previous topic</td>
<td>8</td>
<td>18</td>
<td>8</td>
<td>13</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>4. Continuing in previous theme</td>
<td>8</td>
<td>18</td>
<td>8</td>
<td>13</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>5. Continuing description of concepts in previous audio object</td>
<td>7</td>
<td>16</td>
<td>7</td>
<td>11</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>6. Concepts in the object match user interest</td>
<td>5.6</td>
<td>13</td>
<td>22</td>
<td>36</td>
<td>36</td>
<td>48</td>
</tr>
</tbody>
</table>

The middle column labeled ‘e(h)o’ shows the actual values used in the final demonstration. The distribution of ranking contributions in the ‘e(h)o’ column is used for audio object selection while a visitor remains within the same exhibit. When users change exhibits only the criteria 2, 3, 4, and 6 are used with the relative distribution of 14, 18, 18, and 50% respectively.

The criteria are implemented in the form of forward chaining rules in which the condition part matches semantic characteristics of each audio object with the interaction history and user interests. If the characteristics of the audio object satisfy the condition, the rule is fired and the ranking for the object is increased. Several rules can be fired for the same audio object. After all rules for the matching audio objects fired and contributed to ranking, the object with the highest ranking is selected.

For example, the rule below represents criterion 1 in Table 1. The rule adds ratings to the audio object that describes the same artifact as the object being replaced. The rule checks whether candidate object ?1n2 describes the same artifact ?a as previous object ?1n1. Next, we make sure that ?1n2 is not an exhibition object but an actual artifact within the exhibition. The PropertyValue is a fact representing semantic descriptions in the form of triples (obtained from the ontologies via transformation when loaded into the inference engine). For brevity, we have also used XML entity descriptions to refer to the namespaces of the ontologies.

It should be noted that the levels of interests in the user model are updated with visual concepts in the new exhibit before they are used to calculate the ranking. As a result, the influence of the context of the new exhibit (in addition to 14% for guide objects) is strongly represented in a 50% contribution from the user model.
(defrule artifact2artifact
  (user-group (user ?u) (group 1))
  (replace (user ?u) (context ?e) (object ?in1)
    (sequence ?seq) (time-chosen ?t))
  (test (neq ?in1 nil))
  (in-context ?a ?e)
  (PropertyValue &psch;#describes ?in1 ?a)
  (PropertyValue &psch;#describes ?in2 ?a)
  (not (PropertyValue &rdf;#type ?a &crm;#exhibition))
  (not (replaced (user ?u) (next-object ?in2))))
=>
  (call ?*object-ratings* addRating ?u ?in1 ?in2
    "artifact-rating" ?t))

For more details about representation and information retrieval aspects in ec(h)o see (Halala et al., 2004) and (Hatala et al., 2005).

4.5. USER MODEL UPDATE PROCESS

The rule-based user model provides a generic structure that enables the system developer to consider several inputs that influence user interests. In addition, the model allows the developer to tune the relative influence of each input using a set of parameters. In ec(h)o, we interpreted two aspects of the user interaction with the system and environment: user movement and audio object selection. Each of these actions has different effect on the model of user interests.

Influence of initial interest selection. A new user starts with a blank user model. In order to bootstrap the model we ask each visitor to indicate initial interests. Prior to entering the exhibition space the user selects a set of cards representing concepts of interest that best match their interests (see Section 4.1 Visitor Scenario). An operator enters the chosen concepts of interest into the user model as user's initial interests and from that point the system evolves the user model through the two update channels described below. The parameter controlling the initial interests' weight can be set by the developer.

Influence of object selection on user interests. In ec(h)o each audio object is described by two concepts of interest: primary and secondary. When a user selects an audio object its primary and secondary concepts of interest are used to update the corresponding user's interests if they were already present in the user model, or they are added to the user model if they were not previously included in the model. As a result, the model is dynamic and the number of interests in the model

---

*In a fully implemented system the same could be achieved automatically by asking the user to select a set of initial interests using a computer kiosk system.*
can vary depending on each user’s individual interaction with the system. The model enables the developer to specify the parameters of how much the primary and secondary concepts of interest in the selected audio object increase the level of corresponding interests in the user model.

Influence of location change (context). The second type of input in ec(h)o is user movement. Each exhibit in ec(h)o is annotated with concepts that are visually represented in the exhibit (visual concepts, Section 4.4.1). For example, an exhibit with photos of pioneer explorers is annotated with a concept of ‘History of Collecting’. When a user stops in a particular location (exhibit), the system interprets this as interest in the visual concept. The user model updates or adds the visual concepts as interests to the model. A set of parameters controls the influence of the visual concepts on the model.

The user model uses a spring model to keep interests balanced. The level of interest is represented by the real number and can range\(^{10}\) from 0 to 10. The sum of all interests never exceeds the value of 30. In the model we consider only positive influences from the user interaction that directly increase the level of some of the interests. When this increase causes an imbalance (the sum is above 30), the implemented spring model proportionally decreases values of other interests. This mechanism supports a highly dynamic nature of the user model and guarantees that only a certain number of interests can have a high value. Another characteristic of this mechanism is that it forces the system to ‘forget’ the ‘older’ interests in favor of recently invoked interests. When the interest value drops below a set threshold during the update process the interest is removed from the model altogether.

5. Implementation

Figure 4 shows the architecture of the ec(h)o system. ec(h)o was implemented and tested in a public exhibition space at Canadian Museum of Nature in Ottawa in March 2004. The system used a combined Radio Frequency Identification (RFID) and optical sensing for position tracking. The system tracked the “x, y” coordinates of each visitor approximately every 1.6s with a spatial resolution of 0.3 m. In terms of hardware, the position tracking system used a separate array of video cameras but all sensing data was integrated.

In addition, we used the “eyes” vision system\(^{11}\) to allow for quicker refresh rates. The vision module included color video cameras connected to desktop computers to cover specified interactive zones. A camera positioned on the ceiling

\(^{10}\)The range of the values for individual interests and their total was selected to achieve a desired proportion between object ranking criteria (see Section 4.4.2).

\(^{11}\)http://www.squishedeyeballs.com
above the artifacts was used to detect the rotation of the cube by visitors within one camera zone in combination with the positioning system.

The sound module consists of a sound-file playback and mixing system driven by the position-tracking module. User position information is provided by the position tracking system and used to dynamically mix the soundscapes the user is immersed in. The sound module uses a custom-designed software mixing system implemented on a single computer. We have developed an authoring environment for mapping sounds to the physical topology of an exhibition. The delivery of the audio objects is through a stereo audio interface using FM radio transmission to portable FM receivers. In our testing environment the system served four simultaneous users. The system scales simply by adding more FM transmitters. The vision and audio delivery systems were developed in our lab using the Max/MSP environment.

The reasoning module was fully implemented with all features described in the previous sections. The real-time nature of the ec(h)o environment was the driving force for the selection of an implementation platform that supported the reasoning engine. As shown in Figure 5, the Jess inference engine is at the center of the reasoning module. We have used DAMLJessKB to load DAML+OIL ontologies into Jess (for details see Kopena and Regli, 2003). DAMLJessKB uses Jena toolkit to convert ontologies into RDF triples that are converted to Jess facts. When converted ontologies are loaded into Jess, the rules representing DAML+OIL semantics infer the missing relations in the RDF graph. This happens at startup time and prepares the system to respond to the input in a real-time fashion. In the development version we embedded the reasoning engine in the Tomcat

12 The zone for the camera depends on the height of the mount and height of the hand handling the cube. For example, the zone diameter for the camera mounted at 4 m can be as wide as 15 m with a wide angle lens.
environment in order to facilitate online editing of knowledge models as shown in Figure 5. However, for the final deployment we used the reasoning engine as a standalone application for performance reasons. All communication with the reasoning engine was accomplished through User Datagram Protocol (UDP) connections.

The user model that forms the significant part of the reasoning engine was implemented using a combination of rules and specific Jess extensions via Java classes to support computation tasks such as object ranking and the spring model calculations used to compute the user interests.

We produced over 600 reusable audio objects at a low level of granularity and annotated them with the ontological information. The average length of each audio object is approximately 15 s. The shortest length is 5 s and the longest 31 s. The prefaces typically last 3 s. A majority of informational and narrative audio objects originated from the interviews with researchers and staff from the Canadian Museum of Nature in Ottawa. We subsequently scripted the objects and used actors for the recordings. For details on the content development see Wakkary et al., 2004.

6. Evaluation

Evaluation of ubiquitous computing systems is extremely complex as these systems ‘bridge the physical and online worlds’ and require seamless navigation between the two, without imposing significant cognitive load on the user (Spasojevic and Kindberg, 2001). There is no agreed upon framework for evaluation of such

---

The only part of the user model that was not continually updated in the final prototype was the user type as the size of the final exhibition did not provide enough supporting data for inferring this information.
systems as known in other domains such as information retrieval (trec.nist.gov) or Robocup (robocup.org). Although Burnet and Rainsford (2001) argue for a hybrid approach combining quantitative and qualitative evaluations situated in a well-defined environment, such as a 'smart room' (Pentland, 1996), many projects use ad-hoc evaluation approaches borrowed from other better established domains. These typically include an analysis of log files for various events and user activities, observing user behavior and conducting user interviews. The small number of test users is also an issue in that it does not allow one to make strong conclusions. For example, the evaluation of the deployment of mobile computing systems in the Exploratorium museum project provided 'existence proofs' for certain reactions and phenomena based on a mix of log files, observations and interviews with a small number of users rather than statistical evidence (Fleck et al., 2002).

We have found Miller and Funk's (2001) view of the problem of evaluation of ubiquitous computing systems from the traditional 'validation' and 'verification' perspective very useful. In regard to validation, we evaluate whether the system performs the functions it was built for based on the requirements specification. Verification tests the system against the reality-checking of user evaluation to see whether the system provides the envisioned benefits.

Following Miller and Funk's approach allowed us to focus our evaluation on areas where we researched novel approaches in adaptive ubiquitous systems. We also avoided the evaluation of aspects of the system that are not well defined or understood. Below we describe three validation steps for two main components of the ec(h)o system, the user model and system response:

1. User model updates: the user and environment models are updated with respect to model modifiers that represent observed user actions in the environment. The user model update mechanism interprets the meaning of the actions as conveyed by the model modifiers to adjust modeled user characteristics, i.e. in our case, the level of user interests. In the user model validation we measure how well the model changes user interests with respect to the input and interaction criteria set for ec(h)o.

2. System response: the second validation we performed evaluates how the system selects audio objects based on the user characteristics with respect to the interaction criteria.

3. User interaction: in this validation step we evaluated user interaction. We evaluated the audio objects characteristics the user selected against the interaction criteria.

In the system verification we obtained qualitative data that measured user experience. We developed questionnaires and performed interviews focusing on user's perception and satisfaction with the system from the perspective of our key research questions.
6.1. VALIDATION OF THE USER MODEL FLEXIBILITY

As mentioned in Section 4.5 the rule-based user model provides a generic structure that enables the system developer to consider several inputs that influence the level of user interests in the user model. These inputs influence initial interest selection, object selection, and location change. In addition, the model allows the developer to tune the relative influence of each input using a set of parameters. The spring model implemented in the user model keeps the rest of the model balanced with the maximum values of each interest capped at a value of 10 and the sum of all interest values at 30.

Each of these actions has a different effect on the user interests. In order to achieve a well-balanced user model we designed a series of tests that evaluated how the rules responded to each type of user action. The second series of tests was designed to balance the relative influence of each type of action in the context of typical user interaction. Both tests were performed in a laboratory setting and they used variations of previously observed user interaction.

We performed a series of tests in which we tested the different combinations of parameters for the maximum interest value (maximum-concept), audio object selection contribution (primary-concept and inferred secondary-concept), location change contribution (visual-concept), and initial user interests (initial-concept). Table II shows the range of values for each parameter tested.

The goal of this test was to find a combination of parameters that would establish the dynamics in the user model with the following characteristics: moderate evolution in user interests when listening to audio objects, significant influence of changing context (visual concepts in exhibits), and protecting the user model from the domination\(^{14}\) of a few concepts. Similarly, in the initiation stage we were looking for the balance between concepts initially selected by a new user and how these are combined with visual concepts when a user enters the first exhibit. It should be noted that the user model is only one component used in the ranking of audio objects; there are other factors that significantly influence object selection and overall interaction (as shown in Section 4.4.2).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Tested values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial-concept</td>
<td>5, 7, 10</td>
</tr>
<tr>
<td>Primary-concept</td>
<td>0.7, 1, 1.5, 2</td>
</tr>
<tr>
<td>Visual-concept</td>
<td>1, 2, 3</td>
</tr>
<tr>
<td>Maximum-concept</td>
<td>8, 10, 12</td>
</tr>
</tbody>
</table>

\(^{14}\)As a result this would prohibit exploration of other concepts of interest and lock the user into a few concepts.
In order to simulate user interaction input we used a fixed sequence of steps captured from the users interacting with an earlier version of eecho in our lab. We evaluated all the combinations of parameters by analyzing the graphical representations of the user model as shown in Figure 6. The figure shows the sequence of steps and evolution of interests in each step. In the first step the user selects three concepts as his or her initial concepts of interest. The circle icon indicates concepts introduced to the model by the visual concepts in the exhibit in which the user enters (Step 2, 11, and 15). In the rest of the steps the user selects audio objects. The square icon indicates a primary concept of interest and a triangle icon denotes a secondary concept of interest in the selected audio object. Figure 6 demonstrates some of the significant features of our user model. A broken line on the left shows how a concept of interest ('Adaptation') introduced to the model via listening to audio objects is being continually reduced as other concepts of interest are increasing in value. The dynamics is highlighted for the other two concepts of interest ('Behavior' in the middle and 'People' on the right). The same applies to the initial concept of interest ('Scientific techniques' furthest to the right) that is not selected and its value is reduced continuously. Figure 6 also shows how the value for a concept of interest ('Anatomy') is not increasing once it has reached the maximum value.

The user model proved to be very flexible and responsive to the parameter values and allowed us to control the dynamics of the interest levels. The combination
of parameters that supported the dynamics of the user model closest to our goals was 7 for the initial-concept, 1.5 for the primary-concept, 2 for the visual-concept, and 10 for the maximum-concept. These values supported the contribution of the user model at the level of 36% of the overall ranking of the audio objects (column 'ec(h)o' in Table 1). We kept these values fixed for the rest of the evaluation.

The selected combination of parameter values is specific to our ec(h)o application and not individual users. It is likely that other applications would require different dynamics. Our model is rule-based and designed to be highly flexible. This not only allows us to modify the values of the parameters that suit the application but also to introduce new parameters into the model as needed.

6.2. USER EXPERIMENT SETUP

We installed the ec(h)o system in an existing exhibition about collecting called 'Finders and Keepers'. The exhibition contains seven exhibits, five of which are booth-type exhibits, each with several dozens of artifacts organized around topics. Two exhibits are open exhibits with larger artifacts such as a mastodon skeleton. For the exhibition, we created three interactive zones: two in booth-type exhibits and one in an open space exhibit.

The formal user evaluation included six participants. The participants had previous experience with interactive museum systems such as docent tours (three participants), interactive kiosks (3), audiotape systems (4), film and video (5), seated and ride-based systems (2) and personal digital assistant systems (2). The test group included two men and four women, aged 25–53-years old.

The testing session for each user started with a brief introduction on the purpose and testing procedure. Participants had an opportunity to interact with the system while one of the researchers accompanied them to explain how to use it. We logged all the interactions of this tutorial phase but as this was a "coached" session we did not include this data in our final evaluation. After this short training session the users had an opportunity to ask questions and seek clarification. Next, participants engaged the system as a typical museum visitor would. Users began by selecting their initial concepts of interest and they were then left alone to freely explore the exhibition. We logged all interactions with the system and used this data for the evaluation of the system described in the following sections. After the main testing session, the users were asked to complete a questionnaire. Finally, we conducted and videotaped a semi-structured interview with each participant.

In addition to the six users we tested the system with two expert reviewers. These experts included a senior researcher and senior interaction designer from the museum. Both were familiar with the exhibit and its underlying concepts. The experts tested the system for an extended period of time with specific focus on the depth of the content and meaningfulness of the interaction. After each of the
expert testing sessions we discussed the issues the experts wanted to clarify. Finally, they provided an extensive written report on the system performance.

Table III shows the characteristics of each user session: the total length of the interaction, number of interaction steps, number of selected and listened to audio objects, and number of location changes. As can be seen from Table III the number of location changes for User3 and User6 are exceptionally high. After examining the log files we found that the system repeatedly registered the single event of entering the same exhibit. This may have been caused by either the user moving along the exhibit boundary or by an error in the position-tracking module\(^{15}\). As explained in the previous section, this event caused the user model to be updated with the concepts represented by the exhibit (visual concepts), which skewed the object selection process towards those concepts. Therefore we did not include these two users in our evaluation data.

### 6.3. VALIDATION OF THE SYSTEM RESPONSE (OBJECT SELECTION)

In section 6.1 we showed how the system interprets the user actions and how user actions are used to update the user model and specifically the level of user interests. In this section we present our results of the recommender part of the system that selects audio objects to be offered to the user.

To evaluate the system response capabilities we have used interaction criteria. The level of fulfillment of interaction criteria can be observed from the audio objects offered to the user at each interaction step. To measure the system performance with respect to interaction criteria we defined three characteristics: variety, sustained focus, and evolution. These characteristics measure semantic relationships between offered audio objects with respect to concepts these audio objects represent.

In ec(h)o, at each interaction step three objects are offered $O_{s1}$, $O_{s2}$, and $O_{s3}$. Each object is annotated with a primary and secondary concept of interest it represents $P_{s1}$, $P_{s2}$, $P_{s3}$ and $S_{s1}$, $S_{s2}$, $S_{s3}$ respectively. If we define two sets $P_s = \{p|$
unique interests in \{P_1, P_2, P_3\}, S_s = |s| is unique interest in \{S_s1, S_s2, S_s3\}, and \|M\| denotes the number of elements in the set M then we can define three criteria as follows:

Variety – describes the richness of choices for further interaction at each interaction step. The variety is a basic mean to put users in control of selecting topics of further interaction. It also compensates for an inherent inaccuracy of user interest modeling by providing multiple alternatives. Formally, we define variety in interaction step s as \( Var_s \)

\[
Var_s = c_1 \|P_s\| + c_2 \|S_s - P_s\|
\]

where we set \( c_1 = 1 \) and \( c_2 = 0.5 \). In case of ce(h)o \( Var_s \) can range from (0, 4.5) so we scaled it to (0, 1) for a clearer comparison.

Sustained focus – An ability of the system to sustain the focus on particular interests. Mono-topical systems provide a maximum degree of sustained focus but do not follow shifting user interests. On the other side of the spectrum are systems selecting topics randomly where the sustained focus cannot be reasonably evaluated.

\[
Susi_{s+1} = c_1 \|P_{s+1} \cap P_s\| + c_2 \|P_{s+1} \cap S_s\| + c_2 \|S_{s+1} \cap P_s\| + c_3 \|S_{s+1} \cap S_s\|
\]

where we set \( c_1 = 1, c_2 = 0.5 \) and \( c_3 = 0.25 \). In case of ce(o) \( Susi_{s+1} \) can range from (0, 6.75) so we scaled it to (0, 1) for a clearer comparison.

Evolution – An ability of the system to follow shifting user interests during interaction with the system. Adaptive systems have an ability to continually shift the focus of the interaction by continuously monitoring user’s interaction. We have defined evolution as the weighted number of new concepts introduced between two steps in the interaction.

\[
Evols_{s+1} = c_1 \|P_{s+1} - (P_s \cup S_s)\| + c_2 \|S_{s+1} - (P_s \cup S_s)\|
\]

where we set \( c_1 = 1 \) and \( c_2 = 0.5 \). In case of ce(h)o \( Evols_{s+1} \) can range from (0, 4.5) so we scaled it to (0, 1) for a clearer comparison.

Table IV shows the values of the proposed evaluation characteristics when applied to the mockup data. The rows labeled as primary and secondary concepts represent concepts of interest for three hypothetical audio objects offered to the user. The values in columns 1–10 were chosen to show how different combinations of concepts affect the three measurements. As defined above, variety is measured for each interaction step and has a value of 0 if all concepts are identical (e.g. column 1) and a value of 1 if all values are unique (e.g. column 11). The sustained focus in a particular interaction step is based on the values in this step and the previous step. The sustained focus measures how many concepts from the previous step are repeated in the next step. This information is weighted differently for primary and secondary concepts being repeated as either primary or secondary concepts (columns 7, 8, and 9 demonstrate this clearly). The evolution is also computed from the current and previous interaction step and it captures how many
new concepts were introduced at the primary and secondary levels. It is not a mere complement of the sustained focus as can be seen from columns 6, 7, 8, and 9. Finally, the last three columns in Table IV shows a more realistic distribution of concepts in offered audio objects. In column 12 concepts ‘i’ and ‘j’ are repeated as primary concepts, concept ‘m’ is repeated as a secondary concept and concept ‘i’ is repeated also as a secondary concept. In column 13 two concepts are repeated: ‘i’ as a primary and ‘p’ as a secondary concept.

We have calculated the sustained focus, evolution and variety for each user interaction. Figure 7 shows actual results for one user (graphs for other users show the same trends). The horizontal axis represents interaction steps that trigger object selection. These can be either the user entering the exhibition zone (step number is circled) or the user making a selection of an audio object. When a user enters the space three new audio objects are offered. After a user makes a selection the selected object is replaced with the new one and possibly a non-selected object is replaced if it had already been offered three times.

In Figure 7 we can observe that the system supports high variety of objects in each step without significant changes between the interaction steps. However, trend lines for sustained focus and evolution demonstrate significant changes at the steps representing a change of the exhibit zone. In these points the sustained focus factor decreases significantly indicating that objects offered in the new location represent new topics of interest from those offered in the previous location. This system behavior reflects our selection of the weights established in Section 6.1, specifically the weight for visual-concept, giving a strong influence of the context on the user model. Once the user stays in the same interaction zone the sustained focus increases reflecting continual changes in the user model. The trend changes in the evolution characteristic are caused by the same decision.

Because the changes in the exhibit location caused such significant differences we separated the statistical processing for the ‘location-change’ steps from the ‘object-selection’ steps. Table V shows the statistical values for all three characteristics as obtained from six test subjects.
Table V. Statistical values for variety, sustained focus and evolution

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Selection</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AVG</td>
<td>STD</td>
<td>AVG</td>
</tr>
<tr>
<td>Variety</td>
<td>0.73</td>
<td>0.18</td>
<td>0.77</td>
</tr>
<tr>
<td>Sustained focus</td>
<td>0.50</td>
<td>0.21</td>
<td>0.58</td>
</tr>
<tr>
<td>Evolution</td>
<td>0.30</td>
<td>0.24</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Based on the values in Table V we can conclude that the system offers highly variable objects when users change location and the variety increases as users continue the interaction in a particular location. The high variety during the object selection steps is supported while the system maintains the focus on the concepts of interest as expressed in the user model. The low value of evolution during the object selection stage indicates the continual change in topics offered corresponding to the modest changes in the user model.

This behavior matches our expectations. As described in Section 5.3.2, several ranking criteria are combined to select audio objects offered in the next step.
6.4. EVALUATION OF USER INTERACTION

While in the previous section we evaluated the system's ability to respond in the manner corresponding with our interaction criteria, in this section we examine how users interacted with the system.

As presented in the previous section we have tuned the system to favor sustained focus over evolution. However, a high level of variety enabled users to 'defy' the sustained focus of the interaction by selecting audio objects with newly introduced concepts of interest. Figure 8 shows the dynamics of how the system introduced new concepts of interest and how the users explored those concepts via object selection in the course of interaction. The horizontal axis represents interaction steps and vertical axis represents a cumulative number of concepts of interest introduced and explored up to that interaction step. The zero value for the number of concepts selected in step one is due to the fact that users did not select any object before moving into another exhibit.

The graphs in Figure 8 shows that at the beginning the system introduces new concepts at a more rapid pace. At the same time the user explores objects (and concepts of interest) rapidly until a point is reached where the user explores some of the concepts in more depth (Steps 8–12 and 14–20). Although the absolute values differ between users we have found a similar pattern is present for all users.

Figure 9 shows the percentage of selected concepts of interest by individual users relative to the number of concepts of interest introduced via offered audio objects. The graph shows that after initial steps users quickly converge to a stable
proportion of the selected concepts of interest in the range of 30–70% of concepts offered.

It is difficult to speculate whether with ongoing interaction the level of concept exploration the users reached would remain at a constant level. Theoretically, as the number of concepts of interest in our system is limited to 39 concepts, the users have an opportunity to explore all of them. On the other hand, as we can see from the available data, users tend to explore certain concepts in more detail.

6.5. VERIFICATION: EVALUATION OF USER EXPERIENCE

User experience was evaluated through observation during the sessions, a questionnaire, and a semi-structured interview. The questionnaire included sixty-three questions that assessed user experience related to the overall reaction to the system, the user interface, learning how to use the system, perceptions of the system's performance, the experience of the content, and degree of navigation and control. The questionnaire also provided for open-ended written comments. Majority of the questions were on a Likert scale. Throughout the questionnaire, and especially during the semi-structured interviews we looked for an overall qualitative assessment of the experience based on Bell's ecological components of *liminality* and *engagement* (Bell, 2002). For a summary of the questionnaire results see Table VI.

Participants found the system enjoyable and stimulating, perhaps in part due to its novelty. The general sense of satisfaction was split between those participants who liked the playful approach and those who did not. While our sample was small we noted a clear age difference in that the “younger” participants rated satisfaction higher based on their liking of the playful approach (this was confirmed in the semi-structured interviews).
Table VI. Summary of the questionnaire results on user experience (n=6; 63 questions on Likert scale of 1-5 (being best)).

<table>
<thead>
<tr>
<th>Categories</th>
<th>Average</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall reaction (five questions including “terrible-wonderful; difficult-easy”)</td>
<td>3.60</td>
<td>0.78</td>
</tr>
<tr>
<td>Tangible user interface (seven questions including “uncomfortable-comfortable; difficult-easy to manipulate; annoying-enjoyable”)</td>
<td>4.24</td>
<td>0.50</td>
</tr>
<tr>
<td>Headset (two questions including “comfortable-uncomfortable to wear”)</td>
<td>2.92</td>
<td>0.12</td>
</tr>
<tr>
<td>Learning curve for the system (eight questions including “difficult-easy to get started; risky-safe to explore features; unclear-clear feedback”)</td>
<td>4.07</td>
<td>0.36</td>
</tr>
<tr>
<td>Perception of system performance (eight questions including “slow-fast system response; never-always reliable”)</td>
<td>3.83</td>
<td>0.39</td>
</tr>
<tr>
<td>Quality of the content (fifteen questions including: “uninformative-informative; generalized-customized for me; rigid-playful; predictable-surprising”)</td>
<td>3.78</td>
<td>0.52</td>
</tr>
<tr>
<td>Quality of the audio experience (nine questions including “confusing-clear; mechanical-human-like; wasteful-valuable”)</td>
<td>3.67</td>
<td>0.30</td>
</tr>
<tr>
<td>Navigation and control (eight questions including “never-always able to navigate in an efficient way; always-never found myself lost in the system; always-never found myself uncertain of system state”)</td>
<td>3.23</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Among the factors that stood out as most positive for the participants was that the cube and audio delivery were seen as playful. The open-ended written comments and semi-structured interviews made this point clear as well. The tangible user interface was well received especially in terms of ergonomics and ease of use. This was not a surprise to us since our early testing and participatory design sessions provided us with considerable feedback, especially on ease of use and enjoyment. We went through several iterations of the wooden cube selecting the lightest wood we could find (balsa wood) and going through several form factors tested against different hand sizes. This may have also resulted in the fact that learning to use the interface and navigation were rated highly and participants felt the system had a low learning curve and so it was easy to get started. It should be stated that we provided a short tutorial on the system at the beginning of each evaluation (see Section 6.2) but nevertheless this feedback is encouraging. Interestingly, the audio content was perceived to be both accurate and clear. The issue of trust and delivery style is an area to further investigate. Since we collected the information directly from scientists and staff at the museum rather than a more generic source we wonder if this contributed in part to this result (Wakkary et al.,
These results lead us to believe that the system meets or satisfies many of the current advances of museum guide systems.

The questionnaire did point out challenges and areas for further research. Some things we expected such as the headphones were uncomfortable, yet to such a degree that we are currently rethinking the tradeoff between personalized spatial audio and use of headphones. Other results point to a threshold in the balance between levels of abstraction and local information. Since visitors had difficulties at time connecting what they were listening to and what was in front of them (in part this was an inherent challenge in the exhibition since the display cases had dozens to over a hundred artifacts). In many respects this confirms our finding that the ontological approach did not provide a clear enough contextual link between the artifacts and the audio information. In addition, we see both a threshold point in play versus focused attention on the exhibit in that the question relating to the content asking if it was "distractive-synergistic" scored 2.83. This raises the issue of balance in play and the possibility to shift attention away from the environment rather than play as a means of further exploring the environment.

In an open-ended question in the questionnaire and through the semi-structured interviews we explored the issues of liminal play and engagement. The results here are quite clear that play was a critical experiential factor in using the system. It was often remarked how the experience was similar to a game:

"At first it felt a little bit strange, especially holding this cube that looked like a child's toy, and I felt a little bit awkward about doing that, but I got over that pretty quickly. The whole system to me felt a lot like a game. I mean I got lost in it. I found myself spending a lot [more] time in a particular area than I normally would. And just the challenge of waiting to hear what was next, what the little choice of three was going to be. Yeah ... So I found it over all engaging, it was fun, and it was very game-like." (Participant 5)

The playfulness did in most instances suggest a quality of engagement that led to learning even through diverse types of museum visits including the visitor who browses through quickly but is still looking to be engaged, to the repeat visitor who experiences the audio information differently each time:

"I learned a lot and well you know I'm a scientist here, and I think anybody going through, even people who are in a real rush, are going to pick up some interesting facts going through. And ... I mean, that was good, the text was great and was short enough that somebody in a rush is still going to catch the whole thing. And there wasn't much delay really. I mean once you showed your cube it came up pretty fast, and that is important with museum-goers. I think museum-goers don't stand and spend a bunch of time in one spot so it has to be something that comes up pretty quickly." (Participant 4)

As mentioned earlier, there is a threshold between play in support of the exhibit on display and play with the system that can be an end in itself and even a distraction. For example, one participant occasionally focused more attention on playing with the system than the exhibition due to her enthusiasm for the game-
like quality. In addition, people respond to play differently and can be argued to belong to different types of players (Bartle, 1990). One participant would have preferred a more serious and “non-playful” approach. In this case the playfulness and short length of the audio was seen as anecdotal rather than serious and scholarly.

7. Discussion

At the outset of this paper we acknowledged the challenge to capture the larger context through user modeling, particularly in ubiquitous and mobile computing applications. No doubt Fischer poses the problematic as a description of an ongoing research program than a question that a single project can address (Fischer, 2001). Nevertheless, our strategies along this front included the sensing and inference based on visitor movement, like many other systems, however, we also utilized a mixed criteria, combining ranking of concepts of interest based on direct user selection of audio objects mixed with visual concepts that we mapped to the context (see Sections 4.3.2 and 6.1). Our aim here was to allow for the possibility of new interests to form externally through the context. As it turned out, in analyzing the participants’ selections of audio objects based on the interaction criterion of evolution (see Section 6.3), significant changes occurred less through user selection (this was always possible since we maintained high degree of variability in concepts at all times) than from visitors moving to another exhibit. The criterion of evolution can be said to evaluate internal influences (user’s reflection on content) and external influences (user’s reflection on context). This was possible given our aim to consider user interest as dynamic and evolving based on the interaction with the environment. In fact, we earlier stated that we do not see our system as a museum guide, recommending things based on what people like or know at the outset of a visit, rather we see it as a way to provide enrichment to the ongoing experience of the exhibit and artifacts.

The specific problem we stated at the outset of the paper was how to support the fuller experience design goals as well as functionality with an integrated modeling technique and use of semantic technologies in combination with an audio augmented reality and tangible user interface approach.

In regard to functionality, the user experience results show that ec(h)o was extremely easy to use and quick to learn, and the overall system performed well (see Section 6.5). The validation of the ec(h)o components, namely user model and object selection, showed that these performed at the required level of accuracy and flexibility. While we did not perform a comparative test with other systems, in the verification it was clear that participants had experience with many different museum based systems (see Section 6.2) and we can expect that comparisons were made with past experiences in evaluation of ease-of-use, learning curve, and performance.

In regard to the experience design goals of play and discovery, we feel our integrated modeling approach implemented two techniques to facilitate wider
exploration and the discovery of new topics of interests and the ability to make new connections among topics and artifacts. The first being the aim of keeping interests balanced such that a given topic or set of topics does not dominate and prevent exploration of new topics, for this we used a spring model to proportionately moderate levels of interest (see Section 6.1). As we stated, it is important that the user model learns to “forget older interests” so that newer ones can be invoked. The second technique is to maintain a high level of variety of primary and secondary interests among the objects presented. This affords greater opportunity for the user to evolve his or her interest through a reflection on content as discussed above (see Section 6.3). These techniques contribute to the goal of establishing dynamics in the user model that support exploration and discovery of new interests through moderating evolution in the user interests, maintaining significant influence of changing context (when a visitor moves to another exhibit), and protecting against the domination of a few concepts that would choke off exploration.

We introduced the evaluation of system response or in our case, object selection based on interaction criteria of variety, sustained focus, and evolution. We’ve found these terms useful in the discussion above and we can say that we can measure variety, and rationalize it together with evolution as dependent factors in exploration and discovery of new user interests through interaction. Sustained focus is less clear of a measure at this stage and something we will investigate in future research.

There are cautions in our findings. The first is designers must strike a balance or they run the risk of users engrossed in the playing with the system at the expense of interacting with their surroundings, as one participant commented happened to her periodically. The second caution stems from the results that indicate that visitors had difficulties at times connecting what they were listening to and what was in front of them. It may be that the system did not always provide a coherent story, a resulting tradeoff due to its dynamic nature. Nevertheless, a much richer model of discourse and storytelling could be an option to pursue. In addition, users in the museum settings are significantly connected with concrete artifacts while ec(h)o experimented with the idea of the connection between artifacts and audio objects residing at a higher ontological level. The results indicate that either a much richer model is needed or the hypothesis of linking objects at higher abstract ontological levels is not suitable for ubiquitous context-aware applications or it has to be combined with other approaches.

8. Conclusion and Future Work

ec(h)o is an augmented audio reality system for museum visitors that was developed and tested for the Canadian Museum of Nature in Ottawa. In ec(h)o we tested the feasibility of audio display and a tangible user interface for ubiquitous computing systems – one that encourages an experience of play and engagement. The interface uses audio as the only channel to deliver short audio objects.
We have built several ontologies that richly described the museum environment and artifacts, audio objects and user interests. The knowledge-based recommender system builds a dynamic user model based on user choices and user movement through the exhibit and recommends audio objects to the user.

The findings of this project are positive while also calling for more research in several areas. First, we found that it is possible to build a highly flexible and accurate user model and recommender system built on information observed from user interaction that supports play and discovery as well as functionality. Ontologies and rule-based approaches proved to be a strong combination for developing such systems, yet some museum visitors are looking for more coherent stories that are highly contextualized. The ontological approach did not prove satisfactory and either more extensive knowledge engineering is needed or it has to be combined with stronger narration or discourse models. As museums are highly social places, another area that needs more research is extending the system with support for groups and group interaction.

Acknowledgements

Work presented in this paper was supported by a Canarie Inc. grant under the E-Content program. Authors would especially like to thank Dr. Mark Graham and his colleagues at the Canadian Museum of Nature in Ottawa for their enthusiastic support of this project. We would also like to thank the participants of our several workshops that contributed to the development of the project, and namely our colleagues and students. Kenneth Newby. Dale Evernden, Leila Kalantari, Doreen Leo, Gilly Mah, Robb Lovell, Mark Brady, Jordan Willms and Phil Thomson for their contributions.

References


**Authors’ Vitae**

Marek Hatala

*School of Interactive Arts and Technology, Simon Fraser University, 2400 Central City, 10153 King George Highway, Surrey, BC, Canada, V3T 2W1* Dr. Marek
Halala is Associate Professor at the School of Interactive Arts and Technology at Simon Fraser University and a director of the Laboratory for Ontological Research. Dr. Halala received his MSc in Computer Science and PhD in Cybernetics and Artificial Intelligence from the Technical University of Kosice. His primary interests lie in areas of knowledge representation, ontologies and semantic web, user modeling, intelligent information retrieval, organizational learning and eLearning. His current research looks at how semantic technologies can be applied to achieve interoperability in highly distributed and heterogeneous environments, what are the social and technical aspects of building a distributed trust infrastructures, and what role user and user group modeling can play in interactive and ubiquitous environments.

Ron Wakkary
School of Interactive Arts and Technology, Simon Fraser University, 2400 Central City, 10153 King George Highway, Surrey, BC, Canada, V3T 2V1 Prof. Ron Wakkary is Associate Professor at the School of Interactive Arts and Technology at Simon Fraser University. Prof. Wakkary received his BFA from the Nova Scotia College of Art and Design, MFA from the State University of New York at Stony Brook, and is a Ph.D. candidate in Interaction Design at the University of Plymouth, U.K. The ec(h)o project serves as a case study in his thesis on an analytical framework for interaction design. His primary research interests lie in design of interactive systems in the area of ubiquitous computing including responsive environments, personal technologies and tangible user interfaces, and the study of interaction design related methods and practice.
Rules and ontologies in support of real-time ubiquitous application

Marek Hatala*, Ron Wakkary, Leila Kalantari

School of Interactive Arts and Technology, Simon Fraser University, 2400 Central City, 10153 King George Highway, Surrey, BC, Canada V3T 2V1

Received 6 May 2005; received in revised form 11 May 2005; accepted 12 May 2005

Abstract

The focus of this paper is the practical evaluation of the challenges and capabilities of combination of ontologies and rules in the context of real-time ubiquitous application. The ec(h)o project designed a platform to create a museum experience that consists of a physical installation and an interactive virtual layer of three-dimensional soundscapes that are physically mapped to the museum displays. The retrieval mechanism is built on the user model and conceptual descriptions of sound objects and museum artifacts. The rule-based user model was specifically designed to work in environments where the rich semantic descriptions are available. The retrieval criteria are represented as inference rules that combine knowledge from psychoacoustics and cognitive domains with compositional aspects of interaction. Evaluation results both from the laboratory and museum deployment testing are presented together with the end user usability evaluations. We also summarize our findings in the lessons learned that provide a transferable generic knowledge for similar type of applications. The ec(h)o proved that ontologies and rules provide an excellent platform for building a highly-responsive context-aware interactive application.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Ontologies; Rule-based systems; User modeling; Context aware; Augmented reality; Audio; Museum guide

1. Introduction

Audio museum guides have existed for some time as a means of overcoming the scheduling inflexibility of group tours by museum docents. While beneficial in many respects, the audio guides are limited by their linear, ear sequence and non-interactive structure. Bederson [3] developed a prototype utilizing portable mini-disc players and an infra-red system to allow museum visitors to explore at their own pace and sequence. As museum visitors approached artifacts on display, relevant audio information would be triggered on the mini-disc player and heard through headphones. Hyperaudio [16] provided visitors with palmtop computers and developed specific user models for adaptive systems within a museum setting. MEG [2] is a portable digital museum guide for the Experience Music Project in
Scattle that allows visitors 20 hours of audio and video on demand. Visitors make their selections either by use of the keyboard within the PDA device or by pointing the device at transmitters located adjacent to artifacts.

In the previous works, the relationship of the digital content to the artifacts is either pre-planned and fixed, or the digital content is not networked and limited to the local device; in some cases both limits are true. ec(h)o employs a semantic web approach to the museum’s digital content, thus it is networked, dynamic, and user-driven. The interface of ec(h)o does not rely on portable computing devices; rather it utilizes a combination of gesture and object manipulation recognized by a vision system.

The dynamic and user-driven nature of ec(h)o requires a highly responsive retrieval mechanism with a criteria defined by psychoacoustics, content, and composition domains. The retrieval mechanism is based on a user model that is continually updated as a visitor moves through the exhibition and listens to sound objects. The criteria are represented by rules operating on the ontological descriptions of sound objects, museum artifacts, and user interests.

One of the main goals of ec(h)o is to achieve an enhanced experience for the museum visitors without inserting an extra layer of technology between the visitor and the museum exhibit. Two mechanisms contribute to an accurate retrieval of sound objects in ec(h)o: the user model and ontology descriptions of objects.

With the development of the semantic web [4] the use of ontologies as a formalism to describe knowledge and information in a way that can be shared on the web is becoming common. Adoption of the standard for the ontology web language (OWL) [21] is propelling this trend toward large scale application in different domains. However, the utility of the ontologies is limited by the processing mechanisms that are smoothly integrated with this form of representation. Therefore there is an effort on the way to formalize the logic layer for ontologies. The semantic web rule language (SWRL) [21] is proposed as an important step in this direction, building on the experience of the previous work on RuleML [5]. Eventually the availability of standardized rule language for the semantic web will make it possible to use both ontologies and rules as a basis for innovative applications that are connected to the semantic web. The understanding of capabilities and implications of this combination will be essential for successful deployment and adoption of these technologies. This paper aims at addressing some of these issues through the development of a ubiquitous system with some extreme requirements testing the capabilities of the emerging technological platform.

The paper is organized as follows. First we present the ec(h)o architecture and then we describe ontologies used in the ec(h)o. Section 4 describes the user model and Section 5 outlines the retrieval mechanisms for sound objects. Before we show the results of the evaluation in Section 7 we describe the implementation challenges and lessons learned in Section 6.

2. ec(h)o Architecture

The platform for ec(h)o is an integrated audio, vision, and location tracking system installed as an augmentation of an existing museum exhibition installation. The platform is designed to create a museum experience that consists of a physical installation, an interactive layer of three-dimensional soundscapes that are physically mapped to museum displays, and the overall exhibition installation.

Each soundscapce consists of zones of ambient sound and “soundmarks” generated by dynamic audio data that relates to the artifacts the visitor is experiencing. The soundscapes change based on the position of the visitor in the space, their past history with viewing the artifacts, and their individual interests in relation to the museum collection. To achieve this type of audio experience the overall system must be integrated with a position tracking system that has a frequent update cycle and a high level of spatial resolution. A pattern of the user’s movement can indicate the type of museum visitor [19] as well as user intentions [17].

When the user stops in front of an artifact, she is presented with three sound objects spatially positioned to the left, center, and right. By way of a gesture-based interaction, the visitor can interact with a single artifact or multiple artifacts in order to listen to related audio information. The audio delivery is dynamic and generated by agent-assisted searches inferred by past interactions, histories, and individual interests. The source for the audio-data is digital objects. In the case of ec(h)o, we developed a large sample set of digital objects that originated from the partner
museums. These digital objects were used to populate the network of object repositories.

The ec(h)o architecture (Fig. 1) consists of four independently functioning modules: position tracking module, vision module, sound delivery module, and reasoning module. Two main types of events trigger the communication between the modules: the user's movement through the exhibition space and the user's explicit selection of the sound objects.

3. Semantic description of objects

We have identified two types of information as essential for ec(h)o:

- the content description of the user interests (user model), sound objects, and museum artifacts, and
- psychoacoustics and sound characteristics of the sound objects.

3.1. Ontologies for describing content

The ec(h)o interaction model is based on the semantic description of the content of the sound objects. We have developed a sound object ontology describing objects with several properties. As the ability to link to other museum collections is an important feature of ec(h)o, our ontology builds significantly on the standard conceptual reference model (CRM) for heritage content developed by CIDOC [7]. The CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation. To describe sound objects we use CRM TemporalEntity concept for modeling periods and events and Place for modeling locations. We describe museum artifacts using the full CRM model.

The content of the sound object is not described directly but annotated with three entities: concepts, topics, and themes. The concepts describe the domains that are expressed by the sound object such as evolution, behaviour, lifestyle, diversity, habitat, etc. Since the collections in individual museums are different, so are the concept maps describing these collections. A topic is a more abstract entity that is represented by several concepts, such as botany, invertebrates, marine biology, etc. To facilitate the mappings between topic ontologies in individual museums we have mapped the topics to the Dewey decimal classification [8] whenever

<table>
<thead>
<tr>
<th>Property</th>
<th>Domain</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasTheme</td>
<td>SoundObject</td>
<td>Theme</td>
</tr>
<tr>
<td>hasTopic</td>
<td>SoundObject</td>
<td>Topic</td>
</tr>
<tr>
<td>hasPrimaryConcept</td>
<td>SoundObject</td>
<td>Concept_of_interest</td>
</tr>
<tr>
<td>hasSecondaryConcept</td>
<td>SoundObject</td>
<td>Concept_of_interest</td>
</tr>
<tr>
<td>relatesToTemporalEntity</td>
<td>SoundObject</td>
<td>CRM_TemporalEntity</td>
</tr>
<tr>
<td>relatesToPlace</td>
<td>SoundObject</td>
<td>CRMPlace</td>
</tr>
<tr>
<td>describesArtifact</td>
<td>SoundObject</td>
<td>MuseumArtifact</td>
</tr>
</tbody>
</table>
possible. Finally, themes are defined as entities supported by one or more topics, for example, the theme of "bigness" in invertebrates and marine biology.

Table 1 shows content related properties with their domains and ranges.

In Fig. 2 the sound object ‘IN00327’ is annotated with concepts ‘Anatomy’ and ‘Genus Info’, has a topic


The ontologies for e(h)o were modeled in DAML + OIL. The DAML + OIL representation1 of the IN000327 audio object is shown below

In e(h)o the ontological concepts are transformed into the Jess facts that represent RDF triples (see imple-

1 For readability we use XML entities to refer to namespaces in this paper. For example, &psch; refers to the namespace http://echo.iat.sfu.ca/owl/psychoacoustic.daml, other references are self-explanatory.

---

\[ \text{From Head to Toe, and supports the theme 'What Can You Tell Me About That'. The sound object 'IN00327' describes the artifact 'C3-18' that is modeled as an instance of 'Biological object' type in the CRM model described by the 'Common dolphin skull' object. The exhibit 'E3' from the exhibit ontology holds the in-} \]
mentation section for details). The above DAML + OIL description of the audio object IN000327 is represented with the following facts (with PropertyValue being a fact name used for all RDF triples).

(PropertyValue rdf:type &psch;«IN00327 &psch; ttInfoNarrat\n &psch;«hasPrimaryConcept &psch;«IN00327 &psch;#Anatomy)
(PropertyValue &psch;«hasSecondaryConcept &psch;«IN00327 &concept#Genus_info)
(PropertyValue &psch;«hasTopic &psch;«IN00327 &t:opi c«From_head_to_t:oe)
(PropertyValue &psch;«hasRecord &psch;«IN00327 http://192.168.0.103/sound_objects/00327.mp3)
(PropertyValue &psch;«hasPreface &psch;«IN00327 &psch;#P00327_1)
(PropertyValue &psch;«hasTheme &psch;«IN00327 &theme;«What_can_you_tell_me_about_that)

For details on creation of content and related ontologies see [23].

3.2. Psychoacoustics and sound characteristics ontologies

The auditory interface of echo follows an ecological approach to the sound composition. It provides the basic mechanisms of navigation and orientation within the information space. Three areas are taken into account: psychoacoustic, cognitive, and compositional problems in the construction of a meaningful and engaging interactive audible display. Psychoacoustic characteristics of the ecological balance include spectral balancing of audible layers. Cognitive aspects of listening are represented by content-based criteria. Compositional aspects are addressed in the form of the orchestration of an ambient informational soundscape of immersion and flow that allows for the interactive involvement of the visitor.

Table 2 shows the psychoacoustics ontology that defines the characteristic of the sound objects that are used by the composition rules.

Table 2
Psychoacoustic properties for the Sound Object

<table>
<thead>
<tr>
<th>Property</th>
<th>Domain</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasSpectralDensityCenter</td>
<td>SoundObject</td>
<td>&lt;Number&gt;</td>
</tr>
<tr>
<td>hasSpectralDensityWidth</td>
<td>SoundObject</td>
<td>&lt;Number&gt;</td>
</tr>
<tr>
<td>hasBandwidth</td>
<td>SoundObject</td>
<td>&lt;Number&gt;</td>
</tr>
<tr>
<td>relatesToEnvironment</td>
<td>SoundObject</td>
<td>Physical.Environment</td>
</tr>
<tr>
<td>relatesToEvent</td>
<td>SoundObject</td>
<td>CRM_Event</td>
</tr>
<tr>
<td>hasSource</td>
<td>SoundObject</td>
<td>SourceTypeValue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(e.g. AnimalSound, HumanEnvironmentSound)</td>
</tr>
</tbody>
</table>

4. The user model

In the core of the echo's reasoning module is a user model [22] that is continually updated as the user moves through the exhibition and selects sound objects.

Fig. 3 shows an interaction schema of the user model with other modules. There are two main update sources in the system. First, as the user moves through the exhibition, the speed of the movement due to stopping or slowing down at different artifacts provide updates to the user model. The user's behavior type is computed based on the speed and homogeneity of the user's movement. Stopping and slowing down in front of an artifact are interpreted as interest in topics represented by the artifact. The user interests and intentions influence the presentation of soundmarks. For example, soundmark radius and volume is increased for those artifacts that correspond with current user interests. Another example can be the reduction of the number of soundmarks in the exhibition, if the user's recognized intent is to quickly cross the room.

The second source of updates to the user model considers the user's direct interaction when selecting a sound object. In the model, this maps to an increased user interest in topics presented by the sound object and updates the user's interaction history. We describe the user model and retrieval mechanism in detail below.
4.1. User model components

**Interaction history** is a record of how the user interacts with the ec(h)o-augmented museum environment. Two types of events are stored in the interaction history: the user's movement and the user's selection of objects. The user's path through the museum is stored as discrete time-space points of locations on the path. A second type of information stored in Interaction History is the user's selections in the form of URLs of sound objects.

**User behavior** in the museum context is well studied in museum studies [9] and is used in several systems personalizing the user experience [18,19]. In the case of ec(h)o, several categorizations were used; for example, one user may go through almost every artifact that is on his/her way, and another user may be more selective and choose artifacts that have certain concepts. Our categorization of user types is based on Sparacino's work [19] and it classifies users in three main categories: (1) the avaricious type who approaches artifacts in a deliberate and sequenced manner, (2) the selective type who explores certain concepts thoroughly, and (3) the busy type who wants a general idea of the exhibitions by browsing quickly through the museum.

In ec(h)o, the user behavior is not static. It continually updates by considering the location data accumulated in the previous 5 min; in addition to considering topics of previously selected sound objects.

**User interests** are represented as a set of weighted concepts from the ontology. In ec(h)o each artifact/exhibition is annotated with a set of concepts. The sound objects address a set of particular concepts as well. The system updates the user interests in response to two update channels described above. The interaction of the user and artifacts and sound objects is stored in the Interaction History that together with the user behavior type are used to infer the visitor's interests.

The following rule concept-evol-choose--1 shows an example of how concepts of interest are updated in the user model. The ?*user-model-concepts* object accumulates contributions from all activated rules first and indicates that the user model has to be updated. After all contributions are made, the rule update-user-model---1 (with lower salience value) fires and recalculates the user interests values. It then inserts facts representing values of user interests into the knowledge base. These facts are used in the ranking of sound objects (described in Section 5).
When designing a user model for ec(h)o we considered other application domains where the user model is needed. Another active research area of our lab is eLearning, specifically intelligent support to learners and automatic just-in-time assembly of learning material. A core part of the user model is maintaining user interests that also reappear in other contexts either directly as user interests or as user knowledge, abilities, skills, etc. Recognizing many similarities between requirements from ec(h)o and eLearning domains, we have designed our user model in a modular fashion that benefits from two easily scalable technologies: ontologies and rule-based systems. Fig. 4 shows the generalized flow of processing that keeps track of user interests with generic parts in bold.

The user observations and actions are related to the application-specific objects and the environment that can be modeled using ontologies. In ec(h)o, we use the CIDOC CRM ontology for modeling museum artifacts and the ontologies we developed for sound objects and exhibition (space). In other domains the objects and environment can be modeled in similar ways; for example, in the eLearning domain we model learning objects, courses, curriculum, and learning design (pedagogical processes). We found that user actions correspond to user’s interaction with learning systems.

In the Concept Mapping and Extraction block in Fig. 4, we use inference rules to extract the concepts relevant to user interests and level of user engagement.
with these concepts. For example, when a user selects a sound object annotated with primary and secondary concepts of interest, the system extracts these two concepts and assigns them two different levels of engagement ('activated concepts of interest' link in Fig. 4).

As the name suggests, the Interest Adjustment block is responsible for adjusting the user interest as a reaction to user actions. In our design, this is a generic component that has two parameters: maximum level of individual interest, and a maximum for a sum of all interests. Based on a set of activated concepts and previous values for interest, the algorithm re-computes the values accordingly. Both components are implemented as rule sets and therefore the model can be easily adapted to other applications.

5. Inference-based sound object retrieval

We have identified the following requirements for the retrieval of appropriate sound objects:

1. Content-relevant to the viewed artifact;
2. Content-relevant to the user interests;
3. Content invites to exploration of other areas;
4. Content is plausible from the psychoacoustics perspective.

In addition to the criteria for an individual object the following criteria apply to the sequence of the objects offered to the user:

5. Provide for exploration of a subject in depth;
6. Provide for the fluidity in experience both in content and sound experience;
7. Provide a mix of informational and entertaining objects.

The retrieval process in ec(l)o can be broken into several steps. The input into the process is user interests, interaction history and semantic descriptions of sound objects. In the process the criteria listed above contribute to overall ranking for each sound object.

The following rule \( \text{cl---1} \) contributes to the rating of object \( ?\text{in2} \). The object \( ?\text{in2} \) is a candidate object to replace previously listened to object \( ?\text{in1} \) (represented by the replace fact). The object \( ?\text{in2} \) is a candidate because it matches the concept of user interest \( ?\text{c} \) (fact user-concept) within the context of the current exhibition \( ?\text{e} \) (fact is-about). The object rating is a combination of level of user interest in the concept and level by which the concept is represented by the sound object. The rating is added to the \( ?\text{*object-ratings*} \) java object (see discussion in Section 6.5).

\[
\text{(defrule cl---1)}
\text{(user-group (user ?u) (group 1))}
\text{(replace (user ?u) (context ?e) (object ?in1: (neq ?in1 nil))}
\text{(sequence ?seq) (time-chosen ?t))}
\text{(user-concept (user ?u) (concept ?c) (level ?i) (time ?t))}
\text{(is-about ?in2 ?e))}
\text{(not (replaced (user ?u) (next-object ?in2))}
\text{(test (neq ?in1 ?in2))}
\text{=>}
\text{(bind ?r (* ?i ?n))}
\]

The object-concept facts were created from the semantic representation using rules below. These facts also include different levels for primary and secondary concepts (rules concept-level-cl and concept-level-c2):

\[
\text{(defrule concept-level-cl)}
\text{(PropertyValue & psch:hasPrimaryConcept ?in ?c)}
\text{=>}
\text{(assert (object-concept (object ?in) (concept ?c) (level 1)))}
\]

\[
\text{(defrule concept-level-c2)}
\text{(PropertyValue & psch:hasSecondaryConcept ?in ?c)}
\text{=>}
\text{(assert (object-concept (object ?in) (concept ?c) (level 0.5)))}
\]

The \( ?\text{*object-ratings*} \) is bound to a Java object that simplifies the calculation of object ratings:
The composition criteria considers the next object in the context of the objects the user listened to previously. The selection is based on theme, topic, concepts, and described artifacts. An example of such rules is a rule that increases the rating of the sound objects that context.

```
(defrule guidel--1
 (user-group (user ?u) (group 1))
 (replace (user ?u) (context ?e) (object ?inl & (neq ?inl nil))
 (sequence ?seq) (time-chosen ?t))
 (PropertyValue &s:hasInfoCategory ?in1 &s:Guide)
 (PropertyValue &s:describes ?in1 ?a))
 (is-about ?in2 ?e)
 (test (neq ?in1 ?in2))
 (not (replaced (user ?u) (next-object ?in2)))
 (not (breaks-logical-ordering (user ?u) (object ?in2)))
 =>
```

When all the rules contributing to the ratings of sound objects are applied the object with highest rating is selected to replace the object user listened to (rule calculate-best-object---1). The sound object that describes the same artifact as the object being replaced. The rule checks whether candidate object ?in2 describes the same artifact ?a as previous object ?in1 while ?in2 cannot be an exhibition object but an actual artifact within the exhibition.

```
(defrule artifact2artifact--1
 (user-group (user ?u) (group 1))
 (replace (user ?u) (context ?e) (object ?in1) (context ?e)
 (sequence ?seq) (time-chosen ?t))
 (context ?a ?e)
 (PropertyValue &s:describes ?in1 ?a)
 (PropertyValue &s:describes ?in2 ?a)
 (not (PropertyValue &rdfs:type ?a &crm:exhibition))
 (not (replaced (user ?u) (next-object ?in2)))
 =>
```

Another rule supporting e(h)o’s interaction model is the rule guidel--1 that favors objects annotated as a guide sound object after a previous guide object was offered for a particular artifact. This allows system to keep focus on the artifact. As guide objects are related to specific artifacts the rule makes sure that logical ordering between two consecutive sound objects is not violated.

```
(defrule calculate-best-object--1
 (declare (salience -20))
 (user-group (user ?u) (group 1))
 (replace (user ?u) (object ?in1) (order ?n) (time-offered ?t1)
 (context ?e) (time-chosen ?t2) (sequence ?seq) (category ?c))
 =>
 (bind ?best
 (sym-cat (call *object-ratings* getBest userl ?in ?t2)))
```

The salience value in the rule calculate-best-object--1 guarantees that the rule is applied after all rules contributing to the ratings of sound objects replacing this particular sound object ?in1.

However, in certain cases it is not desirable to offer some guide objects once the user listened to other guide objects. This is prevented by explicitly specifying such undesirable ordering. Second type of objects are expert objects that provide more generic information applicable across several exhibitions, e.g. sound objects describing relation between evolution and diversity.
For more details of information retrieval aspects incho see [11].

6. Implementation

The ec(h)o system was fully implemented, deployed, and tested in the setting of the real exhibition space in Nature Museum in Ottawa in March 2004. The system used radio frequency based position tracking system with an update rate of up to 1.6 seconds. The vision and audio delivery systems were developed in our lab in the MAX/MSP environment.

The reasoning module is fully implemented with all features described in the previous section. During the development we embedded the reasoning engine in the Tomcat environment in order to facilitate online editing of knowledge models as shown in Fig. 5. However, for the final deployment we removed the reasoning engine from the Tomcat environment for the performance reasons. All communication with the reasoning engine was accomplished through a UDP connection.

6.1. Reasoning engine implementation

The real-time nature of the ec(h)o environment was the driving force for the selection of the implementation platform that would support the reasoning engine. As shown in Fig. 5, the Jess inference engine is in the center of the reasoning module. We have used DAMLJessKB to load DAML+OIL ontologies into Jess (for details see [13]). DAMLJessKB uses Jena toolkit to convert ontologies into RDF triples which are converted to Jess facts (see examples in Section 3). When converted, ontologies are loaded into the Jess; the rules representing DAML+OIL semantics (provided by DAMLJessKB) infer all the missing relations in the RDF graph. This happens at the start time and prepares the system to respond to the input in a real-time fashion. However, this nice theoretical assumption was challenged by the reality of our implementation, which we summarize in the following sections.

6.2. Memory requirements of ontological representations

Ec(h)o makes use of several ontologies that need to be loaded into the Jess knowledge base. Table 3 summarizes the number of classes, properties, and instances for each ontology used in ec(h)o.

During the loading process the ontologies are converted into RDF triples and the full DAML+OIL semantics is applied, generating complete RDF tree for the knowledge models. Table 4 shows the number of triples for ontology models only and then for ontology models and instances before and after applying semantic rules.
Table 3
Ontologies used in cc(h)o

<table>
<thead>
<tr>
<th>Ontology</th>
<th>No. of classes</th>
<th>No. of properties</th>
<th>No. of instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concepts of interests</td>
<td>2</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>CRM</td>
<td>62</td>
<td>139</td>
<td>209</td>
</tr>
<tr>
<td>Exhibition</td>
<td>1</td>
<td>3</td>
<td>149</td>
</tr>
<tr>
<td>Psychoacoustics</td>
<td>52</td>
<td>26</td>
<td>2412*</td>
</tr>
<tr>
<td>Theme</td>
<td>1</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Topic</td>
<td>1</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Topic, dewey</td>
<td>107</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* There are 613 instances representing sound objects. The remaining number represents prefaces—short sound objects introducing the main object.

As we can see in the first row of Table 4, the number of facts increased by 75% after applying DAML + OIL semantics. The same wasn’t true for the facts representing instances. We explain this by instances linking to concepts and other instances through properties. As we do not have a rich system of properties in our ontologies the number of inferred facts is smaller.

Table 4
Number of facts representing ontologies in Jess at the startup

<table>
<thead>
<tr>
<th>No. of facts</th>
<th>Before applying semantics</th>
<th>After applying semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology models only</td>
<td>8321</td>
<td>14411</td>
</tr>
<tr>
<td>Ontologies including instances</td>
<td>40910</td>
<td>65505</td>
</tr>
</tbody>
</table>

6.3. Rules

Although the numbers listed in Table 4 are relatively moderate, the real influence of the number of facts is felt in combination with forward chaining rules in Jess. Jess implements the RETE algorithm to build a network to keep track of possible combinations of facts supporting rule activations. With a large number of facts with similar patterns representing RDF triples, the number of possible combinations can be potentially huge.

Another aspect of cc(h)o that was influential for the rule set design is the sequential nature of the retrieval process. The processing chain from the rule perspective is shown in Fig. 6. The processing is triggered by an observation that is inserted into the knowledge base as
a fact. First, the system updates user and environment models, then proceeds with the ranking of objects considering updated user and environment models; finally, it applies the interaction criteria to select next recommended objects. To achieve the sequencing we had to prioritize between groups of rules using salience values which consequently had some undesirable effects. We describe particular challenges and lessons learned in the section below.

6.4. Performance

The final implementation of the reasoning engine ran on a Pentium M 1.5 GHz with 768 MB of RAM. The final demonstration served two concurrent users (of maximum four possible). The reasoning engine received input about the location of each user approximately every two seconds. This input caused a short 50\% spike in processor activity when the user moved within the same exhibit and a short 100\% spike when the user changed exhibits. After receiving input about user selection of a sound object, the processor performance briefly reached 100\% and completed the selection of a new sound object below the 1 s limit (this was well below the time the user actually listened to sound objects, which was typically 5–20 s). The memory usage during load time reached above 512 MB and then stabilized around 372 MB (these numbers measure memory used by Java JVM).

The use of a forward chaining inference engine has proved itself to be an efficient mechanism for responding to the dynamic nature of the user input. The system loading time was relatively long as a lot of parsing and initial inference is performed on the ontologies and object descriptions. After the startup phase the amount of inference is limited to updates from the user input, resulting in quick responses.

6.5. Challenges and lessons learned

From the implementation perspective (we will talk about qualitative evaluation in the next section) the reasoning engine had the only criterion: a real time response to other parts of the echo system. As we developed content incrementally we did most of the reasoning engine design and development with a limited set of 150 sound objects recorded early in the process. As a result some of the challenges showed up when we scaled up to the full set of over 600 sound objects. Another aspect that challenged us was simultaneous support for multiple users. We discuss some of these challenges that have general implications for similar systems.

6.5.1. Problem 1: rich semantics can cause significant computational delays

6.5.1.1. Problem. The rules for selecting sound objects use several criteria for fluency of a dialogue. The criteria depend on ontological annotation of themes, topics, concepts, etc. With richly annotated objects the system was not able to select new sound objects in real-time.

6.5.1.2. Cause. Different criteria are represented by individual rules and when fired they contribute a value towards the final score for the objects. Some criteria are satisfied for many sound objects. For example, the criterion that keeps coherency of theme in the dialogue is activated many times as all sound objects are categorized only into seven themes, which are present in the exhibition. The criterion itself has little decisive power but consumes many resources.

6.5.1.3. Solution. After we analyzed results from the preliminary user testing we eliminated some of the rules/criteria. This had a minimal impact on the quality of the end user experience and significantly reduced the number of rule activations. In general, the semantic annotation that categorizes an object in a coarse manner should not be used in a generative computation but rather used for filtering out of unsuitable candidates.

6.5.2. Problem 2: concurrency has to be treated explicitly

6.5.2.1. Problem. In the case of concurrent users, the reasoning modules waits until sound objects for all users are computed and delivers all of them at the same time. This caused significant latency for individual users.

6.5.2.2. Cause. In echo(h) we had to work with salience values (rules with higher salience value fire before rules with a lower value). In the case of multiple users, the rules interfered with each other. For example, if a second user makes a choice before a computation for the first user is finished then rules with a higher salience
for the second user start firing. This causes computation for the first user to be pending until all the rules for
the second user with higher salience values fire. With
an increasing number of visitors the latency increased.

6.5.2.3. Solution. We found the solution with help
from the Jess community. We categorized the users into
groups and assigned an identical set of critical rules
for each group. The set of rules is activated only for
users belonging to the group, so the users from di­
ferent groups do not block each other. In general, the
same problem can occur when the reasoning engine is
exposed as a web service and a multiple access to the
service is allowed.

6.5.3. Problem 3: know-your-tool or carefully
consider implications of implementation platform
6.5.3.1. Problem. A rule that gets activated many
times with a ‘not’ clause positioned early on the list
of preconditions takes a long time to fire.

6.5.3.2. Cause. ‘not’ Pattern can only match an ab­
sence of a fact. In our case, it is evaluated only when
the fact is asserted (then it fails) or when the pattern im­
mediately before the ‘not’ clause on the rule left hand
side is evaluated. Therefore patterns following the ‘not’
clause are evaluated at the runtime. Combining this
with a large number of candidate facts resulting from
the ontology representations causes significant delays.

6.5.3.3. Solution. Position a ‘not’ clause as the last
pattern on the left hand side of the rule.

6.5.4. Problem 4: do not use rules for extensive
numerical computations
6.5.4.1. Problem. Computing multi-criteria numerical
preferences required assertion of extensive number
of facts and use of salience values resulting in growing
response times for subsequent iterations.

6.5.4.2. Cause. As several criteria are used to con­
tribute preference values to the overall score of each
sound objects, we need a mechanism ensuring that all
contributions are made before making sound object se­
lection decision. There are two possible approaches:
first, add all the contributions as facts and then fire
summation rule; or, keep adding contribution to one
fact, which means retracting and re-asserting it into the
knowledge base. The second approach is more time
consuming. Both approaches require use of salience
values to make sure all contributions were made.

6.5.4.3. Solution. Build a simple extension in Java (or
other language) that will perform the computation and
make it accessible through the inference engine exten­
sion mechanism (direct call to Java in the case of Jess).
This will speed up computation as generating large vol­
ume of facts and build up of the Rete network for rule
activations will be avoided. The salience will still be

---

**Fig. 7. Number of facts in iteration steps.**
needed to ensure that all contributions were made. The rules in Section 4.1 illustrate the solution. Figs. 7 and 8 show the effect of moving computation from the knowledge base to the external Java module.

7. Evaluation and discussion

ee(h)o is a complex interdisciplinary research project that has to be evaluated from different perspectives. As the evaluation of ubiquitous computing systems is extremely complex [20] we have found Miller's and Funk's [14] view of the problem of evaluation of ubiquitous computing systems from the traditional 'validation' and 'verification' perspective very useful. In validation we evaluate whether the system performs the functions it was built for based on the requirements specification. Verification tests the system against the reality by checking whether the system provides the envisioned benefits. Finally, the evaluation of technical aspects of the system implementation can provide insights to the developers of a similar system.

Following Miller's and Funk's approach allowed us to focus our evaluation on the areas where we researched novel approaches in the adaptive ubiquitous systems. We also avoided the evaluation of the aspects of the system that are not well defined or understood and the evaluation results would provide very little value.

Our validation efforts concentrated on the system components for which we either had predicted outcomes or have established the criteria for such outcomes. Specifically, we have validated the flexibility and responsiveness of the user model and effectiveness of the object recommendation component. We have verified our solution with the targeted end user group through extensive questionnaires and videotaped interviews.

In this section we provide an overview of the evaluation results as those are reported in detail elsewhere [10]. A detail account is given for the evaluation aspects related to rules and ontologies.

7.1. Suitability of ontologies and rules for user modeling

In the context of our work, the user model performs a function of a recommender system [15]. "Recommender systems represent user preferences for the purpose of suggesting items to purchase or examine" [6]. Several types of recommendation techniques have been developed: collaborative, content-based, demographic, utility based, and knowledge-based. Often the researchers combine several techniques to achieve maximum effect. The knowledge-based recommender systems perform favorably with respect to the introduction of new users and new items (so called 'ramp-up' problem [12]) which is an important feature for ubiquitous computing environments. The knowledge recommender systems require three types of knowledge [6]: catalog knowledge or knowledge about objects to
be recommended, functional knowledge of mapping between user needs and objects, and user knowledge.

From this perspective we have used ontologies extensively to describe knowledge about objects, environment, and the user. As multiple criteria were used to determine the user interests, a rule-based approach provided us with the flexibility that enabled us to evolve the system through several iterations. Furthermore, to be able to respond to the specifics of the application we have parameterized the influence of inputs from the user and ubiquitous environment such as maximum interest value, object selection, and location change contributions towards user’s interests, etc. The purpose of the parameterization was to fine-tune our generic user model framework. We performed an extensive testing for the suitable combination of parameters in the lab setting with early input from the test users.

The user model uses a spring model to keep interests balanced. The level of interest is represented by the real number and can range from 0 to 10 (the value was set with respect to other values used for ranking objects). The sum of all interests never exceeds the value of 30. In the model we consider only positive influence from the user interaction that directly increases the level of some of the interests. When this increase causes an imbalance (the sum is above 30), the implemented spring model proportionally decreases values of other interests.

Fig. 9 shows the sequence of steps and evolution of interests in each step. In the first step three concepts are selected by the user. The circle icon indicates concepts introduced to the model by the visually represented exhibit concepts (Steps 2, 11, and 15). In the rest of the steps the user selected sound objects. The square icon indicates primary concept and triangle icon secondary concept in the selected sound object.

The rule-based model proved to be very flexible and responsive to the parameters. The representation of the knowledge in the form of ontologies made the

![Fig. 9. Evolution of user interests.](image)
design and implementation of the model very easy with
the clear way of accessing the knowledge. The use
of the DAMLJessKB module accompanied with the
DAML + OIL language semantics made the inference
in the knowledge base transparent, which enabled us
to concentrate on the model implementation instead on
navigating and inferring static knowledge.

7.2. End user verification

As Miller and Funk [14] point out the verification
evaluates the system from the perspective of provided value. Typically, the qualitative methods are used and end user testing is involved. The qualitative methods are more suitable for novel approaches and new areas of research to verify the potential of those.

In ec(h)o we have conducted in depth usability testing of the system while deployed in the real museum setting. An extensive testing was done with 6 subjects. The subjects were briefly trained on how to use the system (learning phase), and then had an opportunity to ask questions. They used the system on their own for a period of 10–20 min. After this session, they completed a modified version of Ben Schneiderman’s acceptance test. Finally, we conducted and videotaped interviews with the subjects. In addition to those tests, we had one museum expert evaluating the content side of the system in depth.

The overall use of the system was rated relatively high. For example, when asked to rank between 1 and 5 on a Likert scale (5 being best) over five different questions relating to the overall reaction to the system, the averaged response was 3.6. The evaluation scored 4.6 for ease of use and 2.8 for satisfaction. Navigation and engagement of the audio information rated high; for example, appropriateness of the audio experience scored 4.0. This leads us to believe that the system meets or satisfies many of the current advances of electronic guide systems. Participants were explicitly asked to compare the system to experiences with other systems and the prototype ranked favorably.

Difficulties exist in relating sound objects to a specific artifact. In certain cases visitors didn’t mind the ambiguity while others clearly found it frustrating. The results also differ in the ‘attitude’ related questions. Some users had strong feelings about their preferred modes of interaction; others approached the system from the more playful perspective.

It is difficult to draw conclusions from the number of testers we had. The expert reviews were strongly in favor of the approach and the system. The reviews were helpful in catching potential inconsistencies and challenges.

Hatala and Wakkary [10] provides more detailed discussion on the ec(h)o evaluation results from the user modeling perspective.

7.3. Efficiency of ontologies and rules for ubiquitous real-time applications

As the implementation section already presented concrete results and lessons learned from using ontologies and forward chaining rules in ec(h)o, in this section we summarize the outcomes and highlight a potential of used technologies for the realtime applications.

The ec(h)o implementation was based on technologies that were available, stable, and supported by tools in 2003. W3C’s Ontology Web Language has since superseded the DAML + OIL ontology language. This would be our candidate language if we were developing the systems now.

The representation of DAML + OIL (or OWL) ontologies in the forward chaining system knowledge base reflects their RDF representation in the form of triples. This form of representation creates an enormous number of syntactically similar facts resulting in potential performance problems. However, these problems can be overcome by using unordered facts [13]. A major benefit for the real-time systems is that the inference applies ontology language semantics at startup time, inferring the full graph representing all existing relations. During the runtime only relations with newly created instances are inferred resulting in speedy updates to the system. From the developer’s perspective the uniformity of the representation and availability of the full relation graph makes it easier to develop rules referring to the ontologies and properties between objects.

There are many best practices available for writing forward-chaining rule systems. With the large number of syntactically uniform facts some of the recommendations need to be observed rigorously otherwise resulting in a big performance hit. A good knowledge of underpinnings of the inference system is needed (in our case a Rete network and algorithm) particularly about ordering facts in the precondition part of the rules and
using the not clauses in the rules. Also, carefully considering the delegation of certain tasks such as numerical computation to the external modules can improve the performance significantly.

One specific aspect of the multi-user real-time application that we were not able to resolve satisfactorily is the possible collision of rules for individual users. The problem occurs when the salience values are used to sequence processing steps. Our approach grouped users and assigned them their own rule sets so the users from different groups did not collide. A more robust solution would call for the dynamic creation of modules for each user with the full management of these modules to avoid exhausting of the system resources.

Another related effort in the Semantic Web community in the area of rules is Rule Markup Language (RuleML) aiming at interoperability between inference environments. However, we have not considered the RuleML since other requirements such as performance had a priority over the interoperability. We also wanted to benefit from the ability to experiment with and extend our selected inference engine.

8. Conclusions

In this paper we have presented the design and implementation of an augmented audio reality system for museum visitors named ec(h)o. Each visitor experience is tailored to the visitor's interests. The user interests are inferred from the user's movement through the exhibition as well as from the visitor's interaction with the sound objects. The sound objects are retrieved based on their relevance to the user interests, narrative criteria, and psychoacoustic criteria. ec(h)o uses ontologies to describe concepts, temporal and spatial characteristics, and psychoacoustic and sound characteristics of sound objects. In the core of the system is a rule-based inference engine that powers the retrieval mechanism and the user model specifically designed for the applications using rich semantic descriptions.

The system is a result of convergent research streams from research in object repositories, interaction design, auditory display, knowledge representation, and information retrieval. The ontologies combined with the rule-based inference proved to be a powerful implementation platform well suited for this type of the systems. We believe this has enabled us to extend works cited through the paper in several directions. First, it extends the work of the Alfaro et al. work [1] by building a rich model of the concepts represented by the sound objects. In ec(h)o, the content presented to the user is not pre-processed for possible linkages as in the systems using Rhetorical Structure Theory [24]. Our approach replaces pre-processed linkages with a retrieval mechanism based on composition and interaction criteria formulated in the form of the rules and applied to semantically-annotated independent objects.

The requirements of the real-time ubiquitous application required us to face the challenges stemming from the combination of two powerful technologies: ontologies and forward-chaining rules. We have summarized our findings in the lessons learned that provide a transferable generic knowledge for similar type of application. The ec(h)o proved that ontologies and rules provide an excellent platform for building a highly-responsive context-aware interactive application.

Acknowledgements

Work presented in this paper is supported by a Canaric Inc. grant under the E-Content program. The authors would especially like to thank Mark Graham and his colleagues in the Nature Museum in Ottawa for their enthusiastic support of this project. We would also like to thank our colleagues and participants in several workshops who contributed to the development of the project, namely Kenneth Newby, Dale Everden, Doreen Leo, Gilly Mah, Robb Lowell, Mark Brady, Jordan Willms, and Phil Thomson.

References


Dr Marel Hatala is an Assistant Professor at School Interactive Arts & Technology and a director of Laboratory for Ontological Research at Simon Fraser University Surrey in British Columbia, Canada. Before this appointment Dr Hatala was a Research Fellow at the Knowledge Media Institute At The Open University, UK and an Assistant Professor at The Technical University of Kosice, Slovakia. Dr Hatala holds his PhD in Cybernetics and Artificial Intelligence.

His main research interest areas are willow representation, ontologies and semantic web, user modeling, intelligent information retrieval, organizational learning and eLearning. In the last four years Dr Hatala focused on building a scalable distributed infrastructure for e-learning systems and middleware enabling interoperability between heterogeneous repositories. The infrastructure connects repositories in Canada, US, UK, Australia and Europe. His current research projects look at how semantic technologies can be applied to achieve interoperability in highly distributed and heterogeneous environments, what the social and technical aspects of building a distributed trust infrastructures, and what role the user and user group modeling can play in interactive and ubiquitous environments.

Ron Wakkary is Associate Professor in the School of Interactive Arts & Technology at Simon Fraser University in British Columbia. His research includes interaction design, and design methods and practice for interactive systems. He has served on the Canadian Culture Online Advisory Board to advise the Minister of Canadian Heritage, and is on the Banff New Media Institute Research Advisory Board. He has lead design and technology research projects for the Canadian Nature Museum, Nokia Research, Museum of Modern Art, the Guggenheim Museum, and Electronic Arts Internex. He is currently leading the Am-I-able Network for Responsive and Mobile Environments, a research network funded by Canadian Heritage, and research on the everyday complexity of design funded by the Social Sciences and Humanities Research Council.

Leila Kalantari is pursuing a master's degree at School of Computing Science, Simon Fraser University, British Columbia. During her undergraduate studies at SFU she worked as Research Assistant for Computational Logic Lab, Intelligent Systems Lab and Laboratory for Ontological Research.
Framing complexity, design and experience: a reflective analysis

Ron Wakkary
Simon Fraser University, Canada
rwakkary@sfu.ca

Abstract

The paper discusses theory and practice in the roles of reflective practice and contextual design in addressing issues of complexity in design. The author defines a new understanding of the role of complexity in design. The paper reviews theories in design and HCI related to reflective practice, context, and embodied interaction. A case story of practice in interaction design and museums is presented as a practice-based investigation of the complex. The paper calls for the framing of larger research agendas in this area with the need to further work on issues of context, reflective practice, embodiment and human activity in order to provide a more comprehensive and integral view of design activity. The paper concludes with the need to reframe concerns in design in order to emphasise situated participation, non-rational design strategies, in situ design and a re-orientation in focus from tasks to experience.

Keywords: complexity, design methods, design theory, interaction design, reflective practice

1 Introduction

This paper argues for a shift in design and design practice through the framing of complexity. The need for establishing this connection arises out of the author’s practice in interaction design, where design is approached socially, contextually, and experientially. Interaction design is understood to be an inter-disciplinary convergence of design and HCI (human-computer-interaction), inclusive of aspects of interactive art, performance, computing science, cognitive science, psychology and sociology (Sanders and Dandavate 1999, Winograd 1997, Löwgren 2002, Precece et al. 2002). This research is part of ongoing investigation in complexity and design. The paper will provide an overview of theoretical starting points from design and HCI for understanding design through complexity. As part of a practice-based investigation, a case story will be discussed describing a design project in ambient intelligence and museums. The paper calls for the framing of larger research agendas in this area with the need to further work on issues of context, reflective practice, embodiment, and human activity in order to provide a more comprehensive and integral view of design activity. The paper concludes with the need to reframe concerns in design in order to emphasise situated participation, non-rational design strategies, in situ design and a re-orientation in focus from tasks to experience.
2 Design situations, ordinariness and complexity

In the contexts of design and HCI, complexity has been discussed in isolated pockets along three dimensions, all of which assume a design process but rarely acknowledge it. Firstly, design outcomes are understood to be complex, as expressed in architecture, evolutionary theory, and human factors engineering (Dawkins 1986, Norman 1998, Venturi 1966). Secondly, attention has been given to the definition of design problems as complex, from Rittel's notion of the 'wicked problem' to Simon's 'ill-stated problems' to Alexander's 'pattern language' (Rittel and Webber 1973, Buchanan 1995, Simon and Súlkössy 1972, Alexander et al. 1977). Thirdly, there has been discussion on the role of HCI and Information Design in supporting end users' complex problem solving (Gay and Hembrooke 2004, Albers and Mazur 2003, Mirel 2004). However, another trajectory is emerging, in which the term 'complexity' is not explicitly used. This includes the ideas of reflective practice and context. Here, design is seen to be boundless and dynamic rather than bounded and quantifiable (Buchanan 1995, Barnard et al. 2000, Nardi and O'Day 1999, Thackara 2001, Dourish et al. 2004, Fischer 2000, Schön 1983). The focus of my research on complexity emphasises this latter trajectory due to its implicit understanding of complexity as a common factor in design activity.

Two key issues emerge from the current state of discourse on complexity and design. The fields of design and HCI are moving closer together and at times discussed interchangeably and at other times understood to be intertwined (Ehn 1989, Norman 1998, Fallman 2003, Coyne 1995, Gay and Hembrooke 2004, Fischer 2004, Zimmerman et al. 2004). Winograd was among the first to identify this trend (1996, 1997). In large part, the motivating factor is the need to acknowledge the unique contextual aspects of interaction and the need to design in response to specific (typically complex) and not generic situations, a shift he coined as the move from machinery to habitat (Winograd 1997). The second key issue is the lack of a coherent theory on complexity in design, especially inclusive of design practice.

While many argue that design produces complex artifacts, and that design practice can be captured as complex formalisms, I argue that we need to understand design as an activity that responds to situations of varying complexity. The key distinction is a question of understanding design as a prospective action, that is actively reflecting within a present moment on future action and contingency, as opposed to a retrospective event from which we view the design process or artifact as a stable past action with little attention given to context. In the former, the relationship between activity and situation in design is integral and dynamic. For example, Schön views design as a conversation (Schön 1983). Rittel understands design as argumentation (Rittel and Webber 1973). In either case, each metaphor implies a dynamic act reliant on interpretation and multiple perspectives. The metaphors explicitly describe an activity in which the actions of speaking/listening, and the nature of what is being said/understood are intertwined and dynamically inform each other. In addition, like a conversation, design, and in turn complexity, is quite ordinary and ubiquitous. And so, an alternate way to consider design is that it is an activity that is integrally related to complex yet everyday situations.

For example, a visit to a museum reveals an everyday yet complex interaction situation. The factors within museum experiences are social, cultural, historical and psychological. The influences on the experience vary from the actions and previous knowledge of the visitor, visitor's learning style, the dynamics of others around them including friends, family and strangers. Naturally, the
experience is affected by the presence of the artifacts and the relationships within collections as an outcome of institutional history, curatorship, exhibition design and architecture. The time of day, duration of visit, room temperature and so on—all have an impact. The experience can be characterised as multivariate, that is, it cannot be assessed by a single factor such as exhibit design, signage, or time spent in front of an artifact (Lehn et al. 2001). Instead, the museum experience is subject to multiple influences and results in multiple outcomes (Leinhardt and Crowley 1998). Identifying a design intervention that may have a direct and positive impact on experience is clearly not easy! Many similar situations have been discussed in design research such as how we work (Ehn 1989), seek information (Nardi and O’Day 1999), learn (Gay and Hembrooke 2004) and live in our homes (Tolmie et al. 2002, Bell and Kaye 2002). We can see that almost any situation is in some form complex, yet few discussions include how we might design for these situations.

In part this lack of exploration of design practice as a response to complex situations is a result of the unwarranted focus given to viewing complexity in design as a quality of artifact or process. Such views tend to abstract or isolate either the design object or process from their context. To a large degree, this complexity in outcome is seen to be the result of a complex design process. Are design processes really complex or do we just assume that a complex outcome is the result of a complex process? And therefore should we assume that a simple outcome is the result of a simple process? Clearly we should not. I think we all understand that while many design outcomes are complex artifacts and actions, many outcomes are extraordinarily simple. And the reverse is true, simple processes can prove to be very effective. We have a tendency to analyse design retrospectively as opposed to prospectively—and in the process over-interpret for rational attributes such as logic and uniformity. If we did view design prospectively, as in fact design practice demands, we would see that complexity precedes, accompanies and follows design action. Complexity is contextual, situated and dynamic and therefore cannot be isolated in processes or artifacts. That is design and designers’ actions respond to complex situations. What we find is that the process is not pre-determined as complex, symmetrical or simple in structure, rather it is a dynamic process that is improvisational and responsive to the changing design situation. An active stance is required in design. Such design strategies have come to be understood as reflective, embodied, or contextual in practice.

2.1 Reflective practice: dynamism, contingency and unfolding in design

Schön’s account of professional practice spanned design, engineering and psychoanalysis. His observations of professional activity lead him to reject a theory of technical rationality and to develop an alternate theory of the professional as reflective practitioner. A reflective practitioner embodies the full scope of the complexity of the design situation, a term given us by Schön that is highly appropriate for understanding complexity and design. The practitioner’s knowing is embodied in action and her practice is understood through reflection. A reflective practitioner focuses on ‘problem-setting’ activities to overcome the limits of ‘problem solving’ (Schön 1983 37–39). The problem and design response is reasoned through experimentation, and fluidly engages in a variety of representations from sketching to scenarios as a mode of experimentation—what Schön refers to as ‘frame experiments’ (Schön 1983 150). According to Schön, the reflective practitioner as designer interactively frames the problem and names the things she attends to within this frame, she generates ‘moves’ toward a solution and reflects on the outcomes of these moves. The cognitive
scientist Gedenryd has expanded on the dialectical nature between Schon's concepts of exploration (setting the problem) and experimentation (framing the problem) (Gedenryd 1998 l69-l72). The designer functions by going back and forth between construction and reflection as a means to understand the designer's situation she is creating, hence the notion of the designer as having a 'reflective conversation' with the situation. Schon argues against the fabled objectivity of a rational approach by asserting that the reflective practitioner, through this type of engagement with the design situation, is shaping the situation (Schon 1983 150). As such, reflective practice accounts for the dynamic, contingent, and unfolding nature of design. For Schon, reflection is a critical element of professional activity and design.

Other theorists have added to Schon's model of the reflective practitioner. Coyne and Gedenryd root the practice in the philosophical pragmatism of Dewey, Heidegger and Rorty (Coyne 1995, Gedenryd 1998). Under the pragmatic account, design takes the form of a hermeneutic process of interpretation and creation of meaning, where designers iteratively interpret the effects of their designs on the situation at hand (Gedenryd 1998, Coyne 1995). Louridas extends the theory through metaphor. He borrows Levi-Strauss's concept of a 'bricoleur' as a metaphor for a designer as a reflective practitioner (Louridas 1999). A bricoleur is someone who makes do with what is available or encountered in a specific situation. Like Schon's notion of reflective conversation (Schon 1983 165-66), the bricoleur operates from the available means (the concrete tools and materials offered by a specific design situation) but treats them as signs, by seeking to determine and redefine the roles they can play in the given situation. The designer as bricoleur interrogates the situation from within and from multiple perspectives while constructing with the heterogeneous means the situation affords.

2.2 Human activity and context: the 'theoretical pinch' in human-computer-interaction

Many HCI theorists and researchers have come to identify issues of 'context', 'situation' and 'practice' as putting a strain on the traditional theories of HCI (Nardi 1996, Dourish 2004, Bodker 1991, Gay and Hembrooke 2004). As Nardi puts it,

we are beginning to feel a theoretical pinch, however—a sense that cognitive science is too restrictive a paradigm for finding out what we would like to know.

(Nardi 1996 13)

The understood need is to move the theoretical trajectory of HCI from a reductivist understanding of human cognition toward an understanding of embodied and situated human activity.

In response to the rigidity of cognitive science, ethnographic and scenario-driven methods have begun to take hold in HCI practice (Suchman 1987, Carroll 2002, Carroll 2000). Further along in this direction, an emerging set of 'context-based' theories for HCI have adapted ideas from an even wider spectrum of psychological, social, political and philosophical theories based on understanding human activity. For example, Nardi, Bodker, Gay and others (Nardi 1996, Bodker 1991, Gay and Hembrooke 2004) have advocated on behalf of activity theory, a theory developed by psychologists in the early 1920s (Vygotsky [1925] 1982), as a research tool and an alternative framework for understanding human activity as it relates to individual consciousness. While the primary concern of activity theory is human activity; the insight is in the view that activities can only be understood through the role of related everyday artifacts, and that artifacts and activities are inextricably situated in a social practice. Dourish (2001, 2004) argues in his concept of embodied interaction that activity and context are dynamically linked—or 'mutually constituent' (Dourish 2004 14).
Based on the philosophical viewpoints of Heidegger and Wittgenstein, Dourish argues against the rational notions of abstracted cognition in favour of understanding human activity as an embodied practice that negotiates (and constructs) meanings in systems and contexts through interaction.

3 Reflective case story of practice and method

In practice, it is clear how design actions are responses to situations of varying complexity. Underpinning the responses is the simultaneous analysis and action that is prospective reflection within a present moment. The case story aims to demonstrate how a theoretical understanding of complexity in design is informed in large part through praxis.

The case story stems from a recent research project in ambient intelligence. The project is an audio augmented reality guide for museums, known as ec(h)o, integrated with a semantic web based and adaptive information retrieval system. The platform is designed to create a museum experience that consists of an interactive virtual layer of three-dimensional soundscapes that are physically mapped to the museum exhibition. The source for the audio is digital sound objects. The digital objects originate in a network of object repositories that connect audio content from one museum with other museums’ collections on the network. The system enables interaction by movement and gestures without the direct use of a computer device (Wakkary et al. 2003, Hatala et al. 2004). The visitors, wearing wireless headphones, experience a real-time soundscape composed of sounds related to artifacts nearby, as they move through the exhibition space. In closer proximity to artifacts on display, visitors hear and select audio information based on what is on display and choices inferred by a reasoning engine based on their pre-selected interests, past movements in the space and previous interaction. Selections are made by the visitor through gestures with a wooden cube.

In this project, the objectives for the system were to evaluate visitors’ experience with ec(h)o in comparison with their past museum experiences. In our case, we were not interested in an objective viewpoint or quantification of experience, rather our evaluation methods were aimed at first-person evaluations by the visitors, contextualised by quantifiable data based on the reasoning engine and location-tracking data. However, the user evaluation is not critical to the
discussion of design/HCl practice in this paper and therefore will not be discussed here.

3.1 Overview of methodology
In combination with emergent actions in practice and the preliminary framework drawn from the relevant theories of reflective practice and contextual design, the author explored the development of a design method that was responsive, improvisational and emergent in structure. The method anticipates dynamic responses rather than outlines separate and sequential steps. The method is best described as having few components that react dynamically to each other until a reasonable design outcome is agreed upon. A key concept behind the method is Schön's notion of 'frame experiments'. Here this idea is instantiated in two forms, scenarios and participatory workshops.

Scenarios
As Schön argues the design process is led by 'frame experiments'. In our case, one form is a scenario. The process begins by enacting a possible outcome based on observation and conceptualising of the design situation. Like traditional use of scenarios in design, the goal is to envision a possible outcome or future as a response to the design situation. The different forms of scenarios include role-playing, storyboarding, scripts/narratives, sketches, videos, and interactive works—however each form is enacted within the same physical and social context of the design situation in order to ensure that the scenario is designed 'in the world'. The design process begins with a scenario, yet subsequent scenarios are created whenever it is required. This occurs often since each subsequent scenario is revised and deconstructed through participatory design workshops. For example, in ec(h)o videos taped in various museums were produced as scenarios 'documenting' the visitor's experience and the system.

Participatory workshops
Workshops are another form of a 'frame experiment', however based on participatory design. Involving people in open but structured workshops allows for exploration of design responses to situations generated by scenarios. Workshops are a response to scenarios focusing on aspects of interaction that are enacted but not actively designed in a social or human context. Workshops can be in response to other workshops and are therefore only planned one at a time in an ad hoc and responsive fashion. Each workshop arises out of the previous design inquiry. Initially, in order to invite participation the workshops include prototyping as a participatory act alongside simple actions, typically adopting a low-resolution approach, such as paper prototypes. Over the course of the development of the design, the workshops and nature of the prototypes shift from generative to evaluative. As an example, in ec(h)o a participatory workshop explored movement, gesture and three-dimensional audio as a possible interaction model, by exploring ways of virtually 'catching butterflies'.

Prototypes and prototyped environments
Prototypes and technical workshops serve an enabling and evaluative function. Early in the process they act generatively, supporting design responses with technology or exploring them through 'Wizard of Oz' approaches. As the design outcomes emerge, components of the eventual system become prototyped and together are evaluated and help to evaluate the interaction through participatory workshops.

The method is purposely simple, involving the generative techniques of scenarios and participatory workshops, and enabling techniques of prototyping. Yet the resulting process is a complex non-linear structure that does not include separation of activities or inherent sequences. It does not privilege planning preceding action,
or decomposition and analysis as a prerequisite for synthesis. The approach is a response to traditional methods in design, methods that simply lack in responsiveness to design situations. The exact sequence is not predictable but the overall pattern is. For example, while you may not know what workshop will follow it is clear that it will connect with ‘parallel’ workshops and eventually a scenario. The process is akin to dead reckoning approach to navigation—setting a general direction, and discovering the world as you go and marking each point in absolute reference to a previous point as you encounter them.

3.2 Description of the design practice in ec(h)о

The aim of the design activities was to develop a gesture-based interaction model, a navigation model, an audio display model and to prototype the required supporting interactive system. The team developed five scenarios excluding a number of storyboards and informal role-playing sessions, six participatory workshops, and a series of iterated prototypes and technical workshops. Preceding the scenarios and workshops we engaged in two ethnographic observation sessions and staff interviews at the Canadian Nature Museum. The project produced a publicly demonstrated prototype system, user testing and technical documentation. The scenarios explored the visitor experience of ec(h)о from the visitor’s perspective. In addition to video, the scenarios included storyboarding and interactive mock-ups. Each scenario responded to the conceptualising and outcomes of the participatory workshops. The scenarios evolved into documentations of the actual system since each re-direction and refinement was reflected in each scenario. All the videos were video-taped in museum environments.

In the workshops, the majority of participants for the workshops were drawn from our university campus. The ages of the participants ranged from 23–63, with the average age range being between 31–36. The backgrounds also varied, we were careful to include in equal numbers, students, faculty, staff, administrators and genders. Participants were screened based on past experiences with visiting museums and experience with museum technologies. The workshops were evenly split between group and individual activities. The later workshops required more specific problem testing and utilised more refined prototypes or models and so tended to be individual, while the earlier workshops were low resolution and required group input. Each workshop was videotaped, and followed by a structured interview, and each participant completed a questionnaire. Sample size for each workshop ranged from 8 to 12 participants.

Each workshop was given a name and an open call was made to the university. The
Figure 3. A series of frame stills from the various scenarios.

design team structured, planned, made the call, and performed each workshop within two weeks or so of deciding another workshop was required. Below are brief descriptions of each workshop.

**Workshop 1: How do you catch butterflies?**
The objective of this workshop was to begin the development of an interaction model based on human gesture in response to spatial audio that was envisioned in the first scenario. The team initiated the co-designing by beginning with the metaphor ‘catching butterflies’. In discussion, brainstorming and ‘bodystorming’ sessions, participants helped the design team come up with two alternate metaphors for considering gesture. They were asked to describe how they would act out these metaphors and to experiment with them. In the group discussion conducted at the end of the workshop the participants talked about the metaphors, resulting gestures and their responses to the constraints of the model.

**Workshop 2: Sticks and stones**
In response to workshop 1, workshop 2 was an exploration of movement with objects in response to the audio display. Participants were split into teams and asked to develop objects that would facilitate hand movements and could ‘function’ with a ‘Wizard of Oz’ audio display system. Participants were given toys, objects, and various materials to modify and construct. Each team was given a few minutes to explain their approach such that the other team could ‘play/test’ the objects. Discussion of these playtypes raised relevant issues such as limits and potential expansion of the playtypes, as well as critiques and responses to the models.

**Workshop 3: House of cards**
The workshop was designed to generate a conceptual model for navigation based on the developing gesture interaction model. *Trivial Pursuit™* cards were modified to provide us with an extensive ‘repository of objects’. Utilising an extreme variation of a card-sorting exercise, three models were generated including one that we ultimately incorporated in the final prototype without modification.

**Workshop 4: Serious play**
In response to the navigational model and
Framing complexity, design and experience

initial interaction prototypes used in video scenarios, this workshop explored the physical and embodied implications of a physical interface. Participants worked together in groups with construction materials such as paper, card, PlayDoh, fabric, markers and various small objects (buttons, seeds) in order to individually create interaction objects. After the design stage, each team played, demonstrated and enacted with each other’s concepts. A set of playtypes were selected as ‘best’ concepts to test with a ‘Wizard of Oz’ version of the responsive system.

Workshop 5: No buttons
The workshop was a response to outcomes of workshop 4 and 3. The aim was to individually evaluate the pairing of our evolved interaction object with the navigational model.

Workshop 6: Preface
The workshop explored and evaluated a series of approaches to the audio display and interaction based on the model of a conversation. We developed several approaches to the idea of a ‘preface’, and ‘telling’ components of a conversation for their effect on the turn-taking dynamics of the interaction model. We aimed to find a range in the audio display that encouraged discovery and play in engaging with the audio information. Participants evaluated the approaches with a desktop prototype of the audio display engine.

3.3 Design outcomes and findings
The design outcome for an interaction model was a combined gesture and turn-taking approach based on a conversation model. In response to the workshops and scenarios we moved toward deictic gestures, such as pointing, or the gesture of a hand holding a glass. We employed a further constraint by introducing an object and thus achieving greater consistency of movement. This approach was suggested in an earlier workshop. Further workshops allowed us to develop the form of the object and mnemonic mapping of the navigation to the object and gesture. An interesting finding was that participants would tend to make device-like objects with analogue push-button or contact-based interfaces, while during ‘testing’ they would prefer the simpler non-device like objects and manipulations. This challenged our notion of what we may now consider a ‘natural’ interface.

The design outcome of a navigation
Wakkary

model emerged from exploring paper prototyping within the emerging conditions of the interaction model, audio display and system. In large part, it was an integrally parallel set of design activities. The resulting model is a simple structure that we referred to as ‘1-2-4’. Based on our conversation model approach, visitors would hear three different topics to choose from in the form of audio prefaces, lead-ins to longer narratives, and if a topic is chosen it is replaced with a new and related narrative, while the previous choices are replayed. In our workshops, we used modified Trivial Pursuit™ cards to emulate the over eight-hundred audio objects within our repositories. Workshop participants felt the model, the ‘1-2-4’ approach was more responsive and ‘tailored’ to their interests.

The audio display model was pragmatically based on the navigational model and the development of inference rules related to the reasoning engine, while theoretically based on a conversational model. The workshops and scenarios explored different approaches in the presentation of the audio with the aim of allowing visitors to effectively discern between thematically or conceptually different information objects. We also confirmed that the conversational approach was appropriate to maintaining a level of playful engagement and dialogue. Part of our findings related to the style of presenting artifact information. The outcome was audio prefaces written such that they had a ‘teasing’, humorous quality, and were recorded with a diverse range of voices. Other findings were consistent with the theoretical insights we had been discussing and researching. The preface or conversational approach clearly piqued curiosity, encouraged more exploration and continual turn-taking. Participants found it easy to understand, to use and more familiar. Comments included, they felt more ‘focused’, “it’s nice to hear the ‘emotion’ in the voice”, “it’s not too dry and academic”, and “it doesn’t feel like an automated machine”.

4 Discussion: the non-rational approach and limitations

Schön makes the argument against what he refers to as ‘technical rationality’, whose basis lies in logic and reasoning in theory outside of practice. The rational in his view is the logical abstraction of thought and action (Schön 1983). Gedenryd, in analysing design methodologies, traces the rational view of design practice to math, logic and cognitive science. In his arguments, design methodologies adopt a rational approach to design methods characterised by the common principles of separation and sequence (Gedenryd 1998). Design, based on traditional cognitive science has to date, adopted the linear model of analysis, synthesis and evaluation. To a large degree this approach has been inadequate in addressing the ordinary complexity of design situations. In contradiction, the practice detailed in the case-story can be described as ‘non-rational’, an approach that guards against abstracting views of practice, and rather grounds interpretation and reflection on practice—in practice. In Gedenryd’s terms, the approach is interactive and in the world, rather than intra-mental. In other words, the movement from design decision to design decision in workshops or scenarios is like dead reckoning in navigation where one determines the next destination in absolute reference to the last. And the movement between workshops and scenarios is like rotating puzzle pieces or Tetris pieces to find the right fit—it can only be done through action. In relation to Dourish, complexity is seen as an ‘interactional’ issue where it is understood through and by action, as opposed to a ‘representational’ issue—it simply cannot be mapped out beforehand (Dourish 2004). Situated in context of place, people, and technical systems, the design approach encourages action to manipulate the context through playtypes such as Play-Doh objects, ‘Wizard of Oz’ systems, and
of course, directly through dialogue with participants and designers.

In the case-story we can see design practice as asymmetrical and non-uniform. Each response has its own degree of exponential possibilities or divergences that are pragmatically explored until a set of relationships and a contingent form occurs that we understand to be the design outcome. In some sense, the described practice is a heterarchical form of networked relationships as opposed to the hierarchical form of traditional design methods.

However, certain limitations in the case-story are evident areas of further research in practice. This includes further research in the analysis and observation of design practice. This would naturally entail work on the methodologies for documentation and analysis of design practice, including first-person methodologies and ethnography in action. It is by far more common to record interactions and real-time observations of ‘users’ but not practitioners. Secondly, in architecture, it is self-evident to build on site and in a sense the need to design in situ cannot be taken too far. The workshops in the case-story could have been enacted in the museum environment in many cases and therefore taken out of the lab and studio. Lastly, complexity in interaction is often in human to human relationships—ec(h)o is still very much an exploration of human to system interaction.

Another alternative to the rational approach in design methods is to consider methods as parts of loose toolkits that designers carry with them to be applied as they see fit (Löwgren and Stolterman 1999). Such ‘second generation’ of design methods has arisen with a focus on collaboration and creativity over the systematic formalising of design processes (Broadbent 1979), yet without a broader conceptual underpinning within design issues such as context, embodiment, or complexity this approach to design is compromised and minimal in its impact in addressing the full range of design situations.

By way of a conceptual underpinning, the idea of complexity is proposed as a framing experiment for a translation, synthesis and augmentation of design practice and the role of context, human activity, and experience in design. Complexity is used as a descriptive term and not in the mathematical sense. The aim is not to quantify or create an abstraction of design practice based on the epistemologies of mathematics or science. I do not believe that is possible. Rather, complexity is proposed as a descriptive term based on the general understanding of the complex and the emergence of the term in design and HCI theory and practice.

Limitations to current theories related to complexity are readily admitted. Aside from the inherent challenges of a theoretical rethinking, methodological concerns question the first-person narrative and or situated accounts of ethnographic or phenomenological approaches (Nardi 1996, Dourish 2004). Advocates of the contextual approaches based on activity theory are quick to point out that the theory is a research framework in psychology for understanding individual consciousness that, while having strong implications for better understanding the nature of human activity and interactive technologies (Nardi 1996, Gay and Hembrooke 2004), has proven far more challenging to mobilise into informed practice. The related ecological approaches have only begun to examine how to bridge the sociologically-based methods of ethnography and participatory design with emergent systems that characterise information ecologies (Nardi and O’Day 1999). Reflective practice has mostly remained a descriptive theory in design. And so the very idea of a theory of design and complexity is in reality a call for a new research agenda.
5 Conclusions

Both design and HCI offer theoretical frameworks as starting points for addressing complexity. The paper offers a beginning point for further synthesis, critique and analysis of these concepts in order to offer a clearer view of where we stand. The challenge of better understanding the role of complexity and the resulting implications is not a trivial pursuit, nor is there a singular or even delineable goal. Rather, it is a question of framing a broad research agenda. Yet it is reasonable to argue for a shift in design and design practice through the framing of complexity that will potentially address human and social experience, the reality of practice, and the need to design in a physical, human and experiential sense.

Design is a human activity that addresses human and social experience. A re-orientation away from tasks and efficiency in design objectives is required to address the fuller aim of responding to complex human experience. Activities are characterised by simultaneous, diverse and cross-boundary tasks and aims. As a consequence, interpretation, contingency, exploration and negotiation are viewed as necessary and favourable attributes of design responses. A key challenge is to develop approaches to evaluation that can identify and incorporate these lesser quantifiable attributes.

Design methods create from within complex and situated actions that are unable to be modelled or represented fully. In turn, this requires methods that are situational and dynamically interact within changing contexts. Design methods do not need to predetermine the process or steps, rather there is a need to anticipate the dynamic nature of the action and responses that happen while designing. These responses are situated and embodied rather than abstracted and conceived as a mental model—they can be seen as non-rational. This calls for understanding design practice as emergent and interactive.

The integral relationship in design of activity and context is characterised by instability, contingency, and interrelatedness. It is not possible to fully model or reconstruct this dynamic or reduce it to primitives and essences. Since situations and contexts are not easily represented, design and analysis is performed in situ. The dynamic of interaction and context is seen in the world and not in the studio or lab.

We design for the activities of people in situations or situated participation, and require design methods to support this approach. In two respects, interaction is social rather than individual. Our actions are predominantly people to people in which technology can facilitate, mediate or intervene. Secondly, even as individuals acting with technology we are social in nature in that we are interwoven into a social context and history. We are social in the sense that our actions through technology are either mediations or interventions into our social world and it should be assumed that such actions involve diverse and simultaneous experiences that extend well beyond the definition of task or user.

Acknowledgements

Work presented in this paper is supported by grants from the Social Sciences and Humanities Research Council and Canarie Inc. I'd also like to gratefully acknowledge the co-investigators in the ec(h)o project, Dr. Marek Hatala and Kenneth Newby, as well as the support and efforts of Robb Lovell, Dale Evernden, Gilly Mah and Doreen Leo. The project could not have happened without the generous support of the Canadian Nature Museum, especially Dr. Mark Graham and his team.
Notes

1 Daniell Fallman contends that in HCl, there is little evidence of interest or concern for the designing of systems or prototypes. A syndrome he refers as ‘it just happens’, see Fallman (2003).

2 Dourish in discussing his related concept of embodied interaction has commented on the ordinariness of conversation as discussed by Sacks in his analysis of conversation. Dourish argues that ordinariness is a feature of context as it is understood within embodied interaction, see Dourish (2004). This argument can equally be applied to complexity. What we then understand is that complexity is not a factor of scale (largeness) and extra-ordinariness.

3 Rittel argues for a hermeneutic approach to design, see Rittel and Webber (1973).

4 For details of the qualitative and quantitative evaluation of echos, see UMAI (User Modeling and User-Adapted Interaction) forthcoming special issue on ubiquitous computing and user modeling.

5 See for example the pioneering longitudinal study on the practice of engineering, specifically an aspect that investigated reflective practice through the use of protocol analysis (Adams, Tums and Atman 2003).

References


Gastronomica - the Journal of Food and Culture 2(2) 42–62.


Ron Wakkary is Associate Professor in the School of Interactive Arts and Technology at Simon Fraser University in British Columbia. His research includes design theory, design methods and interaction design in mobile and ambient intelligent environments. He has presented and published widely, including Computer Human Interaction ACM, Siggraph, Internet and Consciousness Reframed. He was cofounder of Stadium@Dia in New York where he collaborated and co-developed pioneering projects in art and the Internet. He is currently leading the Am-I-table Network for Responsive and Mobile Environments, a research network funded by Canadian Heritage.
A New Educational Model for Interactive Product Design

THE INTEGRATION PROJECT

New technologies are fundamentally changing the way we learn, work and play. Technical knowledge and understanding alone are inadequate to deal effectively with many of the implications of new technology. This raises questions concerning both what technology can do and what technology should do. In either case, the products, systems and services we create with new technology are of little value if we do not readily understand what they are, what they do and how to use them.

One of the keys to addressing the complexity of interaction is a balanced understanding of both the technical (utilitarian and performance) issues and human (social and cultural) considerations. To ensure that the solutions to these complex problems meet design expectations, it is critical to integrate an active prototype testing and validation process into the design development cycle.

Design Education and New Technology

Changes in technology are forcing educators to rethink the role of design education with respect to both business and society. At the Institute of Design in Chicago, John Heskett has been investigating the role of design in everyday life: "Design is simultaneously becoming more specialized in some respects, with more detailed skills in specific areas of application and more generalist ones in others, with hybrid forms of practice emerging....On another plane is the difference between designers as form givers, determining form in a manner that allows no variation... or as enablers using the possibilities of information technology and powerful miniaturized systems to provide the means for users to adapt forms and systems to their own purposes."

The implications of Heskett's analysis are significant. Design education has reached a fork in the road, and each school will have to judiciously plot its own future trajectory. Currently, there are too many alternatives requiring too many prerequisites for the future path to be clearly evident.

In recognizing the impact of the increasing role of computing in people's lives, Terry Winograd at Stanford University was among the first to identify a design practice whose outcome and focus was a qualitative process rather than a thing or an object. He labeled this new prac-
Three members of the CreATE team work on ideas for a new visually-based management tool for designers.
(L-r: Mikki Ho, Jag Poonian and Doreen Leo, missing Tanya Hsu)
Cice "interaction design." Winograd identified the need to focus on the perceptual and psychological aspects of human experience by rooting interaction design equally in graphic design, psychology, communication, linguistics, and computing science.

Much has changed in the few years since Winograd identified the need for teams with diverse skill sets to focus on the individual user in efforts to provide better solutions to problems associated with desktop computing. Today's networked wireless technology represents a more significant challenge that by its very nature will require multidisciplinary solutions. We now have the capability to embed sensor technology in virtually any product. These pervasive computing devices and ambient technologies will be capable of responding to people in everyday situations. In addition, these devices will be capable of talking to one another or anyone else anywhere in the world. The implications are astonishing; the world as we know it will change.

Technology, Context and Experience
In 1999, Joseph Pine II and James Gilmore observed the increasing rate at which products were reaching market saturation, a point at which products were becoming a commodity, and suggested that circumventing the typical product price war associated with commoditization—adding competitive value to product or service offerings—would require businesses to shift from a product- and service-based economy to an experience economy, where business caters to lifestyle experience.

Design theorist Richard Buchanan similarly argues for a paradigm shift in all design fields. He has identified a historical trajectory of moving progressively through "four orders of design," from symbolic to things, actions and environment. Patrick Jordan, an expert in the field of human factors, carried this argument one step further, claiming that it will be critical for designers to develop a richer understanding of people in order to design the kind of pleasurable products to meet people's new lifestyle expectations.

If we combine these factors, it becomes apparent that we are now at the leading edge of technological change that will affect all aspects of everyday life in a profound way. The next generation of designers will need new skills and knowledge to negotiate this new terrain. Design schools must respond to this challenge. Problem-based learning and project-based learning provide useful models. The goal is to develop constructivist, project-based learning environments combined with a reflective practice approach to design.

The key pieces to this puzzle include teamwork; the need to develop a working knowledge of new technologies; the need to develop an understanding of the way people live, work and play; and, perhaps most important, the need to prototype and validate the new design concepts.

Teamwork: As designers consider more knowledge about the product cycle during conception and planning, the act of design becomes a more complex activity. As a result, product design is frequently done by teams of professionals that include social scientists who are trained to study the characteristics and qualities of human experience, along with designers and engineers. Although the necessity to work in teams on complex design problems is recognized in business and industry, the education system has been slow to follow. Initiatives undertaken at Stanford University, Arizona State University and the University of Illinois at Chicago have validated the success of multidisciplinary teams working in industrial design education.

Understanding new technology: We have seen a similar pattern in our abilities to assimilate new technologies into our education system. Pervasive or ubiquitous computing requires the integration of multiple technologies, including software and hardware, and an understanding of human-computer interaction. There is a need to foster a multidisciplinary team-based approach to overcome this hurdle. In order to address the increasing emphasis on the design of functional products within education, design students will require a stronger foundation in the basic elements of technology.

Understanding people: The third piece of the puzzle is the necessity to develop a more thorough understanding of people and the quality of human experience. Elizabeth Sanders, president of Sonic Rim and an adjunct faculty member at Ohio State University, argues that "the people that we design for are the real virtuosos of the 'experience domain.' They are the ones who will create their own experiences." Accordingly, she suggests it is
Small interdisciplinary teams of students design, prototype and field test their own ideas for innovative applications of new technologies.

essential that we use new tools to encourage and engage these ordinary people in the design and development process to help us learn to better satisfy their needs. The late Paul Rothstein, former professor at Arizona State University, developed a research and design method called “a (x4)” to provide designers and educators with a tool to develop design scenarios about user experience. These tools and techniques are providing designers with a better understanding of the implications of lifestyle changes. They are also providing the key to new participatory design methods—to engage the prospective audience throughout the design development process.

Prototyping and validation: The final step is the ability to integrate the individual pieces. The only way to do that is to place the product in the hands of the prospective users. Industrial design has long recognized the importance of prototyping. In addition, the development of effective intelligent interactive products and systems is a complex process with significant social implications. Without prototypes to support the viability of new concepts, many ideas will remain unsubstantiated and highly questionable. For these reasons, prototyping should be an intrinsic part of the development process for this new generation of interactive products and systems.

Bridging the Disciplines
Our research demonstrated that an interdisciplinary approach was needed to effectively address these complex problems. It would also be necessary to work toward a well-balanced understanding of both the technical and the human considerations. In effect, this pointed to a new curriculum built around a core combination of design, information technologies and human-computer interaction, with additional support in the areas of cultural studies, electronics and business. The dilemma was the logistical nightmare of adding all of these requirements to an already overloaded design curriculum.

For the past four years, we have been building the infrastructure for a new interdisciplinary university program to equip a new generation of undergraduate and graduate students with the knowledge and skills to tackle the full potential of interactive products, systems and services. Initially, this program was the central component of the new Technical University of British Columbia, which started in 1999. In 2002, the program became part of the larger Simon Fraser University where it is now part of a department known as the School of Interactive Arts and Technology.
Tools & Techniques

Examples of useful pre-built components or building-block type hardware and software components to help students get started include:

- LEGO Mindstorms Robotics Invention System: The heart of the Robotics Invention System 2.0 is the RCX™, an autonomous LEGO microcomputer that can be programmed using a PC. The RCX serves as the brain of LEGO Mindstorms inventions.
  http://mindstorms.lego.com/eng/products/ris/index.asp

- LEGO ROBOLAB: ROBOLAB is an educational program developed as a joint venture between National Instruments, LEGO Dacta and Tufts University to help develop engineering intuition. ROBOLAB uses a powerful combination of LEGO bricks and National Instruments LabVIEW graphical development software to introduce engineering concepts to students of all ages.
  www.ni.com/company/robolab.htm

- Phidgets: These 3D widgets simplify software development in interface design and give designers the ability to plug together hardware components and to focus on the programming aspects of interactive product development. Phidgets arose out of a research project directed by Saul Greenberg at the Department of Computer Science, University of Calgary.
  http://group/lab.cpsc.ucalgary.ca/phidgets; www.phidgets.com

- e-Gadgets: The EU IST/Future Emerging Technologies Research project seeks to adapt to the world of tangible objects the notion of component-based software systems by transforming objects in people’s everyday environment into autonomous artifacts (e-Gadgets). The e-Gadgets range from simple objects (tags, lights, switches, cups) to complex ones (PDAs, stereos), and from small ones (sensors, pens, keys, books) to large ones (desks, TVs).
  www.extrovert-gadgets.net

- Max/MSP Scripting Language: Max/MSP is a graphical environment for music, audio, and multimedia. In use worldwide for over fifteen years by performers, composers, artists, teachers and students, Max/MSP is the way to make your computer do things that reflect your individual ideas and dreams.
  www.cycling74.com/products/maxmsp.html

- Pure Data (PD) Scripting Language: PD is a real-time graphical programming environment for audio, video and graphical processing. It is the third major branch of the family of patcher programming languages known as Max (Max/FTS, ISPDW Max, MaxMSP, jMax) originally developed by Miller Puckette and company at IRCAM. The core of PD is written and maintained by Miller Puckette and includes the work of many developers, making the whole package very much a community effort.
  http://puredata.info

- Teleo™: This rapid-prototyping and development tool, developed and marketed by MakingThings, consists of a line of modular and networkable hardware components that can be connected to a computer via USB and programmed and controlled using any one of a number of programming languages. Components range from a variety of input and output modules to motor controller modules and accessories.
  www.makingthings.com/teleo.htm

- NADA Rapid Prototyping of Physical Interfaces: NADA lets users integrate sensors, graphics, animation, sound and electrical devices as interactive objects and environments. It is a server-like application that provides hardware I/O services to projects being made in Macromedia Flash MX2004 or later and Java™. Because these services are 100 percent network-accessible, a NADA project can use multiple systems connected by a network for incredible flexibility.
  www.sketchtools.com

- Basic Stamp: A basic stamp is an easy-to-use microcontroller made by Parallax. The stamp contains a microcontroller, memory, clock and voltage regulator in a package that resembles an integrated circuit. All that is needed to program it is a PC and a 9V battery or other power supply.
  www.parallax.com/html_pages/products/basic_stamps.asp

- Arduino: Arduino is an open-source physical computing platform based on a simple I/O board and a development environment that implements the processing/programming language. This is an open-source project, owned by nobody and supported by many.
  http://arduino.berlios.de

- Processing: This open-source programming environment and language is designed to help people program images, animation and sound. The language is used by students, artists, designers, architects, researchers and hobbyists for learning, prototyping and production. It was created to teach fundamentals of computer programming within a visual context and to serve as a software sketchbook and professional production tool. Processing is developed by artists and designers as an alternative to commercial software tools in the same domain. Processing evolved from ideas explored in the Aesthetics and Computation Group at the MIT Media Lab and is an open project initiated by Ben Fry (Broad Institute) and Casey Reas (UCLA Design/Media Arts).
  http://processing.org

- Wiring: This programming environment and electronics I/O board is used to explore the electronic arts, tangible media and teaching and learning computer programming and prototyping with electronics. It illustrates the concept of programming with electronics and the physical realm of hardware control, which are necessary to explore physical interaction design and tangible media aspects. Wiring started at the Interaction Design Institute Ivrea and is an open project currently developed at the University of Los Andes.
  http://wiring.org.co
Our goal in developing a new curriculum is to foster a better understanding of the need to develop integrated solutions to meet the individual social, cultural, environmental and technical issues associated with emergent technologies. Detailed analysis, advanced prototyping and user field-testing are integral elements of the new curriculum. Typical projects focus on opportunities to capitalize on wireless, networked technologies and fall into categories ranging from software applications and electronic games to hybrid software/hardware concepts for ubiquitous computing devices and/or ambient technologies.

Although this course has a face-to-face lecture component, there are no traditional design studio facilities. There is a 1.5-hour lecture per week for the entire class and three 1.5-hour lab sessions for each section of 20 to 24 students. Reference materials and discussion forums are accessible online to complement the lecture sessions. Students also have unlimited access to computer lab facilities.

The Curriculum

The Integration Project runs for a full academic year, with the fall semester devoted to the concept and the spring semester to the realization of that concept. Students have the freedom to speculate and experiment in the first semester with a clear reminder at the beginning of the second semester to recognize the scope of the deliverables due by year-end. In addition, each team is encouraged throughout the project to identify potential faculty or industry mentors to assist with various aspects of the project.

At the end of the course, each team presents a detailed design concept, which typically includes preliminary models and/or proof-of-concept technical models to substantiate the feasibility of the project. Identifying a concept, reflecting on how it has changed over time, scoping out the feasibility and logistics of producing an operational prototype, as well as field-testing and user evaluations are key elements of the development process.

Key Issues

The complexity of the technical aspects of prototyping interactive products and systems is a significant challenge, particularly for those in the design community who do not have an extensive technical background in software and hardware development. Our experience indicates that a building block format for these components can help get students started with some degree of comfort.

For many teams, this level of prototyping is entirely new and requires a significant level of support and encouragement. Yet, results of the first class exceeded expectations: all 12 teams successfully produced and field-tested advanced operational prototypes of new product concepts. The requirement for students to engage a third party to review and independently test and evaluate projects they have developed has added a significant degree of motivation to excel. A particularly gratifying result of the course was that 2 of the 12 teams in the first course entered a commercially sponsored entrepreneurship competition with more than 100 entrants, and one of these teams successfully competed to the semifinal round. One of the project teams from the second year of the Integration Project course placed second in Microsoft’s Imagine Cup competition.

Looking Back

Learning to design with new technologies presented several challenges that forced us to rethink our approach to design education. We focused on developing a course that would place the onus on design students to consider technology based on the needs it served. The additional requirement to assess feedback from real users heightened the sense of reality surrounding the project and forced teams to scrutinize the resolution of their concepts.

As the course progressed, it became readily apparent which ideas held merit and which fell short of the mark. There was much to learn from the new process, as the pros and cons of each project were open to discussion. Overall, there was a clear sense that this student cohort was beginning to develop a better balanced understanding of the relationship between the technical and human issues in their design thinking.

References


Elizabeth Sanders, "Virtuosos of the Experience Domain" in the Proceedings of the 2001 IDSA Education Conference.
Experience and design methods: cross-dressing and border-crossings

Ron Wakkary¹ and Simon Niedenthal²

¹ Simon Fraser University, Canada
² Malmö University, Sweden

rwakkary@sfu.ca
simon.niedenthal@k3.mah.se

Introduction

We are pleased to present this special issue on experience and design methods. This issue has its origins in a workshop organised by Ron Wakkary, Thecla Schiphorst and Jim Budd (2004), titled Cross-dressing and border-crossing: exploring experience methods across disciplines, held as part of the ACM SIGCHI 2004 conference in Vienna, Austria. The workshop was to all appearances among the strangest gatherings at SIGCHI 2004. Blindfolded researchers were spied wandering the fluorescent-lit lower concourse of the Austria Center in Vienna, or dashing from table to food-court table in a tag-like game whose rules remained obscure. But the playful, seemingly aimless movement of the participants belied a serious purpose. The practice of designing interactive objects, environments and systems has outgrown its roots in Human Computer Interaction (HCI), with its residual load of cognitive science. We engage in interactive experiences out of choice, in search of the pleasures of the mind and senses we have hitherto found in performance, art or our experience of architectural spaces. For the workshop organisers and participants, a provocation is required: hence, cross-dressing, the opportunity to try on new roles and experience design methods coming from art and design—disciplines often marginalised by the research community.

The workshop received an overwhelming response and unfortunately only about fifteen percent of the position papers submitted were accepted. The workshop included participants across and within HCI (the academic discipline of the conference), including performance, visual art, education, computing science and interaction design. The participants were research practitioners who in the workshop shared design approaches and projects. This included presentations and body/brain-storming around the concepts; combining playful engagement with opportunities to reflect upon the emergent themes. The goal of the workshop was to explore unique methodological frameworks for designing technologically-mediated experience. What follows in this issue is a selection of papers from the workshop participants. Each paper addresses the need to make discursive, evaluative and ultimately more manifest pleasurable, puzzling, richly ambiguous and essentially more human aspects of experience within their designs.

Experience in design

As designers of interactive systems we find ourselves stretching the limits of methodological practices that enable us to explore, build, communicate and prototype experience. It seems that addressing human experience requires a 're-dressing' of design practice. Domains such as performance, dance, interaction design, industrial design, visual art and education each embody knowledge and enact rigorous methodologies for constructing experience. Each of these domains defines experience, experience qualities and attributes, and defines affordances for enacting (and re-enacting) experience as a fundamental methodological
tool in the respective discipline.

At the intersection of HCI and design, Terry Winograd identified a design practice whose outcome and focus was on the perceptual and psychological aspects of human experience by rooting interaction design equally in graphic design, psychology, communication, linguistics and computing science. (Winograd 1997)

A key genesis point in the evolution of 'experience' as a design concept is the work in the 1930s of the industrial designer Henry Dreyfuss (1967). Dreyfuss’ work in ergonomics lead to the publication of the *Measure of man*, an extensive database of human measurement to facilitate the design of products tailored to a 'standardised' human body. In the late 1960s ergonomics split into the related science, engineering and kinesiology-based field of human factors, the political and social movements in Scandinavia that became known as participatory design (Ehn 1992), and the design methodology of user-centred design (Nielsen 1993, Norman 1988). Design experience was seen in surprisingly different lights, one functional the other social and political. In the early 1970s, the democratic social movements led to concepts of increased participation and assertion of user experience within the design process itself, such as participatory design and in the anti-modernist notion of pattern languages by the architect and urban planner Christopher Alexander (Alexander, Ishikawa and Silverstein 1977). The increasingly critical role of the user in these design processes contributed significantly to the evolution of design. At the same time the phenomenon of space, time and environmental design—clearly the domain of architecture—also began to play an ever-increasing role in design. For example, the ethnographer Edward T. Hall helps us to understand the participatory role of people in communication environments and spaces (Hall 1976). Enabling the audience experience was also a key goal of theorists and practitioners of the fields of performance and theatre, namely Vsevolod Meyerhold (Cooke 1983), and later the work of theorist and theatre director Jerzy Grotowski (1968) and Augusto Boal (1979). This tradition directly informed the concepts of interactive design from the early work of Norman Bel Geddes (Marchand 1995) to today’s interactive technology experiences and environments (Dodsworth 1998, Murray 1997).

Recently, HCI theorists and researchers have identified issues of ‘context’, ‘situation’ and ‘embodied experience’ that have strained the traditional theories of HCI. As Nardi puts it,

> we are beginning to feel a theoretical pinch, however—a sense that cognitive science is too restrictive a paradigm for finding out what we would like to know. (Nardi 1996)

The understood need is to move the theoretical trajectory of HCI from a reductivist understanding of human cognition toward an understanding of situated human activity. Carroll writes:

> We do not now (and in fact may never) understand human activities in enough detail to merely list the attributes computer systems would have to incorporate in order to meet these requirements: Precisely what kind of computer will help people learn microbiology, choose a new job, or relax? Indeed, human society and psychology develop in part as a consequence of the contemporary state of technology, and technology is precisely what is running ahead of our understanding so rapidly now. Thus, we have little prospect of developing final answers to questions about the nature of human activity—certainly not at the level of detail that would provide specific guidance to designers. Our best course is to develop rich and flexible methods and concepts, to directly incorporate descriptions of potential users and the uses they might make of an envisioned computer system into the design reasoning for that
In response to the rigidity of cognitive science, ethnographic and scenario-driven methods have begun to take hold in HCI practice. Further along in this direction, an emerging set of 'context-based' theories for HCI have adapted ideas from an even wider spectrum of psychological, social, political and philosophical theories based on understanding human activity (Dourish 2001, Nardi 1996, Nardi and O'Day 1999). Such attention to the richness of context and human experience has emerged over the years in HCI theory, less so with practical approaches for the designing of interactive systems.

Seeds for experience methods

Not surprisingly, use experience or user experience is a central concern in several of the contributions in this issue. Per Linde et al. view interaction design as a constant re-evaluation of place, material and methods in the service of qualities of use that differ from traditional HCI criteria of efficiency, ease of use and learning. Rather qualities of use in interaction design incorporate aesthetic experience and socially meaningful activity. This paper explores the quality of experience as presence in space and materiality that can be perceived as place. The paper explores a range of design methods from tangible computing games, to improvisation and performance, to augmented mixed media as a presence-making tool. Jacucci and Isomursu add to the approaches of participatory design and ethno-methodology with a concept of design happenings. Drawn from the art movement of happenings in the 1960s of such artists as Vito Acconci and Alan Kaprow, the authors propose design happenings as a resource for experience and its expression in order to more directly capture use experience. Gromala and Shaw address the direct experience of pain and the need for tools that account for its subjective aspects such as type, intensity and duration. They suggest a method for communicating highly subjective experiences.

In addition to tools and use experience, other authors addressed the role of existing design methodologies. Pamela Jennings recast the traditions of computer-supported collaborative work (CSCW) as CSCP, or collaborative play, as a means of attacking complex design problems in her Constructed Narratives project. Her paper provides an account of this approach and delves into the role of interdisciplinary collaboration inherent to such an approach. PARC researchers Allison Woodruff and Paul M. Aoki discussed the use of conversation analysis techniques in the context of designing an electronic guidebook for a historic home and extend its use into social and mobile audio spaces. Eva Hornecker meditated upon the relationship between didactic/facilitation methods in learning and the designing of interactive systems and spaces in light of experience design. Artist Sarah Rubidge and composer Alistair Macdonald present their Sensuous Geographies installation that examines the benefit of our aural experience and the role of performance and audience in interactive installations. Other participants, whose papers are not reproduced here, but who added immeasurably to the workshop, included Phoebe Sengers, Kirsten Boehner, Jofish Kaye and Maria Håkansson.

Like the workshop, the aim of this special issue is to plant the seeds of cross-disciplinary dialogue and to cross boundaries, assuming other roles in order to experiment methodologically and to establish a new common knowledgebase aimed at design and human experience. We believe these papers succeed in planting new ground for new research and practice.

We would like to gratefully acknowledge all the authors in this issue for the privilege of working with them and all
the participants in the workshop. In addition, we would like to thank Thecla Schiphorst and Jim Budd, co-organisers of the workshop for their invaluable contributions, mediation and guidance.

References


RELATED PUBLICATIONS: CONFERENCE PROCEEDINGS


UNDERSTANDING AURAL FLUENCY IN AUDITORY DISPLAY DESIGN FOR AMBIENT INTELLIGENT ENVIRONMENTS

Milena Droumeva
Faculty of Education
Simon Fraser University
mdroume@sfu.ca

Ron Wakkary
Associate Professor
School for Interactive Arts and Technologies
Simon Fraser University
rwakkary@sfu.ca

ABSTRACT

This paper presents the design and some evaluation results from the auditory display model of an ambient intelligent game named socio-ech(h)o. socio-ech(h)o is played physically by a team of four, and displays information via a responsive environment of light and sound. Based on a study of 56 participants involving both qualitative and preliminary quantitative analysis, we present our findings to date as they relate to the auditory display model, future directions and implications. Based on our design and evaluation experience we begin building a theoretical understanding for the unique requirements of informative sonic displays in ambient intelligent and ubiquitous computing systems. We develop and discuss the emerging research concept of aural fluency in ambient intelligent settings.

1. INTRODUCTION

socio-ech(h)o is a prototype environment for group play whose goal is to explore issues of design, use and evaluation for ambient intelligent systems. Ambient intelligence (AmI) computing is the embedding of computer technologies and sensors in architectural environments that combined with artificial intelligence and multi-modal displays, respond to and reason about human actions and behaviors within the environment. Our broader research with socio-ech(h)o is focused on several related issues including motion sensing, light and sound display design, user modeling, and interaction models. The overall research goal is to understand how to support groups of participants as they learn to manipulate an ambient intelligent space, as well as to better understand how to design ambient components of a responsive environment capable of providing this type of support. The research questions are numerous in a project of this nature and yet immersive and embodied interaction does not lend itself to reducible and isolatable variables that can be measured independently. Given this, our study design has focused on an ecological investigation and theory-building, aiming to provide broad, yet particular set of heuristics that help describe and make sense of the different system and display components of AmI environments. The empirical quantitative results, which include a survey instrument, time of completion and accuracy measures, as well as video coding analysis, are only one part of the picture, providing supportive evidentiary data, which, along with our overall observations and design process reflections, contribute to building a fuller picture of the role, unique characteristics and requirements of ambient intelligent systems.

This paper will specifically address the auditory display design of socio-ech(h)o. First, we make room for a methodological discussion related to approaching rich, complex settings in order to situate our inquiry within the larger paradigm of interaction design as well as ecological methods of auditory perception research. Based on the particular model of sonic feedback in socio-ech(h)o, we offer some results to date along with a preliminary attempt to formulate key factors and concepts towards building a framework of auditory displays in ambient intelligent environments. A second reason for a focus on the methodological aspects of the study is that the main contribution of the paper, a concept of aural fluency arose from an emergent set of questions afforded by this approach to research. In this paper, we outline the initial definitions and resulting heuristics for aural fluency that we believe will lead to further investigation into auditory display design for ubiquitous computing systems.

Figure 1. socio-ech(h)o gameplay, level 4. Participants work together to solve the puzzle Big Bang and build a big fire from quiet crackles.
2. METHODOLOGICAL DISCUSSION

There is the design challenge of crafting an informative auditory display for an ambient intelligent system, and then there is the research challenge of building knowledge about the role of sound in ubiquitous computing environments, as well as the more concrete goal of adding to auditory perception research via ecological investigations. Since at present there are still relatively few examples of complex auditory displays where sound is made meaningful, essential, communicative, engaging and social within a ubiquitous computing setting, it is hard to draw methodological and epistemological examples to guide our present inquiry. In addition, the research into actual ubiquitous computing environments is still in the process of bridging knowledge between psychology, cognitive science, human-computer interaction. Controlled experiments into auditory display issues are simply insufficient at this point in time in addressing these complex requirements of ubiquitous computing, as we don’t yet understand enough about their foundational components. The issues that motivate this current investigation have to do with both understanding the role and requirements of informative sound feedback in ambient intelligent systems, as well as determining best practices in perceivability [1] of sound in such contexts.

In traditional psychology experiments on auditory perception, as well as in quantitative human-computer interaction methods, there is rigid stimulus-response formula. Such studies are good for determining absolute, static thresholds and values of human auditory perception. Some auditory display studies, particularly Kramer and Gaver’s work [2] in auditory icons, and more currently the works of Kramer, Walker, Brewster and Barrass in acorns and sonification [3, 4], have gone farther by situating the stimulus-response formula in a real task situation with a meaningful activity that uses sound feedback. Among others, McAdams and colleagues’ work on environmental sound perception [5] also takes the inquiry further by examining how people’s experiences with everyday, environmental sounds happen within different models of perception, cognition and action, than is suggested by laboratory tests with synthetic tones. Sonification studies, as a subset of auditory display research, focus on determining measures and thresholds of auditory perception responses of continuous, changing sound in terms of scaling, polarity, data-to-parameter mapping and spatialization [1, 3]. Contemporary sonification studies take context and other user-centered design issues into account in the research inquiry, as well as emphasize considerations such as time-prolonged exposure to sonifications; auditory and cognitive fatigue; and aesthetic aspects as an element of efficiency of the sonification [6]. The use of non-traditional sound categories for display - using vocal or other non-synthetic content [7, 8] as well as the role of interaction sonification [8] have also been investigated, though by few. In any case, the unit of analysis is always the human response, measured as a threshold, a just noticeable difference, a correct or incorrect response - all in an attempt to capture generalizable human perceptual and cognitive abilities and predict reactions.

On the other hand, while the field of human-computer interaction and interaction design place emphasis on user-centered design in developing and researching technology, they rarely deal with investigating sound in a significant, systematic manner. When they do, it is often through traditional evaluation and usability methods, which still may not provide a setting that is ecological and holistic enough to allow for building and validating contextual, situated knowledge about sound in ubiquitous computing environments. As Gibson points out, “...awareness is rooted in meaningful experience of the environment; thus ecological validity results from studying subjects/people in their own natural environment, in motion, in active exploration. For people this environment is social, cultural, systemic, economic, political, etc.” [9]. Examining social contexts, group interactions, and embodied experiences and their interplay with sound, listening and soundscapes, could not therefore be achieved through traditional psychology or usability methods alone, and require, we argue, adopting the situated, activity-based approach of participatory design methods in order to achieve this task better [10, 11, 12].

3. PROTOTYPE DESCRIPTION AND AUDITORY DISPLAY MODEL

The socio-ec(h)o game is played by a team of four players and features six levels of increasingly difficult word puzzles solved by coordinated body positions and movements – see Table 2 for puzzle titles – for example, in Lo and Behold all four players had to be crouching low in the middle of the space and hold that position for around 3 seconds. For more details on the game prototype see [13]. The environment is responsive to players’ actions through abstract light and sound and the ambient response is driven by an artificial intelligence module that reasons about players’ actions in space. Players’ movements are tracked using a Vicon motion capture system (www.vicon.com). The goal was to create an ambient system that enables users to learn, use and manipulate it for a problem-solving activity. Based on our preliminary participatory design workshop results, already outlined in [14], we established that the interface required a gradient rather than direct response in order to best represent game states, as well as anticipate and direct players’ actions. In keeping with AmI design we wanted to move away from confirmatory "sound/no sound" feedback, which only signals if users do something right. Moreover we wanted to move beyond discrete incremental changes that still promote an “action-reaction” model of communication and are known to pose problems to short-term auditory memory [15]. Instead, we wanted to create a dynamically gliding intensity of sound that not simply responds but directs user actions and relies on perception of change and just-noticeable differences of different sound parameters. We even introduced an artificial 2-second delay in updating sonic feedback to the game state, in order to push the interaction away from an instant-feedback scenario and towards more complex, attentive and reflective listening – acting model similar to a form of data sonification.

We termed this approach intensity-based gradient sound feedback model [more in 16]. This intensity gradient shifts smoothly between layers of sound and provides not only narrative complexity but also relays information and guides player actions through ambient real-time sonification (see Tables 1 and 2) - that is, tells players how ‘close’ they are to solving the puzzle. Naturally, principles of data sonification came at the basis of the auditory display model. The principles we borrowed included the concepts of scaling, polarity, data-to-parameter mapping and spatialization [3, 17, 18, 19, 20]. In addition, we adhered to general principles of psychoacoustics for just noticeable differences, effectiveness of different parameter shifts (pitch, versus amplitude, tempo or timbre, as an example) [19, 21]. In addition to guidelines from sonification frameworks [1, 3], our sonic feedback borrowed principles from soundscape composition and everyday listening [2, 22]. Finally, our choice of activity
format— a game— came because games are suitable for research, particularly one with an ecological onus, since they provide avenues of natural engagement, while at the same time game mechanics allow careful constraints to be built into the experience, and the quantifiable outcome (game solution) allows for testing basic effectiveness, comparing times of completion, and capturing rich interactions. The way sound display fit into the game as a core mechanic was in following the natural skill mastery progression of the game— the soundscape in Level 1 was perceptually easy to interpret (see Table 1) while the soundscape for Level 4 was significantly more ambiguous. Thus ambiguity, perceptual and aesthetic, rather than being avoided, became a core part of the game experience itself.

3.1. Audio Display Schema

The audio display for socio-ec(h)o consists of three components: a real-time ambient sonification engine that has a different soundscape for each level of the game; an anticipatory feedback sound to signal when all participants are working together towards a goal; and a confirmatory feedback sound, which signals the completion of the goal and the progress to the next level (see [16] for details). The latter two feedback signals are consistent through all levels, so as to create a coherent interpretation of success. For example, an answer to the word puzzle "to and behold" is for all players to crouch low and be still. Further, they must hold this configuration for a short period of time, typically 4 seconds. This ensures that the actions were purposeful and knowingly completed rather than an accidental formation. As they crouch lower, the sound feedback intensifies, and if they stand up again the ambient sound will de-intensify. When they achieve the desired position and hold it for a few seconds, an anticipatory feedback sound occurs signaling to the players that they’ve done something right and if they keep at it they will complete the solution. Upon sustaining the configuration, the level changes with an audio reward and a short video reward.

<table>
<thead>
<tr>
<th>Levels</th>
<th>Core soundscape</th>
<th>Approaches to intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Soft abstract musical soundscape</td>
<td>Amplitude only</td>
</tr>
<tr>
<td>L2</td>
<td>Soft abstract musical soundscape</td>
<td>Phaser-tempo (slight pitch and amplitude adjustments)</td>
</tr>
<tr>
<td>L3</td>
<td>Clinking pebbles soundscape</td>
<td>Pitch only (slight amplitude adjustments)</td>
</tr>
<tr>
<td>L4</td>
<td>Fire sounds</td>
<td>Cross-fade between five recorded fire soundscapes of increasing intensity</td>
</tr>
<tr>
<td>L5</td>
<td>Water stream sound</td>
<td>Phaser-tempo (slight pitch and amplitude adjustments)</td>
</tr>
<tr>
<td>L6</td>
<td>Forest ambience</td>
<td>Cross-fade between five recorded forest soundscapes from calm day to thunder storm</td>
</tr>
</tbody>
</table>

Table 1: Detail of the soundscapes and approaches to displaying auditory intensity for each game level.

As demonstrated in Table 1, sonification principles were used in conjunction with one another and alongside more narrative-based intensity approaches. Specifically, the cross-fader between five different recorded soundscapes of increasing intensity was something we wanted to test out against more aurally and perceptually simple approaches to intensity such as sliding pitch or rising/falling amplitude. Further, the introduction of this narrative sonic component supplemented the game mechanic of rising gameplay challenge— becoming harder to interpret.

Several factors posed challenges to our exploration of the effectiveness of these auditory design approaches. The fact that each level has a different puzzle presents a subjective element of experience that comes in addition to each level using a different soundscape and a different intensity model. Table 2 outlines some of the specific scaling values used in the Max/MSP audio engine. Even though sound parameter changes are stacked to complement each other, the scaling values of main variables were kept similar as much as possible with different base soundscapes. In addition, amplitude adjustments were necessary, especially in the negative polarity cases to counteract effects of pitch and timbre for the equal loudness effect. Negative polarity was chosen in the tempo cases to support a narrative and embodied connection to the puzzles— in the case of Level 2 the participants had to slow down in order to solve the puzzle, and at the same time the sound’s pulsing tempo was going down instead of up. Similarly, in Level 5 the soundscape was slowing down to encourage participants to experiment with still and slow movement, which was again required in order to achieve the desired configuration. Alternatively, in Level 3 users had to run faster than normal and the positive polarity of pitch and upbeat, crisp timbre of crunching pebbles provided a good connection between activity and response. Lastly, in the case of Level 4, the fire soundscape, which built up from crackles and sizzles to a full-blowon bonfire supported the motif of the puzzle, which was Big Bang! The physical solution was to start in the middle and ‘make the fire grow’ then move outwards to the edges standing up.

<table>
<thead>
<tr>
<th>Levels</th>
<th>Word Puzzle</th>
<th>Polarity</th>
<th>Exact Scaling Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>Lo and Behold</td>
<td>Positive</td>
<td>20 -140 (max 157)</td>
</tr>
<tr>
<td>L2</td>
<td>Shoe and Low like a Plum Turtle</td>
<td>Negative</td>
<td>1.20 - 0.8 pitch (1.00 is normal pitch)</td>
</tr>
<tr>
<td>L3</td>
<td>All Rolling in a Bowl</td>
<td>Positive</td>
<td>0.7 - 1.9 pitch bend</td>
</tr>
<tr>
<td>L4</td>
<td>Big Bang!</td>
<td>Positive</td>
<td>Cross-fade fire</td>
</tr>
<tr>
<td>L5</td>
<td>Gazing over Waves</td>
<td>Negative</td>
<td>1.4 - 0.8 pitch (1.00 is normal pitch)</td>
</tr>
<tr>
<td>L6</td>
<td>Swinging in the Ring of Fire</td>
<td>Positive</td>
<td>Cross-fade forest ambiance</td>
</tr>
</tbody>
</table>

Table 2: Detail of the game puzzles, polarity and scaling for each game level.

4. STUDY DESIGN AND RESULTS

In order to evaluate the effectiveness of our prototype design, as well as to investigate and build knowledge about interactions in space, ambient intelligence and auditory display for ubiquitous computing, we created a study protocol based on playing the socio-ec(h)o game. As mentioned in the beginning, a project of this magnitude entails many research questions and points of interest. The study design was meant to address a lot of them, however, not all are relevant to our current discussion of auditory display design and auditory perception, and thus would not be discussed here. 14 teams of 4 were enlisted to play the socio-ec(h)
o game. The experience was conducted in two parts— in the first participants have to complete the first four levels within a maximum time of 60 minutes. In the second part participants have to attempt to complete the remaining two levels within 15 minutes. The research questions we had upon entering the evaluation stage of this project were as follows: 1) How informative was the ambient sonic feedback within our responsive environments? 2) Did our approach to designing the auditory display support participants progression in the game and to what degree? 3) Is there any significance in the effectiveness of different soundscapes and approaches to conveying information that we encoded?

4.1. Research Instruments

All study sessions were videotaped for the purpose of video analysis later. The artificial intelligence module also generated a log of participant activity displaying game success, as well as position in space, as outputted by the motion capture system. The same intensity measure drives the environmental response, so by looking at the log we know or can recreate the precise auditory display at any point of the experience for later analysis. Time of completion measures were collected, along with verbal transcripts, video analysis and finally, after the game portion, participants were given a survey to complete. The survey instrument contained a combination of Likert-scale questions and open-ended questions relating to the effectiveness of the system’s response in helping, and guiding problem-solving, and creating an experience for them. Due to the enormity of the empirical data on this project, the numerous research questions, as well as lack of clear models to follow when interpreting results from complex, ecological studies of ambient intelligent systems we are still in the process of analyzing and interpreting results. To date, we have the survey results, time of completion measures, and some preliminary video analysis. In this paper we’ll focus specifically on the survey results, both Likert-scale answers and more open-ended comments. In keeping with design methods process we also investigate emerging research questions and offer conceptual leads and future directions.

5. FINDINGS AND DISCUSSION

At a base level, our evaluation study of socio-ec(h)o aimed at testing the effectiveness of our design as demonstrated by times of completion (ToCs) and successful solution measures. In terms of the auditory display, we hoped to see some consistency in performance in certain levels across teams that could point to a design success or flaw of a particular approach to sound feedback or intensity. The survey and transcripts in turn were meant to serve as indicators of what participants perceived and thought about the system.

![Average Times of Completion](image1)

![Soundscape Level Results](image2)

![Figure 2. Average times of completion, standard deviation and median values for Levels 1, 2, 3, 4, and 5.](image3)

![Figure 3. Breakdown of results of how effective each level was in both its directive qualities and overall like.](image4)

<table>
<thead>
<tr>
<th>Q #</th>
<th>Question Wording</th>
<th>Avg</th>
<th>Med</th>
<th>StDv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q 1.2</td>
<td>1.2. Did the system require a large effort for teams?</td>
<td>3.07</td>
<td>3.00</td>
<td>1.07</td>
</tr>
<tr>
<td>Q 2.1</td>
<td>2.1. How well did the system provide useful cues in helping you evaluate how close you were to solving a puzzle?</td>
<td>3.58</td>
<td>3.75</td>
<td>0.46</td>
</tr>
<tr>
<td>Q 2.2</td>
<td>2.2. How well did the system respond when you were having problems solving a puzzle?</td>
<td>2.9</td>
<td>2.25</td>
<td>0.55</td>
</tr>
<tr>
<td>Q 3.1</td>
<td>3.1. Do the lights and audio work well together to help you in the game?</td>
<td>3.54</td>
<td>3.75</td>
<td>0.67</td>
</tr>
<tr>
<td>Q 3.2</td>
<td>3.2. How effective was the audio feedback in helping you solve the puzzles?</td>
<td>3.92</td>
<td>4.5</td>
<td>0.68</td>
</tr>
<tr>
<td>Q 3.5</td>
<td>3.5. How well does the virtual system (lights, audio, video) integrate with the physical and architectural space?</td>
<td>3.52</td>
<td>4</td>
<td>0.52</td>
</tr>
</tbody>
</table>

Table 3: Detail of survey question wording and results on a Likert scale out of 5.
In fact what we see in the results is no consistency at all in ToCs of different teams. Not only did ToCs vary greatly across levels (see Figure 2.) but they also varied greatly within a given level across different teams. As the most striking example, Level 4 demonstrates a range of completion times between half a minute and 53 minutes! What is curious and cannot be explained by these performance results are participants’ ratings of the effectiveness of soundscapes and experience with audio response in different levels. In contrast to the completion times the questionnaire results are evenly distributed (see standard deviation values in Figure 3) and overall participants felt quite positive about the system’s accuracy and effectiveness regardless of their individual performance. Figure 3 shows most participants both preferred the soundscapes of level 4 and thought it worked best, thus showing no correlation between performance and preference for sonic feedback. So why did some teams take a long time to complete the puzzles, if they thought the sound response was good anyway and the game experience was positive?

To begin to address this question we analyzed the open-ended survey questions as well as the verbal transcripts from the sessions. While occasional players did not find the sound feedback useful, the majority of participants reported the sound being helpful and many described how ‘the warmer sound told us we were close’ while ‘the negative sound meant we were doing the wrong thing’. The majority also reported the sound being more informative than the lights system (which incidentally was not designed as an intensity gradient, but used a pre-determined schema of light cues). The main suggestion for improvement on the sonic feedback was to make it more instant – many players perceived a lag, which they felt was confusing. The second main suggestion was for the feedback to be clearer in its representation of intensity – perhaps utilizing a negative intensity, signifying when players are definitely on the ‘wrong’ track.

Investigating the relationship between soundscapes and intensity approach on one hand, and puzzle theme and physical configuration on another, revealed interesting results. Several participants commented in-game on Level 4’s soundscapes being ‘the sound’ of Big Bang! and related their progress in the solution to building up the fire and the concept of explosion with the word puzzle Big Bang! This aspect could in turn explain the popularity of Level 4’s soundscapes in the survey results (see Figure 3). Alternatively, many participants mentioned being confused or frustrated with the sound intensity approach in level 5, which uses a negative polarity (water pulse gets slower and slower as participants progress in the game). They specifically commented on the sound’s change not being ‘like waves’ – referring to the word puzzle Gazing over Waves. Finally, a few participants mentioned that they liked level 3’s sound because it ‘sounded exactly like what they were supposed to do to solve it’.

These comments, along with the survey results and transcripts begin to reveal the role of sound feedback in socio-ec(h)o study? The concept of aural fluency and first examine it through the lens of existing literature, then look for specific occurrences and pieces of evidence of it in our study results. From auditory perception research, the problem of aural training is well known. In the carcon and auditory icon fields a user must often formally or informally learn abstract sound signals and their association to digital or real actions [15, 17, 20]. Thus concept-to-signal mapping and especially built-in listener training is an essential part of designing auditory display systems. In data sonification, short-term auditory memory poses problems to users’ ability to correctly and precisely interpret data values and trends from a continuous flow of sound. This problem has been explored in a few studies, by introducing auditory scaffolding such a reference tone [19] and contextual information [20] as well as through interactive sonification [8]. None of these studies however address the long-term effects of use on aural competency when interacting with a technological system. Further, the type of auditory literacy and fluency that is needed and seems to develop in more physical, situated technological environments such as intelligent spaces seems more akin to everyday listening, as articulated by Tronix and Schafer (in natural settings), as well as Gaver (in more technological settings) [2, 22]. It is situated in a social, contextual and shared physical environment and takes on characteristics of listening in everyday life, including dynamic shifts of attention, listening comprehension and aural expertise. As Jonas Löwgren writes in a new paper on fluency in augmented spaces, “the nature of our design material [of interaction design] is information and communication technologies, and unlike the products of industrial design and traditional architecture, it is temporal, as well as spatial” [23]. Aural fluency, we propose, is not simply another way of introducing auditory training issues, it constitutes a tangible connection between auditory perception, narrative and embodiment – that is, users build up aural fluency as a holistic experience of entails specific type of sound feedback, related to a specific set of movements and a particular narrative theme. So if aural fluency is a growing and developing competency using a system realized through sound, what instances do we see of it in our socio-ec(h)o study?

5.1 Exploring Aural Fluency

In line with design research, we take the newly problematized concept of aural fluency and first examine it through the lens of existing literature, then look for specific occurrences and pieces of evidence of it in our study results. From auditory perception research, the problem of aural training is well known. In the carcon and auditory icon fields a user must often formally or informally learn abstract sound signals and their association to digital or real actions [15, 17, 20]. Thus concept-to-signal mapping and especially built-in listener training is an essential part of designing auditory display systems. In data sonification, short-term auditory memory poses problems to users’ ability to correctly and precisely interpret data values and trends from a continuous flow of sound. This problem has been explored in a few studies, by introducing auditory scaffolding such a reference tone [19] and contextual information [20] as well as through interactive sonification [8]. None of these studies however address the long-term effects of use on aural competency when interacting with a technological system. Further, the type of auditory literacy and fluency that is needed and seems to develop in more physical, situated technological environments such as intelligent spaces seems more akin to everyday listening, as articulated by Tronix and Schafer (in natural settings), as well as Gaver (in more technological settings) [2, 22]. It is situated in a social, contextual and shared physical environment and takes on characteristics of listening in everyday life, including dynamic shifts of attention, listening comprehension and aural expertise. As Jonas Löwgren writes in a new paper on fluency in augmented spaces, “the nature of our design material [of interaction design] is information and communication technologies, and unlike the products of industrial design and traditional architecture, it is temporal, as well as spatial” [23]. Aural fluency, we propose, is not simply another way of introducing auditory training issues, it constitutes a tangible connection between auditory perception, narrative and embodiment – that is, users build up aural fluency as a holistic experience of entails specific type of sound feedback, related to a specific set of movements and a particular narrative theme. So if aural fluency is a growing and developing competency using a system realized through sound, what instances do we see of it in our socio-ec(h)o study?
5.2. Dimensions of Aural Fluency

While the statistical results above attempt to address, albeit in a limited way, the measure of effectiveness of the auditory display as it informs, helps and directs participants within an ambient intelligent setting, theorizing about the auditory display points to aspects of conceptual knowledge about the role of sound feedback in such environments that we aim to explore and help define. Building upon the emergent notion of aural fluency presented, we articulate a set of preliminary heuristics about auditory display that stem from our socio-ec(h)o study.

We suggest that aural fluency is a level of competency achieved with and supported by dynamic, ambient and continuous sonification. Like a language, the audio is required to provide a consistent and intelligible structure and expression. Unlike a language, the issues are less complex but also the result of conscious design. There is the need to improve design, however the resulting audio must be credible in order to engage listeners to learn it. We offer four heuristics that help in the design of aural fluency:

Epistemic Dimensions: The way players had to make sense of the audio feedback was through both an analytical and embodied involvement. While balance in this respect is critical, a degree of ambiguity plays a role in motivating listeners to learn and understand. Many of the groups reported being on many occasions confused or unsure about the sound feedback; however it was often a matter of degrees of ‘clarity’ rather than complete confusion; and teams did not seem to want to quit because of it, in fact it seemed to invite more attentive listening and feeling of challenge. This concept relates to the notion of everyday listening that people are already proficient in from interacting with their natural environment – making sense of confusing, unclear, complex sonic situations by selectively focusing or shifting attention on different aural elements, and fine-tuning their ear to particular sound structures and qualities. This fine-tuning, in turn, relates to our idea of aural fluency.

Narratological Dimensions: Leveraging past associations and creating new ones mediates expectations of what particular sounds mean and how to engage with them. As mentioned, the concept of aural fluency helps make sense of findings that show a positive experience of the system and appreciation of the informative qualities of the auditory display in the face of long and inconsistent completion times. Further, when examined through the lens of everyday listening practices, as developed by Schafer and Truax [22], we see numerous examples of both players forming a narratologic association with sound feedback – “we got no fire, somebody has to keep the flame!”... as well as a seamless connections to embodiment – “so the fire builds up and we’re all still...what if we move towards the sound? What if... when the fire gets all crazy we have to move more with it?” This narratological relationship could also serve to explain why negative polarity was not a problem in Level 2, where the solutions required moving very slow (so the slower they move the slower the sound gets) but it was a problem in Level 5 where the relationship between puzzle, movement required and sound feedback was more abstract.

Communicational Dimensions: In AmI systems the environment – in our case primarily comprised of sound – represents the means of accessing and understanding the system’s intelligence, rules and behaviour. This concept arose out of an observation we made that players would often anthropomorphize the system via its sonic response – ‘no, not like this, it doesn’t like that, it gets all quiet’. Sound feedback appeared to be something that users saw as the physical manifestation of the system and relied on it heavily in order to understand how it works and responds to them. In a way it showed a conversational style of experimentation and eliciting response, in order to learn to manipulate, interpret and be in the environment. Again, this concept speaks to the idea of everyday listening [2, 22] – selectively paying attention to minute sound shifts to gain information about the soundmaker’s state, thus building up aural fluency. It brings a different perspective to auditory display design to think of sound feedback as an epistemological occurrence in itself - the content and form in which participants understand and think of the system, which is otherwise ambient, invisible and yet physically persistent. This notion in turn connects the perception of competency that players build up from manipulating the system with their aural fluency, also developed over time. Once again, this helps to account for why even teams who took a very long time to complete a level felt competent in mastering the system’s response and positive towards the sonic feedback.

Confirmatory Dimensions – We observed that having feedback at all times, in this case continuous soundscapes feedback that responds with subtle yet perceivable changes to group actions, reinforces and rewards the efforts of interacting with the system. It provides clear scaffolding for achieving aural fluency. Our case of intensity-based gradient soundscape most resembles a sonification scenario where the response is constant and dynamic, and requires constant attention. The anticipatory reward sound – signaling that users are close to completion served as a soundmark, which users became quickly familiar with, and could instantly identify so as to adjust their actions accordingly. The closer they were to a solution the more time they spent in acute listening attention to the anticipatory reward sound. Upon completion of each game level the auditory reward sound – a signal in the soundscape - seemed to be really enjoyed by the players, and important to their leaving the system confident and satisfied, rather than frustrated and defeated.

6. CONCLUSIONS

Despite the vast proliferation of audio-augmented technologies and spaces, relatively little work has been done to date in designing multi-layered informational auditory displays for responsive environments that actively guide human activity towards solving a problem or achieving a goal [24, 16] While intensity-based sound feedback occurs normally in particular everyday circumstances (e.g. paging your cordless phone at home and going from room to room listening to its sound intensifying), there are few studies of these everyday phenomena, and fewer still of their possible translation into design guidelines for sound feedback. Furthermore, few studies focus on perception of complex everyday changing sound, while taking into account context and purpose of activity, level of embodiment, or familiarity and associations with the sound [24, 25, 26]. A methodological investigation into all these different components of the model is needed in order to understand more fully how to better use and design such auditory displays. Finally, the demands and unique requirements of ambient intelligent systems are still in the process of being heuristically and methodologically developed by the human-computer interaction and adjacent communities, which makes for challenges at this point in time to uncovering stable auditory perception and auditory display design guidelines. What we have presented here is an example of a design-based research project attempting to merge design methodologies with auditory display design investigations. Our analysis starts from the empirical data and opens itself up to emerging research questions
and alternative explanations for our results. From that standpoint, we develop and discuss the notion of aural fluency, as well as build up and describe several general heuristics related to the role of ambient, informative auditory display in ubiquitous computing settings.

7. REFERENCES

Making Sense of Group Interaction in an Ambient Intelligent Environment for Physical Play

Ron Vakkary, Marek Hatala, Ying Jiang, Milena Droumeva, Malahat Hosseini
School of Interactive Arts & Technology
Simon Fraser University
Surrey, BC, Canada
{rwakkary; mhatala; yingj; mvdroume; mhosseini}@sfu.ca

ABSTRACT
This paper presents the results of a study on group interaction with a prototype known as socio-cc(h)o. socio-cc(h)o explores the design of sensing and display, user modeling, and interaction in an embedded interaction system utilizing a game structure. Our study involved the playing of our prototype system by thirty-six (36) participants grouped into teams of four (4). Our aim was to determine heuristics that we could use to further design the interaction and user model approaches for group and embodied interaction systems. We analyzed group interaction and performance based on factors of team cohesion and goal focus. We found that with our system, these factors alone could not explain performance. However, when transitions in the degrees of each factor, i.e. high, medium or low are considered, a clearer picture for performance emerges. The significance of the results is that they describe recognizable factors for positive group interaction.

Author Keywords
Groups, responsive environment, play, embodiment, ambient display, games

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCl): Miscellaneous.

INTRODUCTION
socio-cc(h)o is a prototype environment for group play that explores design and use issues of ambient intelligent systems. Ambient intelligence computing is the embedding of computer technologies and sensors in architectural environments that combined with artificial intelligence, respond to and reason about human actions and behaviors within the environment. Our broader research with socio-cc(h)o is focused on several related issues including sensing and display, user modeling, and interaction models. The overall research goal is to understand how to support groups of participants as they learn to manipulate an ambient intelligent space. The research questions are numerous in a project of this nature and yet immersive and embodied interaction does not lend itself to reducible variables that can be measured independently. Given this, our initial aim is to provide a broad, yet particular set of heuristics that help describe and make sense of group interaction in these environments.

The contribution of this paper is that we describe two factors, cohesion and goal focus as descriptors to analyze group interaction. We further detail how transitions between the degrees of each factor, i.e. high, medium, low help explain the performances of groups in an embodied interaction system. The value of our study is that it reports on responsive environments in which actions are deeply physical, i.e. embodied, visual, aural, and tactile, as well as verbal.

We chose to design our prototype as a game since game structures are suitable for this type of research. Games provide a sufficiently open framework to study a range of interactions that are both embodied and intellectual, yet are sufficiently constrained in regards to goals and rules and so provide commonality and identifiable design factors (i.e. the rules). In socio-cc(h)o, the aim is for a team of four players to progress through multiple game levels. Each level is completed when all the players achieve a certain combination of body movements and positions. At the beginning of each level, players are presented with a word puzzle as a clue in discovering the desired body states. The levels are represented by changes in the environment in light and audio. The levels are progressively more challenging in terms of body states.

In our complete study, we conducted an experiment involving fifty-six (56) participants divided into teams of four. The experiment included a two-hour session of playing in the socio-cc(h)o environment. The teams were divided into two groups that each followed a different
protocol. In this paper we only report on one of the protocols and groups of thirty-six (36) participants. We also have chosen to focus our analysis on data from two critical levels of the game rather than all the levels.

The paper provides an overview of related research, a description of our prototype and study. We then devote the remainder of the paper to discussing the results, analysis of our study and future research issues to pursue.

RELATED RESEARCH

We will briefly describe our previous research related to this project and then present an overview of research into frameworks for tangible and embedded interaction and previous embodied and ambient intelligent prototypes. In [19] we discuss a method for constructing group parameters from individual parameters with real-time motion capture data; and a model for mapping the trajectory of participant’s actions in order to determine an intensity level used to manage the experience flow of the game; as well as design strategies for representing intensity via an audio and visual display. We refer readers who would like further technical details on our prototype to this paper. In [5] we explore the use of an intensity scale for audio display.

Tangible and Embedded Interaction Frameworks

Over the years various frameworks have been proposed to better define tangible user interfaces and embedded interaction. Holmquist and others [8] proposed defining concepts of containers, tools, and tokens. Ullmer and Ishii proposed a framework known as the Token+Constraint System [17] that highlighted the integration of representation and control in tangible user interfaces. Shaer and others have extended Ullmer and Ishii’s work to propose their Token and Constraints (TAC) paradigm [15].

Djajadiningrat argues for a “perceptual-motor-centered” approach to tangible interfaces [4]. He argues for a “direct approach” for its “sensory richness and action-potential” of the objects to carry meaning through interaction.

Homecker and Buur [9] address the interweaving of embodied and social aspects with the interaction experience of tangible interaction in a framework that offers design guidelines along four themes of tangible manipulation, spatial interaction, embodied facilitation, and expressive representation.

While each of these frameworks make a valuable contribution to the epistemological discussions on tangible and embedded computing, none at this point provide an adequate analytical foundation for embedded and embodied group interaction. For example, in [18] we explored the limits of TUI frameworks at the convergence of tangibility and embodiment. While Homecker and Buur’s guidelines are a useful starting point, they are simply too high level and not sufficiently granular for our purposes.

Group Interaction in Game and Play Environments

Recent projects have investigated the play space of responsive environments and tangible computing utilizing sensors, audio, and visual displays. For example, Andersen [1], and Ferris and Bannon [6] engage children in exploratory play and emergent learning through sensor-augmented objects and audio display. Andersen’s work reveals how theatrical settings provide an emotional framework that scaffolds the qualitative experience of the interaction. Ferris and Bannon’s work make clear that a combination of simple feedback and control lead children to widely explore and discover a responsive environment.

In the Nautilus project, Strömberg and her colleagues employ bodily and spatial user interfaces as a way of allowing players to use their natural body movements and to interact with each other in a group game within a virtual game space [16]. Strömberg observed in physical and team games such as soccer or dodge ball that players coordinate their physical movements and rely heavily on communication to be successful.

In relation to the above research, socio-cc(h)o builds on the theatrical, simple and physical interaction models in order to develop a game structure approach that lies between exploratory play and a structured game for adults within an ambient intelligent environment. In addition, we extend the notion of a game structure to an interaction model for the physical environment rather than a virtual game space. We also build on the idea that actions, play and learning are linked in such physical environments.

THE PROTOTYPE & DESIGN MOTIVATIONS

The prototype involves interaction of multiple participants (four at one time) in a cooperative puzzle game that is solved by coordinated physical actions of the group. The environment is responsive to the participants’ actions through ambient audio and light (see Figure 1).

Description of the prototype

The short scenario of the socio-cc(h)o environment below is excerpted from a previous paper [19]:

Madison, Corey, Elias and Trevor have just completed the first level of socio-cc(h)o. They discovered that each of them had to be low to the ground, still, practically on all fours. Once they had done that, the space became bathed in warm yellow light and filled with a wellspring sound of resonating cymbals. Minutes earlier, the space was very dim — almost pitch black until their eyes adjusted. A quiet soundscape of "electronic crickets" enveloped them. They discussed and tried out many possibilities to solving the word puzzle: "Opposites: Lo and behold." At Corey's urging, the four grouped together on the edge of the space and systematically sent a player at a time to the opposite side in order to gauge any change in the environment.
Nothing changed. Madison, without communicating to anyone realized the obvious clue of “Lo” or “low”. She lowered herself to a crouching position. The space immediately glowed red and became brighter. The audio changed into a rising chorus of cymbals – not loud but progressively more pronounced. Corey and Trevor stopped talking and looked around at the changing space. Madison, after a pause began to say “Get down! Get down!” Elias stopped down immediately and the space became even brighter. Corey and Trevor dropped down in unison and the space soon became bathed in a warm yellow light like daylight. The audio reverberated in the space. A loud cheer of recognition came from the group. “Aaaaahhh! We got it!”

The game consists of six levels that require the group to achieve specific body states and goals. The body states are the body movements and positions that players must discover in order to complete a level. Goals are the changes in the environment players are aiming to achieve. The goals are implicit and have to be discovered during the game. Each level has a beginning quality of light and audio. As the players progress toward achieving the right body state, the environment incrementally shifts toward the goal state of the environment (see Figure 1). For example, as depicted in the scenario, when Madison lowered herself, the environment gradually shifted toward the goal of creating day. As each of the other three players followed Madison, the environment responded to movements of each player.

The levels were designed to enable the players to gradually acquire generic skills to manipulate the environment. The aim was that a generic skill acquired at lower level is required in order to discover the more complex body states at higher levels. Goals allowed us to design an implied progressive narrative. We intended for this approach to provide a sense of coherency across the levels, and to loosely map increased challenge to the reward of a more complex display.

The physical environment consists of a circumscribed circular space (the area in which we can detect motion), immersive 8-channel audio, theatrical lighting, and three video projection surfaces.

**Design Motivations**

In approaching the design of our ambient intelligent environment we were clearly inspired by play and games. We followed the principles of simple rules and goals with a great degree of variation in how the goals are achieved. For example, consider the game of soccer, in which within a set of simple rules there are a myriad of ways to score a goal. Fans of soccer often refer to the creative playmaking in a wonderful goal. In fact, it is the very combination of limited simple rules and almost limitless paths to the same goal that establishes a creative space that can make a game so appealing. From the perspective of the design of the interaction model and system, we realized it was important to decide where not to specify interaction and system functionality. In many respects, we learned to off-load formalized interaction among participants to the situated dynamics of people mundanely interacting, i.e. people will communicate together in whatever form possible given the resources in the environment without the need of formalizing communication modes. In addition, the system does not need to encode or sense actions or behaviors that are not relevant to the required body-state at a given level, i.e. no response from the system is a perceived response. We employed simplicity with the aim of creating rich and complex responses. For example, limiting sensing actions to whole body positions and movements rather than gestures, opened gestures up to unique, particular and complex communication between participants. Ignoring or not encoding large parts of the embodied action supported a wide range of exploration of body movements. In other words, our approach was not to overwhelm the interaction by over determining the experience.

At the same time we are looking for clear descriptors and factors upon which we can base refinements and future design, and validation of our design strategy.

**THE STUDY AND RESULTS**

**Description of the study**

In our complete study, each of the participants completed Keirsey's Temperament Sorter\(^1\) to identify his or her personality type. We utilized the results to organize people into team configurations that resulted in 56 participants being chosen from which a cohort of 36 participants is reported on in this paper. The other cohort of 20 participants followed a different protocol aimed at studying the explicit support of players during game play (these results have yet to be published). The sample of 36 included 11 females and 25 males where 30 were 18-24 years old, 5 were 25-34, and 1 was 35-49. This produced 9 teams of 4 players.

The Keirsey temperament designations were pursued in order to develop a baseline metric for our group user model. As it turned out we found virtually no significance between

---

\(^1\) See http://www.keirsey.com
our results and temperaments. We discuss this in detail in another paper currently in review that cautions against the use of Keirsey temperaments as a compositional approach to user models or group interaction. In this paper, we focus on an alternate explanation of observed commonalities and as such make no further reference to Keirsey temperaments.

Each session began with a warm-up that introduced the concept of puzzles solved through physical action that was helped by implicit responses. The warm-up was a modification of the child’s game of “hot-cold.” Participants were also played a range of sonic cues and rewards in order to adjust their perceptual hearing to our sound ecologies. Each team of four played the first four levels without any intervention from the research team. After a short break the last two levels were played. A time limit of 15 (fifteen) minutes was given for completing these last levels.

The evaluation was performed with the socio-ec(h)o prototype in our ‘black-box’ lab environment. The sessions were videotaped and audio recorded. In addition, each participant wore a wireless microphone to record conversations. Following the sessions each participant completed a questionnaire.

**Data Types**

The data analyzed included video coding of the sessions that used a scheme that we developed based on levels of cohesion and goal focus, log data from the system that mapped the game state and player actions represented as intensity, and transcripts of the verbal communications between players. We also compared our groupings against the questionnaire data.

**Coding Scheme and Transcripts**

Each session was recorded with three cameras providing sufficient data to code group actions and behaviors. Our coding scheme was based on two main factors: cohesion and goal focus. The combination of these factors in a two-dimensional matrix show the degree of descriptive capacity, see figure 2. Two researchers independently coded the videos and later negotiated the differences to reach a consensus.

**Cohesion** can be described as the extent to which players appear to be acting as a team (all members coordinating together) – whether that is working on a game solution, playing, thinking, or talking to each other. Cohesiveness is a measure of team dynamics and does not necessarily reflect their focus on the game but only whether they are acting in unison as a team. We analyzed different degrees of cohesion:

- **Low**, players are not together as a group or they are temporarily fragmented. They are not communicating or are individually exploring;
- **Medium**, players are in the process of becoming a group or are regrouping. Players are negotiating roles and establishing leadership or consensus;
- **High**, players constitute an established team. They make several agreements and are coordinated in their movements or are communicating with each other about strategy and solving the puzzles.

Cohesion is a common construct in group studies, referring to the affinity an individual feels toward a group [2] and is demonstrated by a high commitment to the solidarity of the group [12]. Cohesion is often studied in an organizational context dating back to Seashore [14], and more recently it has been applied to virtual teams [13].

**Goal focus** can be described as the extent to which players appear to be or are attempting to “play the game” the way they understand it. Game activity is not dependent on whether players are working as a team or not. In addition, game activity does not necessarily only mean that players are actively playing, i.e. in our case moving. If players are still because they believe the game requires them to be still, then they are “playing the game.” We analyzed different degrees of goal focus:

- **Low**, players are not involved in playing the game. They are resting, or are distracted, or engaged in activities not related to the game;
- **Medium**, players are in the process playing the game. They are experimenting with different actions, and communicating with each other about or reflecting on the effects of their actions;
- **High**, players are actively and consciously playing the game and attempting to solve the puzzle at hand. This is reflected in concerted efforts and good communication related to their performance in the game. Many ideas are shared on actions for solving the puzzle.

Malone and Lepper [11] consider games to be intrinsic motivators for learning that suggests an internal state of mind however a sense of focus or concentration is evident.

![Figure 2 Matrix showing the descriptive capacity of the two factors cohesion and goal focus](attachment:image.png)
Huizinga referred to play as invoking a magic circle, a liminal space for games [10] that extends the perceptible concentration to a sense of intense focus that is almost separate from the everyday world. This idea has been captured by psychologist Csikszentmihalyi's notion of flow, which is a high level of engagement, risk and challenge found in play and ritualized in sport [3].

**Intensity Scale**

A component of our system is a reasoning engine that manages the game state of the interaction. The model for this includes mapping the trajectory of the body states to participant's actions in order to determine the intensity level, or proximity to the desired body state. For example, the sensing system sends data on predefined parameters such as velocity and body positions. Based on these basic parameters, the reasoning engine infers higher-level behaviors for the user group such as high-fasting group, middle-low-stationary group, etc. The intensity is computed based on heuristics applied in response to the current state of the game and the inferred behaviors of the teams.

The intensity is measured from 0 to 4 with 4 representing the maximal intensity or state completion. The intensity is reflected in the ambient display of the system, thus when the intensity reaches value 4, the participant hears the cue sound indicating that they achieved the body states required by the current level. In addition, we felt that the overall shifts in intensities toward and away from the goal could be represented in the ambient display in a gradient effect in real-time in order to support player actions in the environment and awareness of the game state. It should be noted that the ability to sustain intensity levels are also monitored, typically a 4 second duration is required to complete the state, see figure 3. The reasoning engine provided a log of intensity values. This allowed us to precisely analyze the game state (players’ actions) and the state of the environment (ambient display).

**Questionnaire**

The questionnaire was comprised of twenty-five questions including Likert scale and semi-structured answers. The questionnaire was structured in four parts with questions on the perceptions how the system facilitated or constrained goals; the social and physical resources of the players, system and environment; the role of internal and external elements in the learning and cognitive transformations; and how the activity and understanding of the system changed over the course of the two hours. Participants completed the questionnaire immediately after the sessions.

**Analysis**

Our analysis looked at the different levels of cohesiveness and goal focus over duration to determine a density value in percentages:

$$\text{density} = \frac{\text{factor}(\text{min})}{\text{duration}(\text{min})}$$

We looked for combinations of density values of the different degrees (high, medium, low) of the two factors (cohesion, goal focus) and compared these to team performance or duration of the game level. Additionally, we correlated the different degrees of cohesion and goal focus factors with team performance (duration) using the Pearson correlation coefficients. The Pearson correlation coefficients measure the degree and the direction of the linear relation between two variables. That is, how much changes in one of the variables related to changes in the other variables. Correlation can be used to estimate the extent to which teams’ performance, cohesion and goal focus factors were related. Lastly, we compared the video coding results with the intensity data from the logs, see figure 4. Based on these comparisons we can isolate key events that we can further examine through transcripts and videos.

**Results**

Our results discuss correlations between high degrees of the two factors, the role of transitions, and players' perceptions. Tables 1 through 4 show correlations between cohesion and goal focus. Note that in each table, column numbers refer to the same values as rows, for example in table 1, row 7 and column 2 shows a significant correlation of .871 between the medium degree of cohesion and completion time.

One might expect that a team that showed high density values of both cohesion and goal focus factors would lead to a fast performance in the game. Indeed, we found that Team H held density values of 93% for goal focus and 97% for cohesion in level 4 and completed the level in less than a minute. However, Team D had significantly more modest density values for level 4, 63% for goal focus and 67% for cohesion yet were able to complete the level in just under a
High level of cohesion and goal focus does not necessarily lead to good performance:

Statistically, we found no significant correlations between high degrees of cohesion or goal focus factors and team performance in game level 3 (see Table 1). We had virtually the same results for game level 4 see Table 2. However, Table 1 shows a significant correlation between Medium degree of cohesion and performance (.871). Table 2 shows a strong correlation between Medium degrees of cohesion and goal focus and performance (.892; .927). This led us to examine the role of transitions, where factors change in degrees such as a team shifts from a high degree of cohesion to medium degree of cohesion.

Transitions as an influencing factor

We found that transitions from different levels of coherence and goal focus held statistical significance when compared against performance throughout level 3 of the game except for transitions to high cohesion, see row 9 in Table 3, and significance in transitions from all degrees of both factors in level 4 of the game except for transitions to low cohesion, see row 9 in Table 4.

Players Perceptions

We examined the relationship between players' perception of the helpfulness of the system and their performance. No correlation was found, thus fast players did not necessarily believe the system to be more helpful than slow players. However, there was a significant relationship between players on teams who completed the most levels and their perception of the support of the system. This suggests that the more "skilled" players (those who could complete the higher levels) perceived the system to be more helpful. The overall rating of the system was quite good, for example on the question of how helpful the system was the median score was 4.0 (SD 1.02) on a scale of 1-5 (5 high).
DISCUSSION

We expected a clearer pattern of what makes players successful and where breakdowns occur. For example, we expected correlations between high degrees of cohesion and goal focus. In our analysis, the results point to three possible patterns of engagement with our system yet only one pattern that leads to better game performance:

- **Get it right:** Certain teams were simply very good. They demonstrated an understanding of the game that the players were able to carry across levels. For example, Team H was consistently quick and their initial ideas for what might work in the game tended to be correct.

- **Luck:** Some teams were simply lucky in certain levels. The players were not engaged yet accidentally they formed a correct body state leading them toward a solution however they may not know how they got there. For example, in some cases the degree of cohesion and goal focus was modest yet the intensity was high.

- **Banging their heads against the wall:** Some teams displayed extraordinary persistence by consistently having high degrees of cohesion and goal focus yet equally consistent in their inability to find a correct path for a solution.

Enabling skilled players is a part of any game as is luck, yet the third pattern is a distinct problem with the prototype. One possibility is that we designed an approach that exclusively looks for positive patterns to support, i.e. as players find a correct path the system provides positive support by increasing the intensity of the ambient display. A design change would be to include a negative intensity response in the ambient display that would warn players they are going in the wrong direction.

Transitions were a better indicator of performance, at least in terms of speed. In other words, teams that made fewer transitions, i.e. shifting between different degrees of either cohesion or goal focus, completed levels faster. This is clearly illustrated in figure 6 showing the affect of transitions on duration. Here teams that take longer have a greater number of transitions, for example the marginally longer time of Team C in comparison to Team B can be
accounted for by the slightly more transitions in degrees of cohesion.

In our preliminary examination, we found that transitions allow players to strategize, analyze the system’s response, and to communicate, in that order. It is clear that fewer transitions may help a team to perform faster but that does not mean that no transitions are an optimal pattern.

Our future research will focus on a detailed analysis of the video and transcripts to better understand what occurs during the transitions and how better to support these activities in terms of design. This next phase of research will rely on analyzing the rich qualitative data we have, based on the questionnaires, video, and audio. Further, we expect the embodiment aspect of our prototype to yield a multi-modal description of what occurs during transitions that are particular to tangible and embedded systems. While these preliminary findings do not fully make the case of generalization, we do believe that our future findings will have general applicability to ambient intelligent systems.

We analyzed only levels 3 and 4 of the six (6) levels of the game. The amount of data across all levels is substantial. We aimed to find patterns in these two critical levels first. Levels 1 and 2 are in large part learning levels. Nevertheless, our future work will look at the full cycle of the game with an eye toward the cumulative patterns such as skill acquisition and team dynamics. For example, a pattern of more effective strategizing and less analysis of the system’s response as the levels progress suggest players clearly learn as they play.

CONCLUSION

We have provided an overview of related research, described our prototype, and experimental study. We concluded by discussing the results, analysis of our study and future research issues to pursue. The contribution of this paper is that we describe two factors, cohesion and goal focus as descriptors for analysis of group interaction. We detail how transitions between the degrees of each factor help explain the performance of groups in an embodied interaction system.

ACKNOWLEDGMENTS

This project was supported by a grant from Canadian Heritage, New Media Research Networks Fund.

REFERENCES

Abstract

In this paper, we analyze the current state of museum guide technologies and applications in order to develop an analytical foundation for our future research in an adaptive museum guide for families. We have focused our analysis on three critical areas of interest in considering group and social interaction in museums: tangibility, the role of tangible user interfaces; interaction, visit types and visit flows; adaptivity, user modeling approaches.

Keywords: museum guides, adaptive museum guides, tangible user interfaces, user models, group interaction, mobile, family

Introduction

Museum technologies have increasingly become the focus of research in such areas as ubiquitous computing, tangible computing and user modeling. Since the advent of audio-based museum guides, much research and development has been placed on increasing the technological capacity of augmenting the museum visit experience. Early prototypes such as Bederson (Bederson 1995) provided evidence that it is possible to support visitor-driven interaction through wireless communication, thus allowing visitors to explore the museum environment at their own pace. Current prototypes and fully functional systems are much more complex, supporting a variety of media, adaptive models, and interaction modalities. However, as Bell notes in her article on museums as cultural ecologies; “The challenge here is to design information technologies that help make new connections for museum visitors” (Bell 2002). Bell further argues for the importance of sociality in museums where visits are as equally social and entertaining as they are educational.

At this stage in our research in adaptive museum guides we are exploring how best to address issues of social engagement and play and learning for groups of visitors such as families. Families are by far the most common visitor type to science, history, and natural history museums. In this paper, we describe our theoretical analysis of the current state of museum technologies and adaptive museum applications in order to provide a detailed understanding in support of our current research goals. We’ve narrowed our focus to three critical trends in adaptive museum guides that we believe will form the theoretical anchors of our future research. The areas include: tangibility, interactivity, and adaptivity. We discuss these trends in order and follow with a discussion on the impacts on our research. For the reader, we believe this paper provides a critical summary of electronic museum guides of the last decade and outlines key areas of concerns for academic researchers.
Situating Tangibility In Museums

Tangible user interfaces (TUI) imbue physical artifacts with computational abilities. Like no other user interface concept it is inherently playful, imaginative and even poetic. In our previous research in adaptive museum guides we developed a prototype system known as ec(h)o. In this project, the simple physical display devices and wooden puzzles at the natural history museum where we conducted ethnography sessions inspired us. As a result, the user interface for ec(h)o was a TUI that coupled a wooden cube with digital navigation and information. In ec(h)o, museum visitors held a light wooden cube and were immersed in a soundscape of natural sounds and information on the artifacts on display (figure 1). Visitors navigated the audio options presented to them by rotating the cube in their palm in a direction that corresponded to the spatial location of the audio they were hearing (we will return to this project below but for a detailed account see (Wakkary and Hatala 2007)).

![Image](image.jpg)

Fig 1: the tangible-user interface of ec(h)o

In our experience with TUIs, they bridge between the virtuality of the museum guide system and the physical surroundings of the exhibition. As such, the adaptive museum guide firstly becomes more integral to the physical ecology of the exhibition including artifacts, display systems, and architecture rather than being a separate technology. It is important to note that our social interactions are in large part mediated by objects and environments as much as by direct contact with others (Kaptelinin, Nardi et al. 2006). Awareness of context is critical to sociality. Secondly, with TUIs our understanding of interacting with technology is informed by our rich experience with physical artifacts and surroundings – since we can leverage existing embodied and cognitive understanding. Analytically, Kenneth P. Fishkin (Fishkin 2004) provides a taxonomy that allowed us to explore these two factors further.

Fishkin’s taxonomy is a two-dimensional space across the axes of embodiment and metaphor. Embodiment characterizes the degree to which “the state of computation” is perceived to be in or near the tangible object. As we discussed, it expresses the integration of computation with the physical artifact and environment. Fishkin details four levels of embodiment:

1. **distant** – representing the computer effect is distant to the tangible object;
2. **environmental** – representing the computer effect is in the environment surrounding the user;
3. **nearby** – representing the computer effect as being proximate to the object;
4. **full** – representing the computer effect is within the object.

Along the second axis, Fishkin uses metaphor to depict the degree to which the system response to user's action is analogous to the real-world response of similar actions – the existing embodied and cognitive mappings. Fishkin divides metaphor into *noun metaphors*, referring to the shape of the object, and *verb metaphors*, referring to the motion of an object. Metaphor has five levels:

1. **none** – representing an abstract relation between the device and response;
2. **noun** – representing morphological likeness to a real-world response;
3. **verb** – representing an analogous action to a real-world response;
4. **noun + verb** – representing the combination of the two previous levels;
5. **full** – representing an intrinsic connection between real-world response and the object which requires no metaphorical relationship.

In figure 2, we have applied Fishkin’s taxonomy to ec(h)o. *Embodiment would* be considered “environmental” since the computational state would be perceived as surrounding the visitor given the spatialized audio display output. In regard to *metaphor*, the ec(h)o TUI would be a “noun and verb” since the wooden cube is reminiscent of the wooden building blocks and the motion of the cube determines the spatiality of the audio, as turning left, in the real-world, would allow the person to hear on the left. If we consider the visitor’s movement, the embodiment factor would still be environmental and we’d have to consider the visitor’s body as being “full” in Fishkin’s use of metaphor. And so in representing the entire system, we plotted ec(h)o between “noun + verb” and “full” on the metaphor axis.

In summary, we believe tangibility is critical in considering the social dimensions of museum visits in museum guides. TUIs can integrate the computational and physical affordances of the museum visit experience, and visitors can leverage embodied and cognitive models of interaction in incorporating new technology. Visitors benefit from augmentation of computation yet maintain awareness and embodied connection to their environment.
surroundings, which ultimately supports social interaction.

Situating Interactivity In Museums

Interactivity can be understood in many ways. We focused on individual or group experiences in museum visits and models of content delivery. In this section, we review a range of systems from the last decade with a particular eye on how these past approaches provide insights into designing for families in which sociality and group activity are paramount. In reviewing the various systems we have developed a matrix that compares the visit type (individual/group) to visit flow (information delivery/game-interaction) which aims to uncover the major factors that affected the design of the various systems.

In visit types, we refer to individual as systems that target interaction with single individuals, whereas group refers to interaction models aimed at groups of individuals. Information delivery refers to the approach of an information corpus or repository that is presented to the visitor whether adaptively or by user selection. Providing the visitor with the ability to interact with the exhibit content in a playful manner we've referred to as game-interaction. Often the game-interaction approach will use games as a means to educate the visitor, as opposed to providing information about an artifact. By employing a cross matrix of both types of categorizations, it becomes possible to understand how both past and current museum guides mediate the museum experience.

Many of the systems we reviewed were information delivery in style, while also falling under the individual category. Typically they involved a Personal Digital Assistant (PDA) and an audio/visual interface such as the Blanton iTour (Manning and Sims 2004). Among the first of such guides was the HyperAudio project (Petrilli and Not 2005), which used an adaptive model that we will discuss again in the next section. In the HyperAudio system, individual visitors were encouraged to walk about the exhibition with the handheld device and headphones, stopping at various exhibits to learn more about specific artifacts in their surrounding. As the authors, Petrilli and Not note,

the presentation (audio message and hypermedia page) would be adapted to each individual user, taking into account not only their interaction with the system, but also the broad interaction context, including the physical space, the visit so far, the interaction history, and the presented narration (Petrilli and Not 2005).

The presentation that is displayed also provides a link that the participant can click on with a palmtop pen to gather further information about the artifact of interest. By doing so, the visitor is provided a map of the museum on which the location of the new artifact is displayed, to allow the visitor to see the artifact in person if he desires. The authors also describe their interest in keeping the graphical user interface to a minimum so as not to distract the visitor. It is for this reason that audio is chosen as the main channel of information delivery.

The PEACH project (Stock, Zancanaro et al. 2007) provides the visitor with a digital character on their PDA that provides information on various artifacts within the exhibition. The PEACH guide also contains rich media in the form of video close-ups of frescoes and detailed descriptions of paintings. Unique to this system is the availability of a printout, which provides the visitor with an overview of the exhibits he encountered while at the museum.

We’ve observed in a commercial system developed by a member of our team as part of Ubiquity Interactive, known as VUEguide that rich media images can at times become more engaging than the authentic object on display. We believe the virtual image can be easily integrated into a narrative world of information delivery. For example, interactive images of Bill Reid’s ‘Raven and the First Men’ sculpture engaged visitors deeply with the screen and encouraged a ‘back and forth’ exploration between the actual and the virtual (see figure 3).
Berkovich et al (Berkovich, Date et al. 2003) developed the Discovery Point prototype, a four-button device that delivers audio to the museum visitor. The device allows the visitor to listen to a short audio clips about artifacts that are controlled by the visitor. The visitor can also create a virtual souvenir by pressing the "MailHome" button, which adds the artifact in question to their personal Web site created upon entering the museum. The device functions without the use of headphones. Instead, audio is delivered through specialized speakers which direct audio to a precise area around the artifact, so as not to disturb other museum visitors.

The Sotto Voce system, produced by Aoki et al (Aoki, Grinter et al. 2002), provided the first instance where researchers focused on group activity. The system contained an audio sharing application called eavesdropping that allowed paired visitors to share audio information with each other while on an information delivery tour. In designing the application with three entities in mind (the information source, the visitor's companion and the museum space), the authors remarked that the system showed that the visitors used the system in creative ways, and with social purpose, while presenting the opportunity for co-present interaction.

In recent years, research has continued on group-based museum tours as an area of interest, but orientation has shifted from information delivery tours to game-interaction activities. This is evident in the CoCicero project implemented in the Marble Museum in Carrera, Italy (Laurillau and Paternò 2004). The CoCicero prototype focused on four types of group activities; (i) shared listening – similar to the Sotto Voce system, (ii) independent use – to allow individuals to choose not to engage in group-based tours, (iii) following – to allow an individual to lead other members of a group, and (iv) checking in – which allows members in a group to know how others are doing through voice communication while not being physically present. The authors state that communication, localization, orientation and mutual observation are four elements that are key to a collaborative visit. The guide functions by providing museum groups a series of games, such as a puzzle and multiple choice questions, which require the visitors to gather clues through viewing the exhibits within the museum.

Similarly, the ARCHIE Project (Loon, Gabriëls et al. 2007) has developed a learning game for school children that allows visitors to trade museum-specific information to gain points in order to win a game. With each player having a specific role that he plays, the visitor must understand various levels of information gained from exploring the museum in order to improve his score. Though the ARCHIE project is only in its prototype stage, it is clear from their initial findings that the game-interaction approach does foster user interest in museum content. Our own system, ec(h)o, is unique in relation to other systems. It is exclusively an individual
type system (which we now see as a significant drawback) yet it used a game-interaction approach. Our interface aimed for an interaction based on an open-ended game qualities or what we referred to as play. Play took on two forms: (1) content play in the delivery of information in the form of puns and riddles; (2) physical play that consisted of holding, touching and moving through a space; simple playful action along the lines of toying with a wooden cube.

Fig 4: Cross matrix of visit flow and visit type

In summary, it is clear that the majority of research has been in the area of information delivery and individual visit type. See figure 4 where we plotted the different systems we reviewed. Sotto Voce is among the first to explore co-visiting or group interaction with a museum guide within an information delivery context. Group interaction is a new area and chronologically represents a trend, especially when combined with a game-interaction approach.

Situating Adaptivity In Museums

This section surveys those projects that deployed user modeling. User modeling is the use of artificial intelligence software techniques to construct conceptual models for users that enable predictions and responsiveness to user interactions. The projects we review include: Hippie, the Museum Wearable, ec(h)o, and HyperAudio. We have also analyzed a system that is currently Web-based, known as CHIP. Each system offers the museum visitor some form of personalized content such as audio or video clips, text, and images, and each uses aspects of visitor interaction with the system to achieve customization. The details of each project are discussed first, where we provided a detailed account of each project’s approach to adaptivity. The overview will be followed by an analytical discussion that includes a cross matrix used to frame the current research in the field.

Adaptive Systems

Earlier we introduced the HyperAudio system from 1999 (Petrelli and Not 2005). Audio content generation was based on a visit model that looked at the group each user was in, whether it was a first or repeat visit, how long they planned to stay and what kind of interaction type they preferred. Underlying the model was a simple boolean variable indicating interest in each item, based on whether the user remained in front of an object after the relevant presentation finished, or whether they turned the presentation off.

Another early attempt to develop a context-sensitive, adaptive museum guide, was the Hippie project
Hippie contained models for 3 distinct components: 1) a static domain model (objects to be presented and processed about), 2) a static space model (physical space), and 3) a dynamic user model. In constructing the user model, the authors assumed that visitors had a "stable interest trait structure", but that environmental and contextual factors played a role in the actual activation of that structure at any given moment. The goal of their user model was "to predict the information needs of a user in a given episode of a visit," thus the model made inferences regarding the next exhibit to visit and the next piece of information to present. Interest in particular exhibits was inferred by time spent there as well as the number of information items selected. They also suggest using "differentiated navigation behavior" to indicate interest, i.e. viewing artworks from multiple locations rather than one indicates more interest. Guide content was classified according to type, and used to infer what kind of information the visitor was interested in. The HIPS project also applied Veron & Levasseur's 1999 ethnography work in museums, which identified four physical movement patterns of visitors: ant, butterfly, fish, grasshopper (Marti, Rizzo et al. 1999).

- Ants proceed methodically through the entire museum space, looking at everything;
- Butterflies fly around from work to work based on interest level;
- Fish swim through the space quickly, glancing at things in a cursory fashion;
- Grasshoppers also bounce around, but with less of a defined idea of what they are interested in.

Visitor type and information preference was proposed as a way to assemble appropriate length audio clips for each individual.

While the Hippie system asked the users what areas they wish to explore, other approaches did not give this level of control to the user, such as the Museum Wearable, which assembles and delivers personalized content to the user without explicit user interaction (Sparacino 2002a) and (Sparacino 2002b). A Bayesian network is used for user classification & decisions about content assembly. Sparacino uses a ternary classification of visitors as busy (cruise through everything quickly), selective (skip some things, spend long on others), greedy (spend a long time on everything). Data was collected and used to set prior probabilities of a model based on path and pause duration patterns. Upon approaching a new exhibit, a sequence of content clips was dynamically assembled using user type to determine how long it should be.

Th ec(h)o project provided visitors with a choice of audio clips throughout their museum visit (Hatala and Wakkary 2005). The ec(h)o user model captured two basic kinds of information about the museum visitor using the guide; interaction history and user behaviour information. Interaction history recorded movement through the museum space along with the sound objects selected for listening, whereas user behaviour information kept track of the user's interests. Interest was set by the user's explicit choice from a number of different concepts at the beginning of the museum tour and then updated based on the exhibits they paused at during their exploration of the museum space and the choices they made regarding which audio content to listen to. If the user showed special interest in a specific concept or two, those concepts would become highly weighted in the model and content related to them would be offered frequently. However, to avoid the complete stereotyping of a visitor, a "variety" element was also included in the recommendation algorithm so there was the possibility of sound objects being offered that would allow the user to move away from their heavily weighted interests.

The most recent project to use adaptability in museums is the CHIP project, which is currently Web-based only. This project allows a user to generate a personal profile via an online rating system for artwork (Aroyo 2007). Images in the system are semantically annotated for artist, period, style, visual content and themes. As users rate them with 1-5 stars, the user model is refined to better reflect their preferences. The classification ratings (e.g. 4 stars for Impressionism) are inferred from the explicit work ratings, and can be viewed at any time. If the user disagrees with the system's calculation of the classification rating, they can adjust it manually. The user can also ask to see "recommended works" which pulls out unrated artwork that the model believes the user would rate highly (4-5 stars). Users can query as to why the system recommended this work, and will be provided with the individual classification ratings that underlie the suggestion. The next step in the project is to integrate this with technology in the physical museum, so that users can access their profile in the exhibit.

The PEACH project also allows the user to actively participate in the construction of her user-model through a
widget (Stock, Zancanaro et al. 2007). The “like-o-meter” widget allows the visitor of the system to state whether she likes or dislikes a given museum artifact, which affects the amount of information the system provides for subsequent related artifacts. The authors suggest that their widget was clearly understood and enjoyed by visitors.

Analysis

The goal of this analysis is to get a sense of what is common practice in user modeling, what is more uncommon and experimental, and what has not been attempted yet. One of the first aspects to consider is what kind of data the systems gather and use as input to the model. There are two basic categories of input commonly used: physical data and content-based data. Physical data includes things like knowledge of the visitor’s current location (Hippie, Museum Wearable, ec(h)o, and HyperAudio), of their overall path through the museum space (Hippie, Museum Wearable), and of their stop duration at each specific exhibit (Hippie, Museum Wearable, HyperAudio). Content-based information includes knowledge of what content the visitors have listened to or selected (Hippie, Museum Wearable, ec(h)o, CHIP, HyperAudio) and how they rate that content (CHIP, PEACH). From this basic input, all of the systems extrapolate and make inferences about categories the user falls into or characteristics the user might have, based on the observed behavior. Every system infers user interest in an artifact/exhibit, usually based on their movement towards it, presence near it, or selection of it in some way. The systems all have the capacity to detect user interest in broader themes or topics, as signaled by the selection of multiple items, which have been annotated similarly. Some of the systems also include the ability to detect interest in a specific information type, as signaled by selection of specific kinds of content (the Museum Wearable, Hippie, and HyperAudio). For example, a user might repeatedly ask for artist biography content, or for details on how an artifact was constructed. The final type of inference found in these systems is the classification of visitors as a certain “user type”. This includes the fish/ant/butterfly/grasshopper path classification from the Hippie project and the greedy/busy/selective user type from the Museum Wearable, both of which used movement patterns to sort users.

Content and Interaction type matrix

One important distinction amongst the ways the projects handle system input is the degree of control that the user has over the cues that the system is picking up on. Another way to think about this would be in terms of the opacity level present in the interaction. With some of the systems, it is easy for users to tell what kinds of information the system is picking up on, such as their presence next to an object or their deliberate selections within the system; we describe this approach as transparent interaction. Other systems have much more opaque interactions that are being picked up on, such as the path through the space or their patterns of stop duration and movement. In these cases, the user of the system may never guess that such elements are being used as input for the user model, and thus it is beyond their control to affect what the system is doing in response to them; we describe this approach as opaque interaction. It’s possible that this opacity will yield a more natural, intuitive interaction with the system, but it’s equally possible that it could result in a frustrating experience where the user does not understand why the system is responding the way it is.

With the collected and inferred data as input, the next concern is what the system presents back to the visitor as output. Several of the guides offer a set of further audio and/or video content that the user can select from (Hippie, ec(h)o). The Museum Wearable also presents audio/video content, but does not allow the user to select what they view; instead the model automatically assembles a tailored presentation for the user. HyperAudio combines these approaches by assembling a tailor made audio presentation and also allowing for individual exploration of topics via the handheld device. A third option for output is to generate recommendations of other pieces the visitor might like based on the model’s understanding of the visitor (CHIP). There is a basic distinction here between systems that have static content and those that have personalized content. The Museum Wearable and HyperAudio actually tailor the content itself to the individual user. In the other systems, the content remains the same, but the order in which it is presented or the options available at any moment are tailored to the individual user.
In summary, Figure 5 shows a cross-matrix of this static-personalized content dimension with the opaque-transparent interaction dimension. As can be seen from the diagram, the majority of research thus far with user modeling in museum applications has involved static content and transparent interaction. There have been a couple of different ways in which the input/inference data affects the output. All the guides have the simple correlation of high user interest leading to more content related to that user interest. Hippie and the Museum Wearable also use user type classifications to affect the duration of the content offered to the user. Room still exists for a range of creative thinking in this area, especially with regards to how collected data can influence the system interaction and output.

**Discussion**

Through gathering research in this study, it is possible to understand the variety of approaches taken to augment the museum space. When faced with group-based activities, the PDA may further distract the visitor from his/her companions. The group-based applications discussed in this study revealed a shift toward a game-interaction approach – while all used PDA-like devices to guide the visitors through the museum, we believe TUIs may be even more effective in this regard. Research within group-based tour applications is fairly limited, and has become the focus of two newer studies (CoCicero and ARCHIE). The game-interaction approach seems not only to affect the way in which individuals browse through the museum, but may also affect the manner in which learning occurs. The game-interaction approaches discussed here encourage the visitors to find artifacts, which match specific criteria in a quest-like fashion. This type of activity might be useful in teaching visitors what singular artifacts are, but may inhibit the communication of the contextual significance that often surrounds artifacts in museums. The earliest of systems to adopt a group-based approach, Sotto Voce, note that co-present interactions should be supported. The CoCicero project takes the concept further by introducing other group-based behaviours to be supported, such as following and checking-in. As our focus is on group-based interaction, using the findings of such projects may prove helpful in designing an application that fosters group visits by families to museums.

With regards to a user modeling approach, the majority of the research lies within supporting transparent interaction, and static content, though it is difficult to state whether one method is superior to another at this time. It could be argued that a push towards even more transparent input is desirable. The CHIP project is the most transparent of all the interactions, allowing the visitor to adjust the user model manually and get feedback from the system as to why certain recommendations have been made for them. No such hybrid adaptive and adaptable system has been implemented within the user space itself. Alternatively, it could be fruitful to
explore the area on the extreme other side of the transparent-opaque interaction continuum, by creating a system where users are unable or unlikely to guess what aspects of their behavior are being used within the system. If done properly, this could yield a very immersive, natural feeling interaction. Projects that have attempted to classify visiting patterns have done so within the individual-based tour context, and there is a lack of research on museum visitor patterns for groups – a potential area of interest for us to address.

There is also little research on adapting to the group conditions of the visitor, e.g., taking into consideration the fact that person A is a mother visiting with three children while person B is a senior citizen visiting with friends. Modeling users on the level of group dynamics is a new and growing area within user modeling, one that we have explored (Wakkary, Hatala et al. 2005) and hope to bring to the museum domain. In terms of content, there is definitely space to explore in the realm of adaptive content, taking the personalization of the museum beyond simple recommendations and on-demand information access and moving it into the realm of individualized presentations.

We began the paper with a discussion on tangibility as a mediating bridge between computation and context that militates the lack of support for social interaction. Tangibility may preclude some of the directions in user modeling such as transparent interaction of systems like CHIP, which is Web-based after all. However, tangibility could be seen as an enabler in game-interaction approaches and group interaction, given its relative playfulness in comparison to graphical interfaces and the ease of sharing tangible devices and the affordance of leveraging past patterns of social interactions mediated by objects.

**Conclusion**

As shown in this paper, there are currently several approaches for adaptive museum guides under exploration. Through analyzing current contributions through the three key areas of interest, tangibility, interactivity, adaptivity, we were able to develop a comprehensive outline that provides a technical grounding for our future research in a family and social based museum guide. The majority of the current literature focuses on the interactions of a single visitor but through analyzing trends in the museum guide research, the focus appears to be shifting.

**References**


Berkovich, M., J. Date, et al. (2003). Discovery point: Enhancing the museum experience with technology. CHI '03 extended abstracts on Human factors in computing systems, Ft. Lauderdale, Florida, USA, ACM Press.


Cite as:


155
How Informances Can be Used in Design Ethnography

Ron Wakkary
Associate Professor
School of Interactive Arts and Technologies
Simon Fraser University
2400 Central City
Surrey, B.C. Canada
rwakkary@sfu.ca

Travis Kirton
School of Interactive Arts and Technologies
Simon Fraser University
2400 Central City
Surrey, B.C. Canada
tkirton@sfu.ca

Corey Julihn
School of Interactive Arts and Technologies
Simon Fraser University
2400 Central City
Surrey, B.C. Canada
cjulihn@sfu.ca

Madison Poon
School of Interactive Arts and Technologies
Simon Fraser University
2400 Central City
Surrey, B.C. Canada
mpoona@sfu.ca

Leah Maestri
School of Interactive Arts and Technologies
Simon Fraser University
2400 Central City
Surrey, B.C. Canada
lmaestri@sfu.ca

Ryan Betts
School of Interactive Arts and Technologies
Simon Fraser University
2400 Central City
Surrey, B.C. Canada
rbetts@sfu.ca

Abstract
In this paper we discuss how we’ve adapted the technique of informance design for use in design ethnography. We detail our design ethnography workflow method and describe our informances.

Keywords
Informance, ethnography, design, pattern language, home life, family, domestic.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
Informance design is a combination of role-playing, improvisation, and bodystorming (an embodied form of brainstorming), which enables designers to generate design ideas through imagining themselves in the minds and bodies of users. In this paper we present an adapted use of informance design for the purposes of design ethnography. We discuss our use of informances and present a workflow method that incorporates the technique into a design ethnography study. We will show how we linked the informances to the development of key design patterns. Our aim is to describe our method such that other designers can understand, use, and modify it.
In our use of informance design we found that it made three key contributions to our design ethnography: 1) informances provide a phenomenological analysis that adds a first-person and embodied representation to the traditionally textual and third-person representations in ethnography; 2) Informances provide early representations of key insights into the interpretation of collected data; 3) Informances provide a shared representation that facilitate a group approach to design ethnography.

Background
Informance design is the use of role-play, which enables designers to enact complex situations through imagining themselves within the minds and bodies of the users [6]. This method was seen as a performance technique for generating novel design ideas that were embodied and in situ [2]. Ongoing use has been documented in organizational settings and mobile communication [5]. To our knowledge, informance design has not been discussed or used as an analytical or ethnographic tool.

Recognition of the social dimension of design has increased the use of ethnographic techniques in Human-Computer Interaction (HCI). Suchman demonstrated the clear efficacy of applying ethnomethodological techniques to design [7]. Discussion has since ensued on the specifics of how to apply ethnography to technology design [3]. Recent studies have utilized design ethnography in trying to understand the home and the role ubiquitous computing might have in supporting domestic routines [4]. We note that as designers have appropriated ethnography for design, design ethnographers have in turn appropriated design techniques for the purposes of ethnography. One such example is the use of pattern languages in design ethnography [4]. Our novel use of informance design follows in this tradition.

Motivation
In the past, we have used informance design in several projects. The technique offered the advantage of wide exploration of design ideas early in the design process. The informances served as preliminary scenarios that would later be refined, scripted, videotaped and used to generate requirements. In more recent projects, team members used informances as an analytical tool to describe and share observations from fieldwork. Due to the nature of the projects, these informances were often of physical and embodied actions like individual interactions with museum displays, or physical play in playgrounds. As we were about to embark on a long-term design ethnography study, we decided it was time to formalize the use of informances as an ethnographic tool and create an appropriate workflow process.

Our study
In order to provide context for the paper we briefly describe our study. The aim of the study (not the subject of this paper) was to describe everyday design in which design is a form of use. We see families as a type of everyday designer who remakes or modifies organizing systems, and who use design artifacts and actions around them as design resources [8].

Our study included three families with young children ranging in age from 5-13 years old. The parents were professionals (elementary and secondary school teachers and a legal aid worker) ranging from new to established in their careers. The study occurred over a five-month period, and included over 350 hours of
observations and interviews. We used 3 ethnographers, each assigned to one family.

**Design ethnography workflow**

The design ethnography process involved several steps. Our process revolved around *sessions*. A session is a single visit by an ethnographer (typical sessions are 1 to 3 hours in length). Ethnographers participated in 3-4 sessions a week. Initially, sessions were scheduled in order to describe the different routines during the weekdays and weekends. Later, scheduling was based on targeting specific activities or events. Each session was documented in three steps.

The first step included documentation and summaries of the observations. Each ethnographer completed this immediately following a session. This involved consolidation of notes, selection of photographed images, videos and audio recordings, as well as further reflection on the observations. A standardized written report was sent to another team member for review to seek feedback to help clarify descriptions and aid further questioning. These comments became annotations that were recorded as part of the document. The completed written report was entered in a database.

The second step of the process was to share observations with the entire team. It was during this step that informances were utilized. The team met weekly. During these meetings each ethnographer reported findings, performed an informance, and the group discussed the significance of those findings. Each meeting (including informances) was documented and entered in the database as analyzable material.

The third step of the process focused on a summative analysis of the gathered data. One of our analysis approaches was the use of a pattern language\(^1\) [1]. The pattern language formed a distillation of repeatable observations and significant attributes of activities.

**Weekly Cycle**

The weekly meetings discussed in the previous section were organized into three activities: 1) oral report, 2) informance, and 3) discussion.

*Oral reports* are when an individual ethnographer summarizes the sessions that he or she participated in during the previous week. The report includes an explanation of significant observations and insights. The ethnographer will also present related images, sound or video that they have captured.

*Informances* complete the report. Each ethnographer re-enacts a specific moment from one of the week’s sessions. An ethnographer will assume the character of one (or more) of the participants (informant or actor in the study\(^2\)), other researchers sometimes acted as secondary participants in order to provide a social context for the informance. While the goal of informances is to describe observations, they provide additional value:

\(^1\) A pattern language is a structured method of describing good design practices in different domains such as architecture, urban planning, computer science, human-computer interaction, and interaction design.

\(^2\) The terms informant and actor are common ethnographic terms. An informant is the person who is relaying an event to the ethnographer or someone who provides key information. An actor is a person involved in the action of an observation.
The performance provides the ethnographer a rich opportunity for reflecting further on the situation and people observed. The degree of empathy required gave the ethnographers a chance to get into the body and mind of the informant(s) or actor(s).

Acting provides the team the opportunity to imagine being with the ethnographer when he or she was observing. This creates a rich impression and leads to members asking more relevant questions and having a clearer understanding of the participants in the study.

For team members who do not participate in any fieldwork, informances provide an opportunity for them to become aware of what is actually happening in the field. These team members can provide fresh insights on an observation.

Ethnographers were asked to explain why a particular observation was chosen for an informance. Predominantly, observations were chosen because they were particularly noteworthy or the actions were primarily physical and an informance was the best way to describe the observation.

Discussions follow reports. Discussions involve the whole team. Typically questions are asked seeking clarification and team members probe recollections for further details and relationships.

Informances
Below are two examples of informances (the names of actors and informants have been fictionalized).

Call Home Informance
In figure 1, one of the ethnographers, Leah, is enacting an observation of an informant, Kevin. Kevin is trying to use the voice recognition feature of his new mobile phone. Leah is in fact relaying how Kevin wanted to show her this feature. In the informance, Leah plays the role of Kevin and uses an eyeglasses case as a prop for the mobile phone. Another researcher plays the role of Leah.

The informance starts with Leah acting as Kevin asking Leah if she wants to see how the voice recognition works: “This is my mobile phone and this is Bluetooth...” Kevin puts the headset around his ear. “It works best for me because I like to talk on the phone while I drive. I’d like to show you how it works.” “Sure” replies Leah. After punching keys on the mobile phone, Kevin holds onto the Bluetooth headset and says, “Call home.” Nothing happens – Leah steps out of character to make some mobile phone sounds. Back in character he mutters to himself, “Ok, it doesn’t work,” then he says to Leah, “sometimes it doesn’t recognize certain syllables...let’s try again.” Again Leah steps out of character to make phone sounds. She resumes her role, holds the headset again and says, “Call home.” Again nothing happens. Furious punching of keys follow and then with some resignation Kevin admits, ”there are too many things on this mobile phone...” Kevin continues to push buttons. “See, you know, certain words it works. I call CleanTech – here, I’ll call CleanTech. CleanTech!” Triumphantly Kevin shows the display of the phone around the room. “See it’s connecting...it’s connecting. So why won’t it call home?” The informance ends as Kevin continues to push buttons. The informance lasted less than a minute.
In Leah’s observations, Kevin never gave up on the voice recognition feature. The informance showed the balance between frustration, satisfaction and perseverance with using a new feature when a need is identified as Kevin remarked about the use of his mobile phone while he drives.

**Preparing a Salmon Filet Informance**

In figure 2, an ethnographer, Corey, enacts the preparation of a salmon filet for cooking by one of the participants, Andrew. As Corey is acting out the informance he explains the lengthy procedure that Andrew goes through in order to prepare fish for cooking on the barbecue. While he is acting out all the tedious details of the process, Corey notes Andrew’s humor about the event as well as Andrew’s meticulous attention to detail throughout the whole procedure. Andrew used needle-nose pliers as a tool for preparing food. The actual event took nearly an hour and involved only preparation (i.e. no cooking).

Corey used paper as a prop for wrapping the salmon and an imagined pair of needle-nose pliers. The majority of the informance was devoted to Corey enacting how Andrew would feel along the top of the fish for bones and then slowly and surgically remove each bone with the pliers. This action was repeated several times. The informance also enacted the methodical process of wrapping the fish in foil. In addition, Corey described the arrangement and proximity of the sink, garbage and other kitchen utensils. He used chairs in the room to reproduce the kitchen configuration. The informance was slightly longer than two minutes. Corey’s observation captured in this informance showed the precise action of removing the fish bones and the need for a specific tool, the needle-nose pliers. Needle-nose pliers are not common in the kitchen and furthermore are here used in an unusual but appropriate way.

**Patterns and other representations**

As discussed in the section, “Design Ethnography Workflow,” the third step in our process was to analyze the data. In our case, we used pattern languages as a method for describing commonalities and key activities across our observations. The informances provided formative representations of key activities and patterns before this step. For example, the Call Home informance described above, presaged and informed the pattern named Training the Mobile Phone (figure 3). The observation in the Preparing a Salmon Filet informance was formalized in a pattern named Using Needle-Nosed Pliers for Pulling out Fish Bones (figure 4). These patterns were significant patterns from our pattern language of over fifty patterns.

**Discussion**

In our use of informances we found some key benefits. The informances such as Preparing a Salmon Filet, provide a phenomenological account of an embodied action. There was no dialogue in this observation, and the insight of the effectiveness of the use of pliers could only be demonstrated as a precise physical action. We feel a traditional written account of this would miss key aspects of this observation. In addition, informances provide early representations of key insights into the interpretation of collected data. While we used patterns in our summative analysis, other studies may represent their observations in traditional vignettes (ethnographic stories), scenarios, video documentary or other forms of representation of data analysis. Informances as we’ve described provide key insights to later refine with
Training the Mobile Phone

Significance: People will invest time and effort in integrating a new artifact if they see a long-term benefit that outweighs the short-term investment of trial and error.

Description: Changing our routines, or taking the time to train others or learn new systems in order to incorporate a new artifact or system.

Example: K made repeated attempts to train his mobile phone voice recognition because he knew it would save him time in the future.

Figure 3 Training the Mobile Phone pattern

Using Needle-Nose Pliers for Pulling out Fish Bones

Significance: Unusual but appropriate use of a tool.

Description: Using a tool that was not designed for the activity but is a perfect match for the task at hand.

Example: A used needle-nose pliers to remove fish bones.

Figure 4 Using Needle-Nose Pliers for Pulling out Fish Bones pattern

more detailed representations. Furthermore, informances will inform these representations from a phenomenological viewpoint that may be complementary. Lastly, our informances provided shared representations that facilitated our group approach to design ethnography. This allowed all members of the team into the process providing for considered and informed discussions of the observations.

Informances can pose challenges as well. Some individuals may not be comfortable performing. Informances, like participant observation is a learned skill. Teams may need to devote time to practice and learning in preparation for a study. Finally, some may question the validity of the data from informances when used as an analytic tool. In many respects, informances used ethnographically assume the validity position of ethnography, which relies on reflexivity. Reflexivity is a conscious self-understanding of the research process and the use of explicit methods to guide the process. This does not satisfy validity and reliability measures of experimental design, yet it is the in-depth interpretations of informants and ethnographers together that provide the rich contextual understanding resulting from ethnography.

Conclusion

Based on our experiences, informances serve as an embodied technique for formative analysis that can guide the ethnographic process as it evolves, can provide early analysis of significant activities, and provides a rich and embodied communication of observed experiences to all members of the team. The contribution of our paper is that we adapted the informance design technique to be an ethnographic tool. We leveraged the analytical strengths of informance design and demonstrated that the technique is as valuable in performing actual events as it is at generating imagined actions.

Acknowledgements

The Social Sciences and Humanities Research Council of Canada funded this research.

References

The Role of Participatory Workshops in Investigating Narrative and Sound Ecologies in the Design of an Ambient Intelligence Audio Display

**Milena Droumeva**  
School for Interactive Arts and Technologies  
Simon Fraser University  
13450 102 Ave. Central City Towers  
Surrey, BC Canada  
mvdroume@sfu.ca

**Ron Wakkary**  
Associate Professor  
School for Interactive Arts and Technologies  
Simon Fraser University  
13450 102 Ave. Central City Towers  
Surrey, BC Canada  
rwakkary@sfu.ca

**ABSTRACT**

We describe two participatory workshops conducted to support design decisions in the making of the audio display for an ambient intelligent game platform. The workshops discussed here explore specific issues of players' interactions with sound and auditory display design. The workshops helped move our design process forward by specifying the role of narrative and sound ecologies in our design. They clarified the role of sound in creating narrative coherence, guiding player actions, and supporting group interaction. We describe the workshops, the auditory display issues we addressed, discuss how the workshops helped inform our subsequent design, and extend recommendations on how participatory workshops can be used by other designers of auditory displays.

1. **INTRODUCTION**

Ambient intelligent (Ami) environments rely on ambient interfaces in which sound plays a critical role. Our current research involves the design of an Ami environment for a physical multi-user game known as socio-ec(h)o. Socio-ec(h)o delivers a responsive environment through sound and light display. In this paper we describe our use of participatory workshops as a method to support design decisions in the making of our audio display.

Ambient intelligence is the embedding of computer technologies and sensors in physical environments that combined with artificial intelligence, respond to and reason about human actions and behaviours within the environment. In our final prototype, the socio-ec(h)o game is played by a team of four players and features six levels of increasingly difficult word puzzles solved by coordinated body positions and movements. The environment is responsive to players' actions through abstract light and sound. Players' movements are tracked using a motion capture system. A custom reasoning engine was developed to track the game state and infer players' actions. An audio engine was developed to create real-time and responsive sonification. The audio display creates the aesthetic feel, represents current game state, and guides future player actions through abstract and ambient real-time sonification. For more technical details of our prototype see [1].

Our overall research goal is to develop approaches for group communication, collaboration and skill acquisition in Ami environments. A game platform like socio-ec(h)o provides both a prototype and test environment.

In this paper we describe two participatory workshops that enabled us to develop an approach for the audio display, which in the end provides a continuous, ambient response along a gradient. Yet the path to this outcome was not a straight line. We explored related ideas of group interaction, players as game avatars, physical movement in responsive environments, perceptions of immersion, game mechanics, and play through an emergent and participatory design approach. This primarily involved an extensive series of participatory workshops and iterative prototyping. Given the systemic and experiential nature of Ami environments, we would argue that few other design approaches would be as effective. In fact, we offer participatory workshops as an alternative to more controlled experiments and usability approaches currently used in employing and testing the audio design techniques we investigated, within similar contexts. The goal of this paper is to describe how we approached designing audio displays for our participatory workshops and to inform readers of the lessons we learned. We hope to encourage the use of participatory workshops in future audio display designs.

The workshops helped move our design process forward by specifying the roles of narrative, game identities and sound ecologies to our design. As a result, the experience of a sound avatar was a significant emergent game element. Sound characters, or unique sound characteristics, enabled players to assume a story world identity that fostered communication, exploration, skill acquisition, and a sense of progression. Ultimately we drew on this experience of unique sound characters in our design for the final prototype, thus providing a narrative coherence to each game level and overall sense of progression that guided participants' actions.

In this paper we provide theoretical background on the role of audio display in games and participatory workshops. We follow by providing the context for our participatory workshops and an overview of the audio design concepts we investigated. We describe each workshop by detailing the issue addressed, the structure of the workshop, our technical set-up, and lessons learned. We end the paper with a discussion, reporting on future work and conclusions.
2. BACKGROUND

Sound is an important channel through which humans perceive their natural or designed environments. There is a proliferation of literature on auditory perception and design for task-oriented, highly computerized environments [2, 3, 4, 5], virtual audio and spatialization [3] and sonification of information clusters [2, 6, 7], yet relatively few of these approaches to sound design have found their way in the design of sound for games [8, 9]. While our conceptual sound design included a variety of existing research in psychoacoustics [2, 10] and the ecological approaches developed by Schäfer and Truax [11], creating participatory workshops was a crucial step in addressing the issue of interactive audio display for an Aml environment.

Due to the complexity of the design concept – to create a game that is played by a team, has structure and rules, offers challenges and affordances, is physical and spatial, and responds to user actions only through its environment, we needed a design methodology that would allow us to explore the richness of the situation. Traditional formative evaluation and usability methods for auditory display simply do not provide a setting that is ecological and holistic enough to allow for actual participation and involvement. As Gibson points out, quoted by Neuhoff [10], “awareness is rooted in meaningful experience of the environment: thus ecological validity results from studying subjects/people in their own natural environment, in motion, in active exploration. For people this environment is social, cultural, systemic, economic, political, etc.” Since we were interested in examining social contexts, group interactions, and embodied experiences, we adopted the approach of participatory workshops.

Participatory workshops can be viewed as a design method based on Participatory Design (PD). PD emerged to address social, technical and power relation issues in designing within organizations [12]. Traditionally, the method involves lengthy involvement with stakeholders within the users’ settings that result in an empowered stakeholder and informed designer co-designing solutions [13]. Participatory workshops adopt the principles of genuine user participation, design within end-user settings, and enabling participants to co-design, however in a severely shortened time period and without the goals of in-depth contextual design, transformation of users into designers, or systemic sustainability. Rather, the workshops are a quick, flexible and powerful tool that allows designers to investigate specific activities, situations and environments. The aim of such workshops is typically to move beyond traditional user-centered design to harness participants’ creativity in understanding what they make as well as say and do [14], or to utilize expertise for hands-on concept development [15], as well as model and manipulate simulated environments through role playing [16]. This last approach is closest to our use of participatory workshops in socio-ec(h)ic. Within our research, we have previously employed participatory workshops to investigate multiple approaches to situated activities [17].

3. DESIGN PROCESS AND WORKSHOPS

The two workshops that we describe came midway through the design process. We had previously hosted several other participatory workshops and development meetings where we developed the conceptual foundations of socio-ec(h)ic, which included core game mechanics, game progression and structure, and narrative development. We had yet to build a working prototype. Our main concern at this stage was the design of a compelling environment based on user engagement, movements in physical space, immersion, and narrative or game progression. We knew at this point that we needed to investigate specifics in the role that the audio display would have. We had determined that the technical preconditions included location tracking, and an ambient interface that might involve body and object movement, location, and gesture. Given the Aml nature of the project we ruled out a graphical user interface of any kind.

Both workshops were set within the physical game space: a black box environment with controlled light and sound displays, delivered via Wizard of Oz techniques. A Wizard of Oz experiment is one that simulates the functionality of a technological system without actually building an autonomous prototype. While participants interact with the “system” as if it were autonomous, human researchers provide behind-the-scenes functionality in response to user actions (i.e. bringing up light or sounds, triggering events). This method allows for exploratory testing of user interactions and experience patterns. It focuses on the effectiveness and possible uses of the simulated prototype, rather than on the usability of an entire system.

Our conceptual starting point for the audio was the use of sound avatars that would allow game characters to be used in the game mechanics and would provide a vehicle for narrative progression. Both workshops were organized around this conceptual starting point. The underlying focus was on the interaction patterns between players and system, and the role of the ambient response in audio and light. We subsequently invited participants to suggest changes in the environment and interaction rules based on their experience of the environment and their avatars.

4. SOUND DESIGN CONCEPTS

Here we discuss design concepts of auditory display that we incorporated into our participatory workshops in order to examine their effectiveness. These concepts were selected on the belief that they might be useful in communicating game information between players and system, signaling progress and changes in the game, and creating an immersive atmosphere. As mentioned above, these concepts are based in applied psychoacoustics, as well as the acoustic communication framework of soundscape design [11], where sound mediates the relationship between listener, environment and soundscape. Below we present the general ideas for audio display, and their relevance to socio-ec(h)ic.

4.1. Keynote sound as “Ground”

Sound is an extremely powerful tool in creating a sense of ambiance in a space, as well as fostering an evocative cultural experience for the users. This is exemplified in various media, especially cinema. In both our workshops, we felt it was important to create an atmospheric keynote [11] sound that would serve as a ‘ground’ for localization of the game’s acoustic space, and situate the rest of the auditory display within its context.

4.2. Musical Expression as Avatar

A core mechanic in our game was player identity. Through previous design sessions we had come to the idea of using unique sounds within a system of display to represent game characters. One way to do that through sound is by using...
musical sound. This "musical expression avatar" approach would utilize a discrete musical (MIDI) phrase to sonify players' identities, their actions and spatial location. This model rests on a long lineage of recognition of a sequence of periodic sounds, and template matching [10]. We used this approach in participatory workshop one, in the form of four individual parts of a counterpoint MIDI composition. Given that the phrases could combine in a number of different ways, we wanted to know how well this approach could work in terms of recognition and identification, and in the formation of sound ecologies, narrative and play.

4.3. Environmental Metaphor as Avatar

This approach, like the previous one (section 4.2), relies on pattern recognition and template matching of a discrete sound, unique to each player. Yet we felt that the richness of environmental sound alone deserves exploration as a vehicle for facilitating recognition, identification with character, and narrative possibilities. In this concept we were inspired both by Schafer and Truax's work in acoustic ecology [11], and their classification of the natural sound environment. We also were inspired by Ballas' work in holistic everyday listening. In which we detect small changes in sound quality when we are required to extract information from sound [10]. Since this approach is the most intuitive and qualitative, we anticipated that it would be hard to gauge its effectiveness.

4.4. Timbre-Based Sonification of Game Events

Besides using avatars to sonify player identities, we wanted the auditory display system to represent subtleties in the play and game shifts that reflected players' groupings, their level of activity, their proximity to one another and to the sound sources. We felt that this could be achieved by using timbre changes as a model for qualitative sound coloration in the real world. Timbre could be affected by applying simple reverberation to a given sound source. Players will have to listen for complexity, colour and quality of the sound. In simple perception terms, this technique could be categorized as a gradient of "muffled" to "bright" or "distant" to "close-up" sound. This approach is based on holistic everyday listening, in which we detect small changes in sound quality when we are required to extract information from sound [10]. Since this approach is the most intuitive and qualitative, we anticipated that it would be hard to gauge its effectiveness.

4.5. Spatialization of Sound

This approach comes from existing literature on sonification of information and attention management [2, 7, 4, 6], which suggest that separating sound in different spatial locations helps in recognition and interpretation of its significance within a rich system of information. In the context of socio-ecological theory, we hypothesized that this approach would work well with our attempt to introduce sound avatars as a game mechanic. We thought that spatializing sound and mapping its virtual location to the physical location of participants would reinforce the connection with individual sound avatars. We used this approach specifically in participatory workshop one.

4.6. Hierarchy of Auditory Display

While using a single sound avatar proved promising, there was the potential to provide more coherence and richness of the sonic characters. As research in auditory icons design suggests [2, 10], a hierarchy of internally and semantically consistent audio icons create better recognition and facilitate navigation and utilization of a system. Thus this approach uses a set of three semantically related sounds that increase in perceived intensity in order to represent each game avatar. Essentially, we constructed a hierarchy of sound signals related to gameplay, movement and character (see Table 2). This approach was introduced in participatory workshop two.

4.7. Delayed Feedback as Core Game Mechanic

This approach developed in response to participatory workshop one, where the feedback was continuous and constant. Instead, utilizing this concept in participatory workshop two, we provided no auditory feedback unless players achieved a specific configuration of spatial positions. Further, the feedback was delayed in that players had to hold their position for at least 3 seconds before they were rewarded with an auditory response from the environment. This model specifically explored the idea of subverting sound's traditional role as auditory reward in computer games by delaying the sonic gratification in order to establish clearer yet subtler gameplay.

5. PARTICIPATORY WORKSHOP ONE

We describe the first participatory workshop by discussing the sound issues we addressed, the structure of the workshop, and design investigations. We provide technical discussion of the sound set-ups and wizard of Oz techniques. We conclude this section by detailing what did and did not work, and the lessons we learned. Participants were students at the university and included both males and females aged between 22 to 34 years old.

5.1. Sound Issues

In this first workshop, the specific sound issues that were explored included: the introduction of personalized sound avatars (including their spatialization patterns); the effectiveness of using different sound categories – music, voice, abstract, or environmental sound, with regard to observed recognition, interaction and play; and the use of audio process [reverberation] and amplitude in sonifying player location and team activity.
5.2. Structure of Workshop

Within the black box space, a circular area was designated as the interaction/play space. Four speakers were placed on the floor back to back, forming four semi-distinct acoustic spaces, or zones (see figure 1). The zones were dynamically created by players' interactions and groupings.

The four participants were first engaged in a pre-discussion during which they were given basic information about the workshop. The participants were told that they would have an individual "sound avatar" but were not told what it would be; and that the avatar would "follow" them in space.

The four participants acted as a team. The workshop consisted of three parts. The first stage was an exploration in which each player individually explored their "sound avatar" and its behavior in the physical game space based on their actions (see figure 2). In the second stage, all four players explore the space together and discover relationships and audio combinations that they can create with their sound avatar and other participants' sound avatars. The third stage included discussions, suggestion for changes, and real-time implementation of some of the suggested changes.

5.3. Design Investigations

While we had spent most of our conceptual design work on developing game and narrative progression, this workshop was open-ended in terms of narrative and game mechanics. We looked to the workshop to explore possible options. The workshop investigated the following game event and narrative components:

- **Discovery of sound avatar** (who you are in the game);
- **Discovery of audio combinations** with other participants (exploration and manipulation of collective identity);
- **Sound Ecologies Challenge**, this challenge addresses players' ability to affect the environment by forming new ecologies (their movement stimulates the dynamic soundscape - in periods of inactivity the environment decays);
- **Ecology As Metaphor**, discovering the right configuration of players and/or activity that will result in a prominent sound ecology. Conversely, the "wrong" combination of players and/or activity could result in a negative ecology or complete decay.

### Table 1

<table>
<thead>
<tr>
<th>Processes</th>
<th>Activities</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplitude</td>
<td>Proximity/Level of activity</td>
<td>Easy to perceive for individuals, harder for groups</td>
</tr>
<tr>
<td>Panning</td>
<td>Player position in space</td>
<td>Easy to perceive for individuals, harder for groups</td>
</tr>
<tr>
<td>Reverberation</td>
<td>Proximity to sound source (speaker)</td>
<td>Harder to perceive for individuals, requires fine timbre recognition and conceptual mapping</td>
</tr>
<tr>
<td>Sound ecologies mix</td>
<td>Player groupings</td>
<td>Easy to perceive for groups. Masking an important issue. Depending on the above components</td>
</tr>
</tbody>
</table>

5.4. Sound Settings One

In the first stage of the workshop, we used the **Musical Expression as Avatar** approach to audio display. A Bach counterpoint piece (in MIDI) was deconstructed into four parts and each was assigned to each of the four participants. The sound avatar (a subset of the Goldberg variations) physically followed each participant, thus reinforcing a sense of association with a sound-based avatar.

The sound ecologies that players constructed by movement formed different musical orchestrations. Here we also tested the use of reverberation to affect timbre of sound and mapped this to distance (proximity to sound source). Amplitude levels were mapped to intensity/level of group activities and movements in the play space.

Our objectives were to observe the perception of musical parts, on their own and in different combinations, as well as the perception of distance through timbre in the form of reverberation, and amplitude levels (volume) as a response to levels of activity. Since we were manually driving the system's response, we were able to adjust those elements fluidly throughout the session. In an exploratory way, we also wanted to test notions of emergent play, free-form play, movement, gesture, and social interactions.

5.5. Sound Settings Two

In the second stage of the workshop we used the **Environmental Metaphor** approach to sonic display. Four environmental sounds signifying earth, fire, water and wind were created. We again used the idea of creating combinations and ecologies with the sound metaphors. Also, amplitude levels and reverberation were used to make subtle changes in the sound, as a response to activity, movement, proximity to sound source, and types of participant groupings. Two known issues we were aware of in working with environmental sound, were the semantic mappings between avatar and its sound representation, and masking. Naturally, water and fire are concepts that could be translated in direct representations, while earth and wind are more metaphoric representations. For example we used footsteps and a processed windy sound, respectively.

The workshop objectives were in large part similar to those of stage one. The idea was to explore how environmental sound, as a different sound category from...
music/MIDI, would influence participants’ experience of exploration, discovery, interaction and play. We also wanted to see whether having environmental rather than musical sound would encourage formation of sound ecologies or game narrative in any specific ways. The display was once again restricted to only one sound for each category. We used amplitude, panning (moving sound around to different speakers), and reverberation to respond to patterns of game events and/or player actions. For mappings of the Timbre-Based Sonification of Game Events in both stages of the workshop see Table 1.

5.6. Wizard of Oz Techniques

For this workshop we created a custom Wizard of Oz sound display tool in Max/MSP (www.cycling74.com) (see figure 3). The Wizard of Oz method is the manual simulation of unimplemented technology. Besides playing the four parts of the Goldberg Variations (stage one) and the four environmental sounds (stage two), the audio tool allowed us to spatialize sound, add reverberation, vary sound levels, and apply granulation. The tool was operated via a UC-33 MIDI controller for faster response time, and operated by two people at once.

We had an audio station set to the side of the interaction space enabling a clear view of the participants’ location, actions and group configurations (see figure 2). In stage one, we only played the sound avatar of the individual who was in the space. We attempted to localize the sound wherever the player was, as well as increased or decreased the sound’s amplitude depending on how close the player was to a speaker. In the second stage, we monitored and responded to the formation of particular ecologies, by mixing in the different musical parts, or different environmental sounds. As well, we introduced reverberation as an indication of distance of player(s) from sound source (a speaker). Amplitude, on the other hand was increased when lots of activity and play occurred, and decayed over time if there was little or no movement.

5.7. What Worked

For the most part we were able to respond instantly to relevant participant interactions, and reward behaviours that we wanted to encourage. The value of using a Wizard of Oz approach, is that in a workshop, which is highly exploratory and quite loosely structured, it allows us the freedom to examine what kind of activities and interactions we want to support, encourage and reward. It also allowed us to spontaneously and dynamically adjust the sound display to match the players’ ease-of-use of the system, thus facilitating their engagement and playfulness. The technique also allowed us the ability to improvise and bring in a special sound reward (a granulated vocal composition) if we felt the players achieved a particularly creative configuration.

5.8. What Didn’t Work

In stage one, we could not make the personalized sound parts follow the individual as we had intended (due to a programming flaw), thus everyone had difficulty identifying with their sound in a spatial sense, and they felt as if the ecologies they formed as a group were random and arbitrary.

In stage two, we were able to make the sound follow the participants across two directions, which dramatically improved the reported individual experience of avatar discovery. However, the sound couldn’t follow participants everywhere, which limited the directionality of the ecologies created. Also the sounds of water and wind proved to be too broadband and partially masked the other two sounds in the ecologies. This resulted in players’ inability to identify their sound, and to develop strong affinities with it.

5.9. Lessons Learned

After the first workshop, our formulation of the core game states were distilled into the following conceptual aspects:

- Evolution (interactions and game state shifts)
- Relationship between play and mastery (skill acquisition)
- Discovery and exploration (as core game mechanics)
- Game types (how different players affect the game)
- Characters and identity (sound avatars/characters)
- Narrative represented by environment (sound, light)
- Sustainability (engagement and generative play)

The ideas of ecology related more to testing the development of narrative and story world. Ecology was understood to encompass both environmental ecology – the internal consistency of the ambient immersive world, and social ecology – the sustainability and engagement in social group models that were formed. Again, even though the structure of the workshop was not formalized, the design had a built-in internal consistency of representation through sound – sound avatars did not mutate or change during the play session, rather they created multiple ‘sound mixes’ based on groupings and activity.

Even though players were not able to identify their sound part within the group musical composition, their movement across space indicated that they didn’t feel constrained by this. Their actions showed that they kept trying to affect the system in some way and get a response or a clear idea of their sound part – what it is and where in the acoustic space it is. We are unable to say whether participants consciously registered the changes in timbre/reverberation and how it may have affected their interactions or movements. It was clear from the discussion that changes in amplitude levels have a much stronger perceptual connection to the sense of responsiveness of the system. We believe that spatialization of the sound was beneficial, especially in the individual sessions, however, since we could not fully simulate it again, it is unclear how this would have affected the game.

While players seemed much more playful with the music components in the workshop they reported a high
satisfaction with having environmental sound characters. Participants felt they were able to identify with them. The players with fire and water avatars reported feeling particularly attached to their "character" and feeling a strong sense of projected identity.

In the second stage’s group experience, all participants expressed a desire for sound to better support the "narrative" formations that they were trying to construct. For instance, the fire avatar employed gesture and composition-like movements, and acted "threatening" to other sound avatars (except water, since a participant declared that water quenches fire). All participants seemed frustrated with not having strong enough feedback and requested a more "clear feedback" system response. In terms of sound narrative with environmental sounds, participants suggested that a greater array of sounds should be used to represent each character, rather than have only one sound.

6.2. Structure of Workshop

Participatory workshop two was also set in a black box space. A circular play space was marked out on the floor and a white curtain tied in the middle, hung from the ceiling. Overhead lighting kits were used to create 4 circles of light within the darkened space. A 4-channel sound system was implemented in the space (see figure 4).

Similar to workshop one, participants were assigned sound avatars and the first stage allowed for individual exploration of the avatar. As well, participatory workshop two consisted of two stages, each containing two parts: individual and group exploration.

6.3. Design Investigations

In participatory workshop two we decided to employ restraints in order to encourage specific types of interactions related to game mechanics, namely, exploration, discovery and achievement. We aimed to build on the free play of the previous workshop and see if a more explicit audio and visual display could shape the play and ultimately encourage skill acquisition in the game and conscious exploration of the game rules.

We focused on clearly describing different aspects of gameplay by explicitly directing what players can and cannot do. In the individual and group sessions, the four circles of light were the only places where players would get auditory feedback, which might be their sound avatar or combinations of sound ecologies. Other areas of the space were non-interactive. The group play was more restrictive, since we had decided to only encourage even groupings of 2 and 4 participants in the lit zones but not of configurations of 1 or 3 in lit zones (see purple diamonds in figure 4). Thus whenever only one participant occupied the spotlight, or two participants were joined by a third, the system decayed and stopped auditory feedback (see figure 5).

An additional nuance to workshop two was time as a variable. We designed the sound and light feedback to be delayed, requiring a player or group of players to linger inside a circle of light for a fixed amount of time (between 2 and 4 sec) before the system responded. Again, this was an
attempt to shift attention to actions and system response. The workshop followed this sequence:
1. Skill acquisition of game rules (delayed feedback rule – player have to be in a circle for at least 3 secs)
2. Discovery and exploration of individual avatar (players discover their sound avatar and its three levels of intensity, see table 2)
3. Forming ecologies – skill acquisition (players learn that combinations of 2 and 4 participants create an auditory and visual response)
4. Forming ecologies of sound (players explore sound ecologies and learn other players' sound avatars by entering into combinations with them)

6.4. Sound Settings One
In workshop two we based our auditory design approach entirely in the Environmental Metaphor as Avatar concept. We used environmental sounds to create a system of auditory icons using the four different avatars. We enriched the complexity of the sound characters by adding increasing intensity to the base sound (see table 2). This approach aimed to support participants' request from the first workshop for more narrative dimensions to the sounds and stronger and richer qualities in the sound avatars.

6.5. Sound Settings Two
This setting was identical to the first sound setting (section 6.4), with the exception of a constant and light ambient sound of frogs' chorus. We also introduced a wildcard auditory event triggered when a single player spent more than 3 seconds in a circle of light. In this case, a composed musical sound (light marimba musical phrase) was played.

6.6. Wizard of Oz Techniques
Our Wizard of Oz set up was virtually identical to that of participatory workshop one (section 5.6). Only in this workshop we included four sets of three environmental sounds and several “wildcard” pre-composed auditory rewards.

In terms of system response, the environment was “silent” unless one of the four circles of light was “activated” by participants. Only one circle of light could be active at a time, and the first player or group to achieve a desirable configuration and hold it, would determine the audio and visual response. As mentioned above, participants had to stay inside the circle for at least 3 seconds in order to hear a sound. The longer they stayed in the circle, the more “intense” the sound would become; the sound would cross-fade through its three intensity levels (see table 2) and would increase in amplitude over time.

6.7. What Worked
This time our sound set up was relatively easy and we managed to respond through audio display precisely. In addition, we were able to reward a few moments of play/exploration with a pre-composed vocal/musical sequence. For example, one player started interacting with the scrim in the middle of the dark space (a white transparent curtain) and we were able to respond to her touching gesture with great precision. Soon other players joined too and explored this newly discovered system feature.

6.8. What Didn’t Work
Because this iteration of the participatory workshop had a tight structure and system rules, the only form of feedback we provided were sound avatars in various combinations, and two accent sounds for each avatar. While we still varied amplitude depending on duration of player actions and movement, this session resulted in a rather dull soundscape. Players later reported being quickly bored with it, after they discovered their avatar in all its dimensions.

6.9. Lessons Learned
Individual explorations with sound avatars went quite well in this workshop. Players typically spent an average of five minutes getting to know their avatar and learning how to elicit system responses. Participants reported being frustrated at having to restrict their movement to the lit circular areas and felt disappointed at the absence of sound triggers. They were also disappointed at having to discover the system’s behaviours, rather than the opposite of the system responding to their free flow of behaviours. However, as participants reported, the response patterns of the system were very clear and the learning/exploration and discovery were greatly facilitated by introducing clear restrictions and feedback. Players did a lot of lingering within a light circle and listening (or what appeared to be listening) to the sounds in great attentiveness.

Some players exhibited more compositional objectives, and others more exploratory activity; it became clear that both should be supported by our system. It was also clear that people were thinking about narrative and different groupings of sound avatars and expected more of a reward to different groupings. They expected the system to respond with something more than just the mix of the present sound characters. The challenge is to give them more, yet not confuse them by creating conceptual incoherence in the sound content choices. Gesture mapping to sound could afford for action and effect, and measure effort/speed, thus appearing that players are “making something happen.” Yet since gesture is a compositional tool, it provides a challenge for motion tracking, sound mapping and recognition of cause and effect.

7. DISCUSSION AND FUTURK WORK
The participatory workshops allowed us to make preliminary conclusions about auditory display in our Aml environment
related to narrative, identity and sound ecologies. What was most valuable to us was that the workshops helped move our design process forward in a structured and meaningful way. As a result, the availability, manipulation and experience of a sound avatar were a major emergent game element. Sound characters or unique sound characteristics for individual players enabled players to assume a story world identity that fostered communication, exploration, skill acquisition, and a sense of progression. This approach encouraged specific exploration of the system response based on one's character, making participants more aware of the subtleties and narrative aspects of their experience. This enabled participants to form internal representations, associations and expectations about how a sound could or should act in the game. Ultimately, we drew on this experience of unique sound characteristics and mapped it to game levels in the final prototype, thus providing a narrative coherence to the level and overall sense of progression that guided participants' actions.

To summarize our key observation, the relationship between narrative elements and sound ecologies repeated itself throughout the two workshops. We were later able to draw on the distinct aspects of the sound avatars and game narrative in creating an internal narrative coherence, thus supporting skill acquisitions (learning behaviours) and communication resulting from awareness of manipulation and creation of sound ecologies.

Specifically, at the points of ecology creation, narrative associations were especially evident with environmental sound. In discussion, a participant commented that the "fire" avatar seemed to keep running away from the "water" avatar because water would put out their fire.

We had not anticipated these narrative developments, and therefore had no way of supporting them through the system's auditory display response. Yet these results became extremely useful in future explorations of the use of sound in fostering and developing narrative constructs and an immersive story world. Both these workshops are an example of using representational sound (whether music or environmental) in tapping into evocative individual memory, from where narrative structures are bound to emerge.

Some issues to think about when attempting to explore sound's narrative qualities in gameplay situations, are improving on the internal consistency of sound character "databases" (as exemplified in table 2) and their clear delivery in the game; improving on the mappings between player interactions and system response; working on play/game/event time and state shifts; and supporting different ecological activities inside the play space.

8. CONCLUSION

In this paper we have described two participatory design workshops that were a platform for exploring auditory display issues for Ami environments. We have described our use of participatory workshops, lessons we learned, key observations we identified in sound perception and their impact on the design of our game environment. We hope that we have shown that using participatory workshops as a formative design tool is useful in exploring the multiple dimensions of auditory displays in Ami environments, and situate explorations in sound design in a more ecological, contextual and holistic setting.

9. REFERENCES

ec(h)o: Situated Play in a Tangible and Audio Museum Guide

Ron Wakkary
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada
rwakkary@sfu.ca

 Marek Hatala
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada
mhatala@sfu.ca

ABSTRACT
In this paper we discuss an adaptive museum guide prototype in which playfulness is a key design goal for the interaction experience. The interface for our prototype is a combined tangible user interface and audio display. We discuss how we determined the specific requirements for play through an ethnographic study and analysis based on ecological concepts of Bell and Nardi & O’Day. We found that we could consider play in two main forms in regard to the interface: content and physical play. We also found that play is highly contextual. Designers need to consider the situated nature of play for two reasons: 1) to best serve the overall design purpose; 2) in order to understand the nature and degree of play required. We augmented traditional user experience evaluation methods of questionnaires and interviews with observational analysis based on Djajadiningrat’s descriptions of aesthetic interaction.

Author Keywords
Tangible User Interface, Ludic design, adaptive museum guide, audio display, ethnography, information ecology

ACM Classification Keywords
Augmented Reality and Tangible UI, Interaction Design

INTRODUCTION
In our adult lives play is an experience set apart from our everyday activities: Huizinga referred to play as invoking a magic circle, a liminal space for games [16]. Do we play in museums? Anthropologist Genevieve Bell identifies the notion of play together with learning in museums [2]. She describes museums as different cultural ecologies in which the museum visit has the qualities of liminality (a space and time set apart from everyday life) and engagement (where visitors interact to both learn and play).

Guided by the notion of play in a museum experience we considered playfulness as a primary design goal in the design of the interface for an adaptive museum guide. However, since play is a very rich and open-ended concept, a large part of our research was aimed at first determining and later evaluating the specifications for play in regard to an adaptive museum guide. In this regard, we are discussing a passive sense of play where enjoyment is a key element but it is not a goal in itself, as one would describe playing a game or a child playing with a toy.

We found that play is highly situated. It is important that the playfulness is not perceived to be separate from the museum environment to the point that it is distracting or does not make sense. With respect to our interface, we found that play can take two forms: (1) content play such as puns and riddles in informational content; (2) physical play that in our case consisted of holding, touching and moving through a space. The content play was dependent on surrounding informational cues or referents. The physical play was a simple playful action along the lines of toying with a wooden cube where play is open, subtle and implicit. In both cases, play creates a higher degree of engagement with the museum artifacts.

Our case study known as ec(h)o, includes a tangible user interface (TUI), spatial audio, and an integrated user modeling technique combined with semantic technologies. The focus of this paper is on the user interaction experience, for a discussion of our user modeling approach see [14]. We discuss related research to our case study followed by a discussion of our design motivations. We describe the ethnographic study, which led to our requirements for play. We provide an overview of the prototype that was installed and tested at Canadian Museum of Nature in Ottawa. We conclude with an analysis of our evaluation and a discussion of lessons learned.

RELEVANT RESEARCH
We aim in our prototype system, ec(h)o to maintain a standard level of functionality with the exception of media richness. We sacrificed the ability to deliver diverse types of media in order to gain the opportunity to move

1 The name ec(h)o comes from the audio experience of the work and the underlying idea of a museum as an ecology.
away from a graphical user interface (GUI) and the personal digital assistant (PDA) in the hopes of creating a more playful interaction through a physical and embodied interface. We also focused on the potential of audio to create imaginative and ludic possibilities.

Previous work most relevant to our case study includes museum guide systems that utilize an adaptive approach, GUI and PDA interfaces in museum guides, and a discussion of work outside of the museum domain that utilizes audio interfaces in ubiquitous and mobile computing contexts. Equally important to our discussion are the ludic qualities of tangible user interfaces, and related ideas of aesthetics and play in interaction.

**Adaptive museum guide systems and audio display**

Adaptation and personalization approaches have been successfully applied to museums in the context of the World Wide Web [6, 25] and in handheld museum guides. ec(h)o shares many adaptive characteristics with the handheld systems of HyperAudio, HIPS and Hippie [3, 24, 27]. Similar to ec(h)o, the systems respond to user’s location and explicit user actions through the interface. All systems adapt content to the user model, location and interaction history. Among the differences relevant to this paper is that these systems depend on a PDA graphical user interface, ec(h)o uses audio display as the only delivery channel and a tangible object as an input device, ec(h)o treats user interests as dynamic, we look to evolving interaction history. Among the differences relevant to this paper is that these systems depend on a PDA graphical user interface, ec(h)o uses audio display as the only delivery channel and a tangible object as an input device, ec(h)o treats user interests as dynamic, we look to evolving interaction history. Among the differences relevant to this paper is that these systems depend on a PDA graphical user interface, ec(h)o uses audio display as the only delivery channel and a tangible object as an input device, ec(h)o treats user interests as dynamic, we look to evolving interaction history.

In museum guide systems there has been a strong trajectory of use of the PDA graphical user interface. Typically, hypertext is combined with images, video and audio [1, 30]. Yet, a PDA is essentially a productivity tool for business, not a device that lends itself easily to playful interaction. Nevertheless, museum systems are typically PDA-based despite shifts in other domains to approaches that better address the experience design issues most prominent in social, cultural and leisure activities [20]. The play constraints of these devices are too great for the level of interaction that goes beyond playing a software game on a mobile device. For example, in the area of games and ubiquitous computing, Björk and his colleagues have identified the need to develop past end-user devices such as mobile phones, personal digital assistants and game consoles [4]. They argue that we need to better understand how “computational services” augment games situated in real environments. The same can be said for museum visits.

Non-visual interfaces, particularly audio display interfaces have been shown to be effective in improving interaction and integration within existing physical contexts. For example, Brewster and Pirhonen [7, 29] have explored the combination of gesture and audio display that allows for complicated interaction with mobile devices. The Audio Aura project [21] explores how to better connect human activity in the physical world with virtual information through use of audio display. Audio is seen as an immersive display that can enrich the physical world and human activity while remaining integrated with the surrounding environment. In addition, audio tends to create interpretive space or room for imagination as many have claimed radio affords over television. In the HIPS project, different voices and delivery styles were used to create an “empathetic effect” between the user and the artifacts they engaged [20]. We’ve adopted a similar approach to our use of audio content.

**The play of tangible user interfaces**

Tangible user interfaces have a strong potential to be playful, imaginative and poetic user interfaces. Ishii’s and Ullmer’s notion of ‘coupling bits and atoms’ was informed by earlier work in graspable interfaces [11] and real-world interface props [15]. ec(h)o’s tangible user interface draws on this notion by coupling an everyday and graspable object, a wooden cube with digital navigation and information. Ishii was inspired by the aesthetics and rich affordances of scientific instruments [17] and the transparency of a well-worn ping-pong paddle [18]. Simple physical display devices and wooden puzzles at the natural history museum where we conducted ethnography sessions inspired us as well.

In 1992, Bishop’s Marble Answering Machine [8] was an early embodiment of the immediate and playful qualities of tangible user interfaces. The prototype uses marbles to represent messages on the machine. A person replays the message by picking up the marble and placing it in an indentation in the machine. Jeremijenko’s Live Wire is a strikingly minimal and whimsically simple demonstration of digital bits transformed into physical atoms [33]. Jeremijenko dangled a plastic wire from a motor attached to the ceiling. The motor accelerates or decelerates based on traffic across the Ethernet network. Ishii’s PingPongPlus [18] explores the intertwining of athletic play with imaginative play. The ping-pong table becomes an interactive surface, the ball movement is tracked and projections on the table of water ripples, moving spots, and schools of fish react to where the ball hits the table.

**Aesthetics of interaction**

Researchers in human-computer interaction (HCI) have recently explored enjoyment [5] and ludic design [13] in interface approaches. Nowhere is this need more evident than in the richly interpretive and social environments of museums [12, 20]. Our emphasis is on the qualities of interaction that result in play that facilitates discovery. While we address this on an informational level in regard to our use of audio content and information retrieval [14], we found we equally explored the embodied and situated aspects of interaction or aesthetic interaction as expressed by Djajadinigrat [10] and Petersen [26].

Djajadinigrat argues for a “perceptual-motor-centered” approach to tangible interfaces [10]. He argues for a “direct approach” for its “sensory richness and action-potential” of
the objects to carry meaning through interaction. He describes this notion of meaning in interaction as *aesthetics of interaction* whereby the “beauty of interaction” as opposed to the beauty of the artifact or interface, tempts the user to engage as well as “persevere” in their engagement [10].

Petersen and her colleagues’ description of aesthetic interaction shares the embodied aspects described above as well as the sense of aesthetic potential that is realized through the action or engagement [26]. For example, Petersen developed a playful interaction approach as part of the WorkSPACE project utilizing a ball that is thrown against a floor projection of documents and work materials as a way of manipulating and exploring the information. Inherent to the ball are kinesthetic challenges, affordances and the situated relationship with the environment. These aspects are realized in action with the object including the playful actions of aiming, throwing and bouncing.

**DESIGN MOTIVATIONS**

Historically, links have been established between play and learning. For example Dewey argued for the *construction* of knowledge based on learning dependant on action [9]. Piaget, through his child development theory described the development of cognitive structures through action and spontaneous play [28]. More recently, Malone and Lepper consider games as intrinsic motivators for learning [19]. In the museum context, Bell’s *cultural ecologies* [2] ascribe the museum visit with qualities of *liminality* (a space and time set apart from everyday life) and *engagement* (where visitors interact to both learn and play). Yet for designers, the specific requirements and attributes of play described above are elusive and too general. We utilized design ethnography informed by Bell’s *cultural ecologies* [2] and Nardi and O’Day’s *information ecology* [23] as a means of defining play for the design of our prototype.

This approach led us to be inspired by simple physical displays, puzzles, and connections between key people and the artifacts that we observed in our ethnographic sessions. Key observations included:

- Highly tactile and hands-on approaches to artifacts and displays including holding, and manipulating;
- Visitors are not spoon-fed factual information in the form of didactics, rather they engage in play and learning through puzzles, games and physical displays;
- Lively storytelling of the museum staff and researchers about the museum collection was often humorous as well as informative.

Our team spent over seventy hours conducting interviews, video walkthroughs, and site visits with over thirty researchers, staff, and administrative staff at the museum. We also observed exhibitions and museum visitors and conducted an analysis of interaction devices in the museum.

**Museums as ecologies**

Bell sees the museum visit as a ritual determined by space, people and design [2]. She decomposes the visiting ritual into three observational categories: space, visitors, and interactions and rituals. Different types of museums have different ecologies, for example Bell describes different attributes between art and science museums. These ecologies are seen to be distinct and supportive of very different kinds of museum visits. Bell also describes interaction concepts that are common to all museum ecologies. We have drawn on two of these concepts in developing our approach, *liminality* and *engagement*.

- **Liminality** defines museums as places that embody an experience apart from everyday life. Positive museum experiences are transformative, spiritual, and even moving. A museum visitor should be inclined to pause and reflect, thus liminality can be seen to permit a deeper engagement.
- **Engagement** is a key concept for museums as people go to museums to learn, however this engagement is often packaged in an entertaining way; museums are a balance between learning and entertainment spaces.

Nardi and O’Day draw on activity theory [22, 31] and field studies to develop their concept of *information ecologies*. The concept they describe strives for a systematic view of organizations based on the relationships among people, practices, technology, values and locale. For example, a library is an ecology for accessing information. It is a space with books, magazines, tapes, films, computers, databases and librarians *organically* organized to find information. Nardi and O’Day utilize the concept of ecology in order to depict the complex relationship among elements and influences of which technology is only one part. Constituent elements of information ecologies include a system, diversity, co-evolution, locality, and *keystone species*. Two of these elements were essential in supporting our design:

**Locality** can be described as participants within the ecology giving identity and a place for things. For example, the *habitation* of technology provides us with a set of relationships within the ecology, to whom a machine belongs determines the family of relationships connected to the technology. In addition, we all have special knowledge about our own local ecologies that is inaccessible to anyone outside thus giving us local influence on change.

**Keystone species** are present in healthy ecologies; their presence is critical to the survival of the ecology itself. Often such species take the role of mediators who bridge institutional boundaries and translate across disciplines.

**Design implications of our design ethnography**

Our observations that fall within Bell’s categorization of *interaction and ritual* emphasized that our system should be open to multiple forms of input such as movement and
physical interaction with the displays, and responsive to different learning styles.

The displays and installations revealed diverse forms of interaction: microscopes with adjustable slide wheels that could be turned to explore different specimens; wooden puzzles which, once completed, would fall apart at the pull of a handle, creating a loud crashing sound that captured the attention of others (Figure 1); a collecting game called The Rat Pack Challenge which tasked visitors to search the room and discern collectable artifacts from non-collectable ones; discovery drawers filled with objects like fossils, fur pelts, and minerals which visitors could touch and inspect at close range (Figure 1); magazines, coloring books, and a small library of natural history artifacts that were lent to students.

Bell notes that an attribute of science museum ecologies is to support the fact that people learn in a variety of ways. Alternative approaches to learning turned up throughout our observations, such as the interactive puzzles, quizzes, and games that require visitors to explore and think about the artifacts on display.

Stories and information we heard in our interactions with museum staff and researchers were examples of the ecology concepts, locality and keystone species as defined in Nardi’s and O’Day’s information ecology. We observed numerous informal yet engaging stories that communicated the specialized knowledge of the researchers. These were first hand accounts and discussed in a wide-ranging manner. Factual or thesis driven accounts of artifacts were mixed with anecdotal and humorous tales related to the discovery, processing or research of the actual artifact. This experience deeply struck us since our shared perception of the public exhibition display space was quite the opposite. Not unlike many exhibitions, the artifacts and contextualizing information appeared static and lifeless, the puzzles and games notwithstanding. In locality terms, it was evident to us that once the artifacts were connected to people, the understanding of these artifacts became deeply connected to all aspects of the ecology and came out in the form of storytelling that covered activities related to the artifact, conservation, storage, research and display technologies, meaning and values associated with the artifacts.

**Design implications:** As a result we felt the need to bring this degree of liveliness to the artifacts on display. We modeled our content delivery and audio experience on the informal and humorous storytelling we had experienced, extending it through riddles and word play. We aimed to create a virtual cocktail party of natural history scientists that accompanied the visitor through the museum.

**Prototype**

In order to better understand the prototype we tested, we provide a typical visitor scenario and describe the system. The prototype is composed of a tangible user interface, spatial audio display and an integrated user modeling technique combined with semantic technologies. While arguably other interface approaches could have been utilized, such as a simple push-button device for input or a mobile text display device for output, such a strategy would be incongruent with our experience design goals.

**Visitor scenario**

The scenario refers to an exhibition about the history and practice of collecting natural history artifacts:

Visitors ec(h)ose selected topics related to the exhibition to establish their interests for the system. An attendant gives the visitor a wooden cube that has three colored sides, a rounded bottom for resting on her palm and a wrist leash so the cube can hang from her wrist, and headphones connected to a small, light pouch to be slung over her shoulder. The pouch contains a wireless receiver for audio and a digital tag for position tracking.

Our visitor moves through the exhibition space. Her movement creates her own dynamic soundscape of ambient sounds. She passes a collection of animal bones and hears sounds that suggest the animal’s habitat. The immersive ambient sounds provide an audio context for the collection of objects nearby.

As she comes closer to a display exhibiting several artifacts from an archaeological site of the Sigit people, the soundscape fades quietly and the visitor is presented with three audio prefaces in sequence. The first is heard on her left side in a female voice that is jokingly chastising: “Don’t chew on that bone!” This is followed by a brief pause and then a second preface is heard to her center in a young male voice that excitedly exclaims: “Talk about a varied diet!” Lastly, a third preface is heard on her right side in a matter-of-fact young female voice: “First
dump...then organize.” The audio prefaces are like teasers that correspond to audio objects of greater informational depth.

The visitor chooses the audio preface she heard on her left side in order to learn more about it, by holding up the wooden cube and rotating it to the left. This gesture selects and activates an audio object. She hears a chime confirming the selection. The audio object delivered in the same female voice as the related preface yet in a relaxed tone. It describes the degree of tool making on the part of the Siglit people: “Artifact #13 speaks to the active tool making. Here you can actually see the marks from the knives where the bone has been cut. Other indicators include chew marks...”

After listening to the audio object, the visitor is presented with a new and related audio preface to her left side, and the same prefaces that she did not choose earlier, are heard again to her center and right side. The audio prefaces and objects presented are selected by the system based on the visitor’s movements in the exhibition space, previous audio objects selected, and her current topic preferences.

**Tangible user interface**

The tangible user interface is a shaped wooden cube with three adjacent colored sides (Figure 2). The visitor makes a selection by holding the cube in front of them and rotating it. The cube was carefully designed to ensure proper orientation. The “bottom” of the cube has a convex curve to fit comfortably in the palm of the visitor’s hand and a wrist leash is attached to an adjacent side to the curved bottom suggesting the default position of being upright in the palm and at a specified orientation to the visitor’s body. The leash allows visitors to dangle the cube when not in use and frees the use of their hand. The opposite side of the bottom of the cube is colored and when this side is held up the audio preface to the visitor’s center is selected (additional support is provided by an icon denoting a pair of headphones with both channels active printed on this side).

If the visitor rotates the cube to the left or right, the audio prefaces on each side are selected. The sides to the left and right are each uniquely colored and display icons showing active left and right channels of the headphones, respectively. The cube is made of balsa wood and is therefore very light (approximately 100 grams or 3.5 ounces), less than a typical networked PDA.

The ergonomic design of the cube and biomechanics of arm and wrist movement form a physical constraint that ensures that the selected cube face is almost always held parallel to the camera lens above and so highly readable. We experienced no difficulties with this approach.

Technically, the input of the selection is done through video sensing. We used the “eyeclick” vision system (http://www.squishedeyeballs.com), which included an array of color video cameras connected to a desktop computer in order to cover specified interactive zones. Each interactive zone included a single camera positioned on the ceiling above in order to detect the rotation of the cube by visitors.

**Audio display**

For the prefaces and audio objects we used a simple spatial audio structure in order to cognitively differentiate between the options heard. Switching between the stereo channels created localization: we used the left channel audio for the left, right channel audio for the right, and both channels for the center. It is an egocentric [9] spatial structure that allowed the three prefaces to be distinguishable and an underlying content categorization structure to exist. The spatialization was mapped to the tangible interface for selection. In addition, we provided simple chimes to confirm that a selection had been made.

The prefaces were written to create a sense of surprise, discovery and above all play, especially in contrast to the informational audio objects. In order to create this sense we utilized diverse forms of puns, riddles and word play, for example:

- **Ambiguous word play:** “Sea urchins for sand dollars” (preface); “Other then the morphology, the sea urchin and the sand dollar are very similar species...” (abridged audio object);
- **Simple pun:** “It’s like putting your foot in your mouth” (preface); “The word gastropod comes from two different roots: *gaster* for stomach, and *pod* for foot” (audio object);
- **Literary pun:** “Dung beetles play ball!” (preface); “Dung beetles turn dung into balls and are equipped with their forehead and legs to push these balls for some distance...” (abridged audio object);
- **Turn of phrase:** “An inch or two give or take a foot” (preface); “Dung beetle nests are usually underground, and can range from a few inches to a few feet deep” (audio object);
- **Definition pun:** “There’s a cat in the garden!” (preface); “Specimen #129 is a John Macoun sample, it is known as a pussy toe because the plant flower and fruit represent a cat’s foot” (audio object);
- **Riddles:** “What is always naked and thinks on its feet?” (preface); “Where gastropods are shielded critters with...
through the ambient sounds that are dynamically created. If navigates the exhibit exploring it on a thematic level animals such as the flapping of crane’s wings. The visitor display. The concepts are translated into environmental sounds such as the sound of an animal habitat, and sound of zones and mapped concepts of interest to each zone and (figure 3). We divided the exhibition space into interactive displays, and a multi-channel editor, mixer and server in the Max/MSP™ environment that functioned as an audio engine. This engine created dynamic soundscapes and delivered unique channels of stereo audio to individual users over FM transmitters. Each visitor carried a small inexpensive digital receiver in a pouch. We produced over 600 reusable and annotated audio objects. The average length of an audio object is approximately 15 seconds. The prefaces typically are 3 seconds in duration.

Navigation
We structured navigation at a macro level, where visitors move throughout the exhibition space in between artifact displays, and a micro level, where visitors are within a specified interactive zone in close proximity to an artifacts display.

On the macro level the input is the visitor’s movement, which is tracked using a combined Radio Frequency Identification (RFID) and optical position tracking system developed by Precision Systems (http://www.precision-sys.com). The movement triggers an ambient soundscape that is made of sounds related to artifacts near the visitor (figure 3). We divided the exhibition space into interactive zones and mapped concepts of interest to each zone and display. The concepts are translated into environmental sounds such as the sound of an animal habitat, and sound of animals such as the flapping of crane’s wings. The visitor navigates the exhibit exploring it on a thematic level through the ambient sounds that are dynamically created.

The audio recordings of the prefaces and audio objects used a diverse set of voices that were playful in tonality and style. This added to the conversational feel and created an imaginary scene of a virtual cocktail party of natural historians and scientists that followed you through the museum. As we discussed above, we identified natural history scientists as our key stone species. We organized sessions of recorded walkthroughs of the exhibition asking each scientist to provide commentary [32]. These sessions became the basis for the discrete audio objects that were categorized by topics and relationship to artifacts on display.

A set of concepts strongly matches the visitor’s interest the related audio is acoustically more prominent.

On the micro level, when visitors are within a meter or more of an artifacts display. The navigation is as follows: as previously discussed, a visitor is played three prefaces, one to his left, another to his center and the third to his right (figure 4). He selects the preface on his right side by rotating the tangible object, and listens to the linked audio object. On the subsequent turn the visitor hears the same two prefaces he did not select, and again he hears them to his left and to his center. Since he previously chose the preface to his right he now hears a new preface in that location. If the visitor then selects the center preface, on the subsequent turn only that preface is replaced by a new preface in the center position. If a preface has been replayed three times without being selected, it is replaced by a preface and audio object of the next highest-ranking topic according to the user model.

User model
For an in depth discussion and evaluation (not included in this paper) of the adaptive user model approach in ec(h)o we refer readers to [14]. Our approach can be summarized as the use of a modeling technique supported by ontologies and rules for information retrieval.

EVALUATION
We installed the ec(h)o system in an existing exhibition about collecting called ‘Finders and Keepers’ at the Canadian Museum of Nature. We created three interactive zones and a complete soundscape.

We performed two sessions of evaluation, a short session and an in-depth session. We evaluated user experience through observation and video analysis. We added questionnaires and semi-structured interviews for the in-depth sessions.

<table>
<thead>
<tr>
<th>Turn</th>
<th>Prefaces played</th>
<th>Preface audio object selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1  1  1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1  1  4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>1  1  4</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 4. Illustration of micro level navigation
The physicality of objects interacts with our bodies in which a myriad of interaction paths coexist. Actions and skill development; and in freedom of interaction unique and varying kinesthetic combinations in which richness of motor actions found in the potential space of rhythm, and flow between the user and the object; the of timing, intercl!>t1>n pallern an aesthetic interaction: the action-potential of physical objects [13].

Djajadiningrat describes three factors as having a role in aesthetic interaction: the interaction pattern of timing, rhythm, and flow between the user and the object; the richness of motor actions found in the potential space of actions and skill development; and freedom of interaction in which a myriad of interaction paths coexist.

The physicality of objects interacts with our bodies in unique and varying kinesthetic combinations in which optimal efficiency gives way to play and experimentation. In simple actions of holding and rotating the cube, we observed a diverse set of interaction techniques when selecting prefetches. We identified at least five basic types all of which successfully operated the system:

- **Hold and rotate**, one hand holds the cube resting on the palm while the other hand rotates it in place (see Figure 3(a-b));
- **Hold, rotate and cover**, one hand holds the cube resting on the palm while the other hand or both hands rotate the cube. The topside is uncovered until the selection is made and then the topside is covered again until its time to make another selection (see Figure 3(c-d));
- **Cradle and hide**, two hands rotate and cradle the cube, after selection is made the colored side is rotated and hidden against the visitor’s body (see Figure 1(e));
- **Rotate wrist**, one hand holds the cube between fingers and thumb, and rotates the wrists to make a selection (see Figure 3(f-g));
- **Rotate with fingers**, one hand holds the cube and rotates it by rolling with the fingers and thumb (see Figure 3(h)).

It is important to note that we observed combinations and variations of these techniques, as well as individual experimentation with the different approaches. As one might expect, we also observed a range of methods for holding the cube when not selecting prefetches or walking through the exhibition such as cradling it in hands, holding it at one’s side or behind one’s back, dangling it from the wrist, or holding its leash to gently sway it from side to side.

This sense of play extended to participant’s movements through the exhibition space. In the interviews participants commented on how they returned to zones to see if the system would indeed not repeat audio objects already heard. We observed participants moving from zone to zone straddling the boundaries of the interactive zones and soundscapes appearing to experiment with their movements and the soundscapes created by entering and exiting zones.

**Evaluation of user experience**

Our formal evaluation of user experience took place in the in-depth session. In addition to observation, we added a questionnaire, and a semi-structured interview. The questionnaire included sixty-three questions that assessed user experience related to the overall reaction to the system, the user interface, learning how to use the system, perceptions of the system’s performance, the experience of the content, and degree of navigation and control. Majority of the questions in the questionnaire were on a Likert scale yet it provided for open-ended written comments. For a summary of the questionnaire results see Table 1.

Overall, participants found the system enjoyable and stimulating, perhaps in part due to its novelty. The general sense of satisfaction was split between those participants
who liked the playful approach and those who did not. We noted a clear age difference in that the "younger" participants rated satisfaction higher based on their liking of the playful approach (this was confirmed in the semi-structured interviews).

Among the factors that stood out as most positive for the participants was that the cube and audio delivery were seen as playful. The tangible user interface was well received especially in terms of ergonomics and ease of use. This was not a surprise to us since our early testing and participatory design sessions provided us with considerable feedback, especially on ease of use and enjoyment. We went through several iterations and form factors of the wooden cube and tested it against different hand sizes. This may have also resulted in the fact that learning to use the interface and navigation were rated highly and participants felt the system had a low learning curve:

Umm, I found it was really easy. Sometimes I got so engaged in listening to what they were saying that I forgot in which orientation I was holding the cube. And I found that I would have to occasionally look down. But the way it was designed with the round part to go in your palm... it was really easy to quickly reorient myself to how I was holding that cube. (Participant 5)

Interestingly, the audio content was perceived to be both accurate and clear. The issue of trust and delivery style is an area to further investigate.

The questionnaire did point out challenges and areas for further research. Some things we expected such as the headphones were uncomfortable, yet to such a degree that we are currently rethinking the tradeoff between personalized spatial audio and use of headphones. Other results point to a threshold in the balance between levels of abstraction and local information. Since visitors had difficulties at time connecting what they were listening to and what was in front of them (in part this was an inherent challenge in the exhibition since the display cases had dozens to over a hundred artifacts). In addition, we see both a threshold point in play versus focused attention on the exhibit in that the question relating to the content asking if it was "distractive-synergistic" scored 2.83. This raises the issue of balance in play and the possibility to shift attention away from the environment rather than play as a means of further exploring the environment.

It was often remarked how the experience was similar to a game:

The whole system to me felt a lot like a game. I mean I got lost in it. I found myself spending a lot more time in a particular area than I normally would. And just the challenge of waiting to hear what was next, what the little choice of three was going to be. Yeah... So I found it over all engaging, it was fun, and it was very game-like. (Participant 4)

The playfulness did in most instances suggest a quality of engagement that led to learning even through diverse types of museum visits:

I learned a lot and well you know I’m a scientist here, and I think anybody going through, even people who are in a real rush, are going to pick up some interesting facts going through. And... I mean, that was good, the text was great and was short enough that somebody in a rush is still going to catch the whole thing. (Participant 1)

However, we feel a future study in which we compared experiences with and without ch(h)o against visitors’ previous knowledge would better measure the degree of learning through use of the system.

As mentioned earlier, there is a threshold between play in support of the exhibit on display and play with the system that can be an end in itself and even a distraction. For example, one user’s enthusiasm for the game-like quality led her to at times pay more attention to the interaction with the system than the exhibition. One participant would have preferred a more serious and “non-playful” approach. In addition, participants observations on the liminality of the experience manifested in comments suggesting that play was more natural for children rather than themselves, however as expressed they soon overcame this issue:

At first it felt a little bit strange, especially holding this cube that looked like a children’s toy, and I felt a little bit awkward about doing that, but I got over that pretty quickly. (Participant 5)

### Table 1 Summary of the questionnaire results (n=6; 63 questions on Likert scale of 1-5 (5 being best)

<table>
<thead>
<tr>
<th>Categories</th>
<th>Avg</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall reaction (5 questions including “terrible-wonderful; difficult-easy”)</td>
<td>3.60</td>
<td>0.78</td>
</tr>
<tr>
<td>Tangible user interface (7 questions including “uncomfortable-comfortable; difficult-easy to manipulate; annoying-enjoyable”)</td>
<td>4.24</td>
<td>0.50</td>
</tr>
<tr>
<td>Headset (2 questions including “comfortable-uncomfortable to wear”</td>
<td>2.92</td>
<td>0.12</td>
</tr>
<tr>
<td>Learning curve for the system (8 questions including “difficult-easy to get started; risky-safe to explore features; unclear-clear feedback”)</td>
<td>4.07</td>
<td>0.36</td>
</tr>
<tr>
<td>Perception of system performance (8 questions including “slow-fast system response; never-always reliable”)</td>
<td>3.83</td>
<td>0.39</td>
</tr>
<tr>
<td>Quality of the content (15 questions including: “uninformative-informative; generalized-customized for me; rigid-playful; predictable-surprising”)</td>
<td>3.78</td>
<td>0.52</td>
</tr>
<tr>
<td>Quality of the audio experience (9 questions including “confusing-clear; mechanical-human-like; wasteful-valuable”)</td>
<td>3.67</td>
<td>0.30</td>
</tr>
<tr>
<td>Navigation and control (8 questions including “never-always able to navigate in an efficient way; always-never found myself lost in the system; always-never found myself uncertain of system state”</td>
<td>3.23</td>
<td>0.29</td>
</tr>
</tbody>
</table>
The expert reviews found e(h)o to be a “pleasurable immersion experience consistent with free-choice leisure learning” establishing its “potential as an effective educational media in a museum setting.” The key concerns were the requirement of different audio content for different ages of museum visitors, and occasional “dislocation” of audio content from the artifacts.

DISCUSSION

In this paper we’ve explored play in a tangible and audio museum guide. Our approach in e(h)o was to create a coherent space for play and discovery across all components of the design including reasoning, audio delivery and interface. The space suggests actions and meaning but maintains an openness and interpretation that requires playful interaction on the part of the user in order to realize the action-potential or relevancy of the information. The results of our ethnography and evaluation of our prototype tell us that play is highly contextual. Designers need to consider the situated nature of play for two reasons: (1) to best serve the overall design purpose; (2) in order to understand the nature and degree of play required.

When is a good thing too much? In our results playfulness was identified positively in all aspects of the interface yet overall satisfaction was split between those participants who enjoyed playing and those who did not. Considering who wants to play and who does not is important. The question of too much play is also important. There is a need to find the balance between play in support of the exhibit and play with the system that can be a distraction and even an end in itself. Otherwise designers run the risk of users engrossed in playing with the system at the expense of interacting with their surroundings, as one participant commented happened to her periodically.

What makes sense as play? Playful interaction lends itself well to integrating with the context and in many cases depends on it, as in bouncing a ball off the floor or wall. Yet attempting to bounce our cube off the wall of the museum is not sensible play in a museum. We persuasively designed the cube to be held and not thrown through its shape, tactility and the leash. Moving the cube in one’s hand or dangling it from the wrist is seen as playful in our museum context – in another context this would be too passive a set of actions to be seen as play. The puns, riddles and word play we used were also contextual. The visitor required the visual artifacts to support the word play and to make sense of the humor. Culture and language make this type of play not for everyone.

What forms of play? Our interface integrated two forms of play: (1) content play such as puns and riddles in informational content and used in the audio display; (2) physical play that in our case consisted of holding, touching and movement exploring the range and richness of the interaction.

These observations connect us back to the wooden puzzle that fascinated us in our ethnographic study. The puzzle incorporated the forms of content and physical play. It was well balanced with its context providing sensible and engaging play that did not overwhelm its purpose of teaching about collecting natural history artifacts.

CONCLUSION

e(h)o is an augmented audio reality system for museum visitors that utilizes a tangible interface. We developed and tested the prototype for Canadian Museum of Nature in Ottawa. In e(h)o we tested the feasibility of audio display and a tangible user interface for ubiquitous computing systems – one that encourages an experience of play and engagement. In this paper we have presented relevant work in the domains of adaptive museum guides and audio displays, ludic approaches to tangible user interfaces, and aesthetic interaction. We provided an overview of our design motivations rooted in ethnography that led to our approaches in audio delivery and tangible interface. We described the components of our prototype and we described our implementation and evaluation.

We found that we could consider play in two main forms in regard to the interface: content and physical play. We also found that play is highly contextual. Designers need to consider the situated nature of play for two reasons: (1) to best serve the overall design purpose; (2) in order to understand the nature and degree of play required.

ACKNOWLEDGEMENTS

e(h)o was funded by Canarie Inc. We thank the Canadian Museum of Nature in Ottawa for their enthusiastic support of this project. We also thank our many colleagues and students for their contributions.

REFERENCES


Sound Intensity Gradients in an Ambient Intelligence Audio Display

Milena Droumeva
School of Interactive Arts and Technologies,
Simon Fraser University,
Surrey, B.C. Canada
mvdroume@sfu.ca

Ron Wakkary
Associate Professor
School of Interactive Arts and Technologies,
Simon Fraser University,
Surrey, B.C. Canada
rwakkary@sfu.ca

Abstract
This paper describes the prototype of a real-time responsive audio display for an ambient intelligent game named socio-ec(h)o. The audio display relies on a gradient response to represent and anticipate player action. We describe the audio display schema, and discuss results of our current experimentation in guiding player actions through types of audio feedback, for creating sound recognition, perceptions of change and sound intensity.

Keywords
Ambient intelligence, auditory display, game, play, responsive environments, sound design

ACM Classification Keywords
J.5 [Arts and Humanities] H.5.2 [information Systems]: interfaces and presentations H.5.5 [Sound and Music Computing]: Systems

Introduction
Our current research involves the design of an ambient intelligence (AmI) environment for a multi-user physical game known as socio-ec(h)o. Our overall goal is to explore approaches to group communication and skill acquisition in AmI environments. AmI requires meaningful and clear yet ambient feedback to users.
We believe that gradient responses can provide this type of feedback. We have developed an audio display schema, and experimented with gradient response and intensities in the audio display in order to confirm and anticipate users' actions.

**Background**

Sound is an important channel through which humans perceive their natural or designed environments. There is growing literature on auditory perception and design for task-oriented, highly computerized environments [1, 2, 10], virtual audio and spatialization [3] and sonification of information clusters [1, 6, 13], yet relatively few of these approaches to sound design have found their way in the design of sound for games [5, 9]. As a game platform, socio-ec(h)o draws on a variety of existing sound design patterns [1], as well as classical psychoacoustics, investigations into perception of environmental sound, and spatialization [8]. Our sonic design benefited also from the acoustic communication approach developed by Schafer and Truax [11], where sound is seen as a ubiquitous communication channel between listener, soundscape and a physical (and cultural) environment.

**Design approach**

Ambient intelligence is the embedding of computer technologies and sensors in physical environments that combined with artificial intelligence, respond to and reason about human actions and behaviours within the environment. The socio-ec(h)o game is played by a team of four players and features six levels of increasingly difficult word puzzles solved by coordinated body positions and movements. The environment is responsive to players' actions through abstract light and sound (figure 1). Players' movements are tracked using a Vicon motion capture system (www.vicon.com). A custom reasoning engine was developed to track the game state and infer players' actions. For more technical details of our prototype see [12].

Our goal was to create an ambient feel and function that is in keeping with AmI and supports group interaction. We felt that the interface required a gradient rather than direct response to players in order to best represent game states and anticipate players' actions. This approach poses a challenge for the audio displays. Together with the visual display, the audio display must create the aesthetic feel, represent current game state, and guide future player actions through abstract and ambient real-time sonification. After a number of design workshops, we arrived at an approach for the audio display that provides a continuous, ambient response along a gradient.

In its simplest form, the approach can be described as dynamic soundscapes that are not only recognized by players, but in turn interpreted to determine the state of the game (figure 2). The soundscapes are made of sounds that can be categorized as musical, abstract and environmental. One question is what combination of sound categories would be recognized and perceived as meaningful in the context of the game?

The meaningful sounds aid an understanding of the game state and the affect of players' actions. We felt we could modulate this understanding by increasing or decreasing the rate of intensity of or changes in the sound along a gradient. An example of this is the game of "hot and cold", where players use words along the continuum of hot to cold to signify the proximity of another player to a solution. The question is what
sound techniques would work best to signal increased intensity and change along a gradient?

Audio display schema
The audio display consists of three components: a real-time ambient sonification that has a different soundscape for each level of the game; an anticipatory feedback sound to signal when all participants are working together towards the goal; and a confirmatory feedback sound, which signals the completion of the goal and the progress to the next level. The latter two feedback signals are consistent through all levels, so as to create a coherent expectation of success. These responses work in conjunction with each other. The gradient intensity is applied to the real-time ambient sonification component and builds up to the two types of feedback.

For example, players in socio-ec(h)o must achieve a specific goal at each level, which is a coordinated configuration of movements and body positions. For example, an answer to the word puzzle "lo and behold" was for all players to crouch low and be still. Further, they must hold this configuration for a short period of time, typically 4-6 seconds. This ensures that the actions were purposeful and knowingly completed rather than an accidental formation. In figure 3 we illustrate players' actions and the mapping of intensity levels and the audio display components. Reading from left to right, the players explore different combinations of movements and actions. As some players discover a combination, the intensity of the real-time sonification increases, in this case to a value of 2 out of a maximum of 4. As more players discover the solution, the intensity level briefly drops but then increases to the maximum level of 4. At this point, an anticipatory feedback sound occurs signaling to the players that they've done something right and if they keep at it they will complete the solution. However, we see that the players are unable to hold the combination and the intensity level rapidly decays to 1 (in our observations players often stop after an increase in intensity or anticipatory feedback in order consider what they just did). The players quickly resume the configuration and the anticipatory feedback sound occurs again. This time the players complete the solution by maintaining the configuration for the required duration, and the confirmatory feedback sound occurs. The real-time sonification fades away then back as it transitions to the next level. Here, a new soundscape is composed and the intensity level is reset to a minimum.

Current progress
To date we have developed the socio-ec(h)o prototype and have implemented preliminary user testing. This includes two three-hour sessions with eight participants, and an additional two-hour session with four other participants. The group included three
females and nine males ranging in ages from twenty-one to fifty-nine. Each session began with a warm-up to introduce the concept of puzzles solved through physical action and support through implicit responses. Each team of four played two levels followed by discussions. After all levels were completed or a total of two hours of interaction (60 minutes in the short session), the game was stopped and an open-ended interview was conducted.

Results
We've organized our results by first discussing what combination of sound categories and audio processing techniques created recognition, followed by a discussion of techniques that best achieved an intensity of sound, gradient of change, and feedback. While visual display was part of the interface we gave greater attention to the role of audio.

Sound Category
The first two game levels featured an abstract musical soundscape. Level 3 had an abstract non-musical sound (crackling of rocks), and the last three levels had environmental soundscapes. Level 4 and 5 were represented by fire and water respectively, while level 6 featured a more complex soundscape of a forest that gradually turned from a calm night to a full-fledged storm. The socio-ec(h)o's soundscapes evolved from one-dimensional, internally coherent sounds, to complex ecologies. In the course of our testing, it became clear that the more abstract sounds (such as a crackling of rocks, or highly processed musical abstractions) were hard for participants to recognize. As a result, these sounds posed difficulties in determining the level of change in the gradient response. In the discussion after levels 2 and 3, players commented that they could tell the sound display was intensifying, but found it hard to tell "by how much." In contrast, after levels 4 and 5 (featuring water and fire sounds) players reported a noticeable positive difference in the "sharpness" of the ambient response. The environmental sounds were more recognizable, and players had an easier time interpreting changes in them. Players also found environmental sound more contextualized and richer in narrative. As one participant commented: "the sound of fire didn't just get louder, it was different, more intense, and I automatically associate it with a positive thing, with success - building a big fire!" It seems, if players have prior experience with a sound, they have a sense of its inherent scale of intensity, i.e. they know what full-blown fire sounds like, as opposed to a faint crackle. This accounts for the difficulty with unfamiliar abstract sound. The players lacked a sense of scale and therefore found it hard to interpret the gradient of change.

Perception of intensity
socio-ec(h)o uses sound that is coherent in its basic characteristics of pitch, rhythm and timbre within the same game level, and represents success rate in the game by intensifying these basic sound characteristics. In Table 1 we detail the range of audio design approaches we experimented with to achieve gradient sound display in each game level.

The Layer fader is a simple cross-fader between 5 layers of sound which could be specifically chosen or pre-composed to represent increased intensity of the same group of sounds (i.e. dripping water gradually changing into a fast river stream, over 5 steps). The Low-pass filter affects the perception of intensity.
through adding and subtracting frequencies from the sound's spectrum. The result is a sense of brightness or muffled sound that is quite recognizable. The Phaser approach relies on running a select sound source through a simple phaser, which varies the sound's phase - i.e. its perceived tempo heard as a pulsation. This approach was made even more complex with culturally associative environmental sounds such as water. Intensity here was related inversely to tempo - shifting from a pulsating beat to full, continuous sound at the completion of the goal state.

As research in classical and contemporary psychoacoustics suggest [8], amplitude change, followed by pitch change and tempo change are the most readily and easily perceived sound variance characteristics. Thus game levels are based on these approaches (and combinations thereof) in order of their “ease of perception”. For example, the bi-quad filter's effect was effective as an approach to gradual change

<table>
<thead>
<tr>
<th>Level</th>
<th>Method</th>
<th>Description of Effect</th>
<th>Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amplitude</td>
<td>Sound levels slowly come up</td>
<td>Volume</td>
</tr>
<tr>
<td>2</td>
<td>Phaser Layer fader</td>
<td>Tempo goes up, crossfade of 5 sounds</td>
<td>Tempo timbre</td>
</tr>
<tr>
<td>3</td>
<td>Layer fader Pitch shift</td>
<td>Crossfade of 5 sounds</td>
<td>Timbre Pitch</td>
</tr>
<tr>
<td>4</td>
<td>Layer fader Low-pass filter</td>
<td>Sound crossfades and muffled-to-bright</td>
<td>Timbre Association</td>
</tr>
<tr>
<td>5</td>
<td>Phaser Pitch shift</td>
<td>Sound goes up in tempo and pitch</td>
<td>Tempo Pitch</td>
</tr>
<tr>
<td>6</td>
<td>Layer fader</td>
<td>Crossfade of 5 sounds</td>
<td>Timbre Association</td>
</tr>
</tbody>
</table>

our design workshops, yet didn't work as well in preliminary testing. We believe the reason for this is the abstract nature of the sound category used with this effect. We still believe low-pass filtering could be effective in representing a gradient, only with less abstract sounds. The phaser effect (slowing down or speeding up of a pulsating beat) produced a perception of change, yet participants reported it hard to the gauge the gradient scale. When prompted during discussion, participants suggested pitch shifting from low to high as a clear way to represent rising intensity. This reinforced one of the approach we had already incorporated in the system.

Perception of change

8-channel spatialization was used to create a perception of change in the soundscape. This was done by gradually moving the sound in space, the audio display cross-faded sounds of different intensity in and out, giving the players the impression that sound is “going away” or “getting closer” to their relative position in space. Participants commented positively on the sound's spatial movements, saying it made the feedback appear “more crisp” and it helped them interpret change easier. Participants responded the most positively to a combination of environmental sound (whether intensifying fire or going deeper into the forest) and a multi-channel diffusion. All participants commented positively on the immersion quality of the play space.

Anticipatory and Confirmatory Feedback

An issue that became clear from participants' comments was the importance of the anticipatory and confirmatory feedback sounds. As players put it, they wanted to know when they were “on the right track,”
but they also wanted an indication of when they “got it,” in order to “celebrate after all the hard work.” Recognition of the sound, as well as overcoming masking, i.e. sound has to be heard clearly over the ambience were key issues. The sounds were random pitch variation of two soft abstract sounds (granulated tapping of glass). We played the anticipatory and confirmatory feedback sounds to all participants before the game began. We found that prior listening was sufficient in creating recognition of sounds when heard again later in the game.

Conclusion and future work
In this paper we have outlined a model of sound design that uses an intensity gradient approach to signifying system and game information to players. We have presented a design approach to interactive sound that we feel is a good fit with AmI systems.

AmI environments rely heavily on a meaningful and clear, yet ambient response. In socio-ec(h)o we introduce a new concept of providing intensity-based audio display with a gradient approach. In this respect we were encouraged by the results of our preliminary tests. Even when the display was less clear, participants were still able to use the audio as a reliable feedback system throughout the game. In addition, specific approaches, such as using environmental sound and certain audio processing techniques proved very effective in contextualizing the game and providing game state information to players. Future work will include further study of the relationship between the audio processing of different sound categories, and the auditory perception of recognition, intensity and change.

References
An Ambient Intelligence Platform for Physical Play

Ron Wakkary, Marek Hatala, Robb Lovell, Milena Droumeva
Simon Fraser University
2400 Central City
Surrey, British Columbia, Canada, V3T 2W1
604 268 7425
{rwakkary; mhatala; lovell; mvdroume}@sfu.ca

ABSTRACT
This paper describes an ambient intelligent prototype known as socio-ec(h)o. socio-ec(h)o explores the design and implementation of a system for sensing and display, user modeling, and interaction models based on a game structure. The game structure includes word, puzzles, levels, body states, goals and game skills. Body states are body movements and positions that players must discover in order to complete a level and in turn represent a learned game skill. The paper provides an overview of background concepts and related research. We describe the prototype and game structure, provide a technical description of the prototype and discuss technical issues related to sensing, reasoning and display. The paper contributes by providing a method for constructing group parameters from individual parameters with real-time motion capture data; and a model for mapping the trajectory of participant’s actions in order to determine an intensity level used to manage the experience flow of the game and its representation in audio and visual display. We conclude with a discussion of known and outstanding technical issues, and future research.

Categories and Subject Descriptors
J.5 [Arts and Humanities]

General Terms
Documentation, Design, Experimentation, Human Factors

Keywords
Ambient intelligence, responsive environment, user model, physical play, puzzles, embodied, auditory display, motion capture, sound ecology

1. INTRODUCTION
This paper describes the research of an ambient intelligent system known as socio-ec(h)o. socio-ec(h)o explores the design and implementation of an ambient intelligent system for sensing and display, user modeling, and interaction models based on game structures. Ambient intelligence computing is the embedding of computer technologies and sensors in architectural environments that combined with artificial intelligence, respond to and reason about human actions and behaviours within the environment.

Ambient intelligent spaces lend themselves extremely well to physical and group play. In this paper we describe our interaction model and technical prototype. The overall research goal of this project is to understand to what degree physical play and game structures such as puzzles can support groups of participants as they learn to manipulate an ambient intelligent space. To date we have designed and implemented the prototype and interaction model. We have incorporated formative and summative feedback through a participatory design process and preliminary user testing.

The aim of our game is for a team of four players to progress through seven game levels. Each level is completed when the players achieve a certain combination of body movements and positions. At the beginning of each level, players are presented with a word puzzle as a clue in discovering the desired body states. The levels are represented by changes in the environment in light and audio. The levels are progressively more challenging in terms of body states and more complex in terms of the audio and visual ambient display. The physical environment consists of a circumscribed circular space (the area in which we can detect motion), immersive 8-channel audio, theatrical lighting, and two video projection surfaces, see figure 1.

The paper provides an overview of background concepts and related research. We then describe the game structure and prototype; include a technical review of the system, and a discussion of technical issues related to sensing, reasoning and display. We discuss our movement-based interaction and display in the context of aesthetic interaction. We describe how we utilized selective responses that were real-time, gradient, provided rewards, developed a composite model for reasoning on different groups of users, and how we customized a motion capture system for real-time bodily and positional sensing rather than gestures. We conclude with a discussion of known and outstanding issues and future work. Our contributions in this paper include a method for constructing group parameters from individual parameters with real-time motion capture data; and a model for mapping the trajectory of participant’s actions in order to determine an intensity level used to manage the experience flow of the game; and design strategies for representing intensity via an audio and visual display.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM’05, November 6–11, 2005, Singapore.
Copyright 2005 ACM 1-59593-044-2/05/0011...$5.00.
and position was “new and exhilarating.” In addition, playing as a team in an interactive virtual space was found to be engaging, natural and fun.

In relation to the above research, socio-ec(h)o builds on the theatrical, simple and physical interaction models in order to develop a game structure approach that lies between exploratory play and a structured game for adults within an ambient intelligent environment. In addition, we extend the notion of a game structure to an interaction model for the environment rather than a virtual game space. We also build on the idea that action, play and learning are linked in such physically-based environments.

Technically, Nautilus employed a theatrical lighting approach similar to socio-ec(h)o. However, position tracking was done through the use of a sensor floor that tracked movement across x and y coordinates only. In socio-ec(h)o we utilized a motion capture system in order to support more complex actions and locations across three dimensions. Motion capture has primarily been used in contexts where 3D data can be captured for later analysis and re-processing. The technology is used with game construction and movie making to create realistic motions of animated characters. However, recently there is ongoing research into the real-time use of this technology in live performance and interactive arts [3, 18]. Particularly, motion capture was successfully used by part of the research team in a live dance performance titled “Immanence” where a dancer manipulated an animated figure in front of a live audience [14].

Another related topic to our research is user modeling in respect to ubiquitous computing and group models. Our work builds on adaptive and rules-based approaches to context-aware modeling [13, 22]. In addition, our approach has been influenced by previous research in group modeling [12, 17]. Specific to games research, we have utilized Richard Bartle’s model of game types [2].

We have also benefited from the research in soundscape studies and acoustic communication through the works of Schafer and Truax [24, 27, 28], and in the area of cognition and psychoacoustics [5, 7, 19].

The approaches in this project, especially the further development of an ambient intelligent platform, and the use of user modeling and interaction models originated in earlier work by this research team on an ambient intelligent museum guide [11, 29].

Lastly, in respect to the links between action, play and learning there is a substantial amount of literature. Dewey argued for the construction of knowledge based on learning dependant on action [8]. Piaget, through his child development theory believes in the development of cognitive structures through action and spontaneous play [23]. According to Piaget, constructivist learning is rooted in experimentation, discovery and play among other factors. Papert extends Piaget’s notions by investigating the knowledge-construction process that emerges from learners actually creating and designing physical objects [20]. Malone and Lepper consider games as intrinsic motivators for learning [16]. Subjective motivations like challenge, curiosity, control and fantasy may occur in any learning situation; other motivations like competition, cooperation and recognition are considered to be inter-subjective, relying on the presence of other players/learners.

**Figure 1. Above are scenes from a user testing session**

---

2. RELATED WORK

Key contextual issues in socio-ec(h)o stretch across many disciplines and research areas. It would seem that research in ambient intelligence in a social-cultural context is inherently interdisciplinary. The range of related topics include research in the areas of play and ambient intelligent spaces, motion capture systems, user modeling, auditory display, and literature linking play and learning.

Björk and his colleagues have observed progress toward fully ubiquitous computing games yet they identify the need to develop past end-user devices such as mobile phones, personal digital assistants and game consoles. Accordingly, we need to better understand how “computational services” augment games situated in real environments [4]. Interactive art projects such as works by F6am and Sponge have explored social and mixed reality environments incorporating gesture and wearable computing [15, 25].

Recent projects have investigated the play space of responsive environments and tangible computing utilizing sensors, audio, and visual displays. For example, Andersen [1], and Ferris and Bannon [10] engage children in exploratory play and emergent learning through sensor-augmented objects and audio display. Andersen’s work reveals how theatrical settings provide an emotional framework that scaffolds the qualitative experience of the interaction. Ferris and Bannon’s work make clear that a combination of simple feedback and control lead children to widely explore and discover a responsive environment.

In the Nautilus project, Strömberg and her colleagues employ bodily and spatial user interfaces as a way of allowing players to use their natural body movements and to interact with each other in a group game within a virtual game space [26]. Strömberg observed in physical and team games such as soccer or dodge ball that players coordinate their physical movements and rely heavily on communication to be successful. In their findings, participants reported that controlling a game through one’s body movement...
3. GAME STRUCTURE AND PROTOTYPE

Below is a short scenario of participants in the socio-ec(h)o environment:

Madison, Corey, Elias and Trevor have just completed the first level of socio-ec(h)o. They discovered that each of them had to be low to the ground, still, practically on all fours. Once they had done that, the space became bathed in warm yellow light and filled with a wellspring sound of resonating cymbals. Minutes earlier, the space was very dim — almost pitch black until their eyes adjusted. A quiet soundscape of “electronic crickets” enveloped them. They discussed and tried out many possibilities to solving the word puzzle: “Opposites: Lo and behold.” They had circled the space in opposite directions. They stood in pairs on opposing sides of the space. At Corey’s urging, the four grouped together on the edge of the space and systematically sent a player at a time to the opposite side in order to gauge any change in the environment. Nothing changed. Madison, without communicating to anyone realized the obvious clue of “Lo” or “low”. While Corey was in mid-sentence thinking-out-load about the puzzle with Trevor, and trying to direct the group into new body positions, Madison lowered herself to a crouching position. The space immediately glowed red and became brighter. The audio changed into a rising chorus of cymbals — not loud but progressively more pronounced. Corey and Trevor stopped talking and looked around at the changing space. Madison, after a pause began to say “Get down! Get down!” Elias stooped down immediately and the space became even brighter. Corey and Trevor dropped down in unison and the space soon became bathed in a warm yellow light like daylight. The audio reverberated in the space. A loud cheer of recognition came from the group. “Aaaaahhh! We got it!” Corey asked everyone to get up. As soon as they were all standing, the space became pitched black again. They dropped down again and the space was full of light. They had learned how to “create daylight” in the space. They had completed level one.

Soon after, a new word puzzle was presented to them in a short video projected on two scrims hanging from the ceiling: “The opposite of another word for hello but never settles.” The lights have become very dim now and the audio is filled with “electronic crickets” enveloped them. They discussed and tried out many possibilities to solving the word puzzle: “This way slow —low to high.” Madison lowered herself to a crouching position. The space immediately glowed red and became brighter. The audio changed into a rising chorus of cymbals — not loud but progressively more pronounced. Corey and Trevor dropped down in unison and the space soon became bathed in a warm yellow light like daylight. The audio reverberated in the space. A loud cheer of recognition came from the group. “Aaaaahhh! We got it!” Corey asked everyone to get up. As soon as they were all standing, the space became pitched black again. They dropped down again and the space was full of light. They had learned how to “create daylight” in the space. They had completed level one.

Soon after, a new word puzzle was presented to them in a short video projected on two scrims hanging from the ceiling: “The opposite of another word for hello but never settles.” The lights have become very dim now and the audio has a slightly more menacing quality to it. Level two will clearly be more challenging...

We formalized our game structure into a schema of levels, body states and goals, see table 1. As earlier described, the game has seven levels. The body states are the body movements and positions that players must discover in order to complete a level. Goals are the change in environment players are aiming to achieve. The goals are implicit and not explicitly stated for the players. Each level has a beginning quality of light and audio. As the players progress toward achieving the right body state, the environment incrementally shifts toward the goal state of the environment. For example, as depicted in the scenario, when Madison lowered herself, the environment gradually shifted toward the goal of creating day. As each of the other three players followed Madison, the environment responded to movements of each player.

Table 1: This table describes the socio-ec(h)o game schema.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Levels</th>
<th>Body State</th>
<th>Goal</th>
<th>New Game Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery of light</td>
<td>1</td>
<td>“high-low”</td>
<td>create day</td>
<td>body position</td>
</tr>
<tr>
<td>Day for night</td>
<td>2</td>
<td>“moving low”</td>
<td>create night</td>
<td>movement/duration</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>“loosely moving”</td>
<td>create day</td>
<td>proximity</td>
</tr>
<tr>
<td>Rhizome</td>
<td>4</td>
<td>“dense center - scattered edge”</td>
<td>create spring</td>
<td>sequencing</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>“this way slow — low to high”</td>
<td>create winter</td>
<td>sequencing/duration</td>
</tr>
<tr>
<td>Biota</td>
<td>6</td>
<td>“two low moving — two high”</td>
<td>create summer</td>
<td>composition</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>“ringing around the rosec”</td>
<td>create fall</td>
<td>composition &amp; location</td>
</tr>
</tbody>
</table>

In addition, the schema includes new game skills and themes. We assigned each level a generic skill in relation to each body state and level. Despite the specific body state, the generic skill acquired at each level is required in order to discover the more complex body states at higher levels. For example, the specific body-state in level 1 is to crouch down low or sit down. The specific body-state in level 2 is to move low and moving or not moving. In level 2, these generic skills are used to achieve the desired body-state of crawling or moving low on the floor.

Themes allowed us to design an implied progressive narrative based on natural evolution. Again, the specific themes and even the narrative are not known to the participants, rather they provide an underlying structure for body states, goal states and game skill acquisition. We intend for the progressive narrative to provide a sense of coherency across the levels, and to loosely map increased challenge to the reward of a more complex display. The content of our display systems, including light configurations and sound material also contributed in the higher levels toward creating and extending a narrative context.

4. TECHNICAL PROTOTYPE

The technical system for socio-ec(h)o includes three key components, a sensing system, reasoning engine and display engine, see figure 2.

4.1 Sensing System

The sensing engine is comprised of a twelve-camera Vicon MX motion capture system (www.vicon.com) (see figure 3) and a custom program written in Max/MSP. Each participant is differentiated by unique configuration of reflective markers worn on their backs. Data is transmitted to the reasoning engine for high-level interpretation.

The motion capture system data was extracted via a proprietary protocol that Vicon uses to pass data packets between machines. Two Max/MSP objects were written that processed the information needed to allow the system to make decisions about game play.
Reasoning Engine

and 

Table 2: This table details the parameters, threshold ranges, timing, and values utilized in our sensing system.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Threshold</th>
<th>Timing</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity</td>
<td>n/a</td>
<td>n/a</td>
<td>Visible/Lost</td>
</tr>
<tr>
<td>Level</td>
<td>700-900 mm</td>
<td>n/a</td>
<td>High/Low</td>
</tr>
<tr>
<td>Speed</td>
<td>0 &amp; 1.5 mm/sec</td>
<td>2 sec</td>
<td>Still/Slow/Fast</td>
</tr>
<tr>
<td>Space</td>
<td>140-170 mm</td>
<td>2 sec</td>
<td>Stationary/Travelling</td>
</tr>
<tr>
<td>Position</td>
<td>600-800 mm (g) 00.0</td>
<td>n/a</td>
<td>Middle/Outside</td>
</tr>
<tr>
<td>Path</td>
<td>2-3 changes</td>
<td>2 sec</td>
<td>Direct/Indirect</td>
</tr>
<tr>
<td>Orientation</td>
<td>.5 radians</td>
<td>1 sec</td>
<td>N-S/E-W/Horizantal</td>
</tr>
<tr>
<td>Density</td>
<td>600 or 1250mm</td>
<td>1 sec</td>
<td>Loose/Dense</td>
</tr>
<tr>
<td>Duration</td>
<td>n/a</td>
<td>4 sec</td>
<td>Short/Long</td>
</tr>
</tbody>
</table>

Relational associations between people are another set of activities that involved determining what sets relationships have formed among the players. Since this involves measuring relationships between people instead of direct parameter measurements of individuals, the number of sets varies depending on how the associations form and un-form. With four players, this means that up to two associations could be active at any given time. One final set tracked who is not in a relational association with others in the group. Density is the only relational parameter tracked and it is based on the proximity of players to one another.

4.2 Reasoning Engine

The reasoning engine provides the intelligence for the system. It interprets the sensing data samples in real time, identifies the level of body state completion, and manages the narrative flow of the experience, see figure 4. The engine receives sensing data from the sensing system and interprets it in terms of high-level group behavior. For example, the sensing system sends data on predefined parameters such as velocity and body positions. Based on these basic parameters the reasoning engine infers higher level parameters for the user group such as high-fast-moving group, middle-low-stationary group, etc.

The group parameters are further evaluated with respect to the individual user player types and the group composition model (group user model) that is dependent on the combination of user types as identified by Bartle’s classifications [2]. As a result, each state completion is determined by its own function that depends on both individual and group characteristics. The function computes a single value we call state ‘intensity’ which is sent to the display engine.

Another role of the reasoning engine is to manage the flow in the game by sequencing of the states using the interaction model that defines the states and their sequencing at each game level. The single state levels simply require the group to complete the state in order to progress in the game. Multiple state levels require a group to complete a sequence of states within a certain time limit. The engine manages the timing of the state and level transitions.

The reasoning engine is rule-based and allows seamless modification and extension as the game evolves or expands. The reasoning engine feeds its output, state intensity and state transition to the display engine.
constraints led to the choice of a motion capture system because of the past [29], the challenge of performance efficiency, accurate movement. While we have used general video sensing systems in the speed of response and ability to accurately and reliably track three-dimensions, and provide immediate response. These also be able to uniquely identify individuals, track accurately in space the main restriction on the sensing system is that it acquire communication protocol that is transferred over the User Datagram Protocol (UDP) communication channel. We consider unidirectional UDP communication appropriate for real-time applications.

4.3 Display Engine

The display engine has two components, an audio and a lighting component. The audio display engine for socio-ec(h)o provides a sound ecology for each individual level of the system. It is custom software programmed in Max/MSP. We developed and structured the audio content on the principles of acoustic ecology and feedback-as-communication [27]. In addition, the audio display provides a gradient response to the participants, telling them how close they are to achieving their goal. The audio display system can alternate between stereo and multi-channel formats and localized and ubiquitous sound. The audio content follows the theme of evolution by utilizing sampled sound and several different sound processing techniques creating a shifting ambient soundscape that moves from simple, abstract sound to rich, environmental sound.

Lighting is manipulated with a DMX 512 controller via a Max/MSP patch. A small light grid and theatrical style lighting instruments and color scrollers are used. A lighting console was created to control multiple lights and color in concert through a cue list mechanism. Cues were written to simulate the various themes at each level.

Both the audio and the lighting systems take their cues from the reasoning engine, and respond to game aspects and configurations specified in the reasoning engine. Thus, the response of the display systems can potentially be used to provide feedback based on a variety of parameters such as how well participants are working together as a group.

4.4 Integration

Integration of the three components is achieved by lightweight communication protocol that is transferred over the User Datagram Protocol (UDP) communication channel. We consider unidirectional UDP communication appropriate for real-time applications.

5. TECHNICAL ISSUES

5.1 Sensing

In constructing a system to follow activities within the socio-ec(h)o space the main restriction on the sensing system is that it acquire information in a transparent way for participants. The system must also be able to uniquely identify individuals, track accurately in three-dimensions, and provide immediate response. These constraints led to the choice of a motion capture system because of the speed of response and ability to accurately and reliably track movement. While we have used general video sensing systems in the past [29], the challenge of performance efficiency, accurate individuation of multiple objects and high resolution three-dimensional tracking in a real-time environment proves too challenging despite advances in recent research [6].

However, the choice of the motion capture system restricted our movement space to an area that could be covered by twelve cameras - approximately 15 feet in diameter. In addition, several problems needed to be overcome. The Vicon motion capture system uses cameras to track reflective markers in the near-infrared spectrum (www.vicon.com). The system is a "passive" system, meaning that the markers do not send a signal that identifies the marker's identification. This is in contrast to an "active" system where markers can be tracked based on a signal it sends to the system either through a blinking rate or through an electromagnetic signal. Both systems are vulnerable to occlusion of markers by objects in the space that come between the marker and the cameras.

A passive system tracks markers in space and over time through calibrated models of where the markers are located in relationship to each other. Two approaches allow the markers to be identified and associated with the particular person to whom it is attached. First, the system tracks very quickly so that the movement from frame to frame is minimized. Second, the system has a model of what kinds of motion are possible. This forms a constraint model that the system can use to eliminate false positive matches of markers to people. The model contains information about how the markers are positioned in relationship to each other such as distance and joint styles.

Normally, motion capture systems are used to track the locations of limbs of one person as they move. However, with socio-ec(h)o we needed to track four or more people. This requirement is very difficult for a passive motion capture system to accomplish in even constrained environments. Some experimentation verified that the system easily confused movements of multiple people. It would often exchange arms of participants and even legs for arms, heads for hands etc. The system could track individual people for a short period of time if they stayed away from each other.

However, we decided not to track shapes of bodies moving in space but instead to track locations of where people moved. Identification of who is moving together with how they are moving is too difficult for a passive motion capture system to accomplish within the unconstrained socio-ec(h)o environment.

The motion capture system uses relationships of markers to identify individuals. A concept called a "rigid body" is used. Here, the system is told that certain markers will move in concert together, never changing their positions relative to each other. Another technique used to track markers is to specify how markers can move by the types of joints that are between the markers. For instance a hinge joint only allows movement back and forth through an angle.

A multiple marker identification system was created that allowed the system to track people in the space. We used four unique rigid body configurations of five markers each to identify and track participants as they moved. These marker configurations became identification tags, which are worn on the back. Five markers are used so that if any of the points became occluded the system might still be able to match its model to the data it is receiving. The more points that are used by the model the fewer false positives matches occur. Tags were created out of poster board Velcro and the markers supplied by Vicon, see figure 5.
The type of information extracted was qualitative in nature. We needed to have the system evaluate game play within human level perceptions of differences between things. It was decided that for group activities this would be information extracted in dualities such as high/low or very course categorizations such as fast, medium, and slow.

The tags were calibrated into the system prior to game play and remained reliably recognizable by the system. Overall, the system worked extremely well given the moderately constrained nature of the environment and the situation of game play. Information was extracted close to 30 frames per second but used by the reasoning engine at a rate of approximately 10 frames per second. It was found that this was adequate for game play.

5.2 Reasoning

The three major processing steps in the reasoning engine were described in section 4.2 above. The reasoning engine was implemented in Jess (http://herzberg.ca.sandia.gov/jess/). As the real time interpretation of large quantity of sensing data is required, careful memory management and activation/deactivation of the rules involved in the particular state are needed. The rule base for the engine is designed with this effectiveness in mind where only those group parameters that are needed for determining the completion level of the active state are computed. The engine also cleans all the history data in each processing step.

The rules for the system were hand-made and are specific for our set of sensed parameters and play states. There are 61 rules overall. The sample rule in Jess language that computes the group parameter 'high-stationary' overall. The sample rule in Jess language that computes the group parameter 'high-stationary'

\[
\text{intensity} = \text{intensity} + (4 - \text{intensity}) \times 0.2
\]

Similar adjustments are applied for other user types and game states. As a result, the display response in socio-ec(h)o is adjusted with respect to the group composition. We are currently investigating our hypothesis that by considering the Battle types participants have a better experience and more quickly become skilled interactors.

As mentioned earlier in section 4.2, the reasoning engine manages the narrative and experience flow of the interaction. The model for this includes mapping the trajectory of the body states to participant's actions in order to determine the intensity level, or proximity to the desired body state. The intensity level is measured from 0 to 4 with 4 representing the maximal intensity or state completion. As shown above, the intensity function is computed by applying several rules that modify the intensity level with respect to the state completion and group composition. Therefore, the intensity function is not computed by a single formula but is defined by heuristics that are applied in full response to the current state of the game. This provides for a powerful mechanism that enables us to evolve and expand our framework without the need to modify the game mechanism.

In addition, the ability to sustain intensity levels are also monitored, typically a 4 second duration is required to either complete the state or a state within a multi-state sequence (for example see level 4 in Table 1 in section 3.1). The overall shifts in intensities toward and away from the goal must be represented in a gradient effect yet be sufficiently real-time in order to best support actions in the environment. This overall model is utilized for sending data to the display engine in a managed flow, see Figure 6. In addition, the reasoning engine modulates the transitions from one level to another.
5.3 Display

5.3.1 Audio display

The audio display system provides an ambient, immersive auditory space that envelops participants in the play world and signals proximity to achieving the goal state of each game level. The system has three types of responses: an ambient, gradually changing soundscape that is distinct at each level; an intermediate reward sound to signal when all participants are working together towards their goal; and a final reward sound, which signals progressing to the next level. These responses work in conjunction with each other and complement the communication and aesthetic aspects of the users’ experiences. In terms of Schafer's classification of sounds belonging to a local sound ecology, the three types of responses from the audio engine could be considered keynote sounds, sound signals and soundmarks, respectively [24].

The basic idea of our audio display system is to signal participants’ success gradually, by intensifying the environment. Since the gradient response is key in this system, most thought had to be put in choosing the approach to representing intensity, as well as in the selection of the content both from aesthetic and cognitive perspectives.

The system uses several approaches to gradient response. The first is a simple cross-fader between 5 layers of sound which could be arbitrarily chosen to represent increased intensity of the same group of sounds (dripping water gradually changing into a fast river stream, over 5 steps). This type of soundscape design conforms with Trux’s idea of variance and coherence [27] – the balance of sameness and diversity, a core idea in perceptual design techniques. In other words, we use a group of sounds that are the same or very similar according to their basic characteristics of pitch, rhythm and timbre, and we represent success rate in the game by intensifying these basic sound characteristics.

The second step in representing intensity is realized through intensifying of pitch, rhythm and timbre, and wc represent success rate in the environment.

Another approach to representing sound that was introduced as an evolution of previous levels of perception was using a low-pass bi-quad filter. This type of auditory perception is based in timbre differentiation, and while it entails a subtler, less precise representation, it proved quite useful in our system. The common recognition of this sound change was the feeling of “opening” and “closing” of a filter, a sliding from muffled sound to sharp, bright sound through attenuating different frequency bands in a given sound. Another useful technique, which seemed to help in better perceiving change, rather then representing change itself, was the use of 8-channel spatialization of sound layers. This was done by gradually moving the sound in a circle giving the users the impression that sound is “going away” or “getting closer” to their relative position in space. The reason why we thought this approach aided in perception was because the combination of localization and sound change would be greater than sound change alone, a notion, again suggested by some current studies in virtual audio [19].

Table 3. Audio display methods used at different game levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Method</th>
<th>Description of effect</th>
<th>Perceptual aims</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bi-quad Filter</td>
<td>Change from muffled to bright sound</td>
<td>Perception of timbre</td>
</tr>
<tr>
<td>2</td>
<td>Phaser + Layer Fader</td>
<td>A gradual cross-fade of 5 sounds and tempo change</td>
<td>Tempo and timbre perception</td>
</tr>
<tr>
<td>3</td>
<td>Layer Fader + Pitch Shift</td>
<td>Cross-fade between 5 sounds (creaking sound)</td>
<td>Pitch and timbre perception</td>
</tr>
<tr>
<td>4</td>
<td>Layer Fader + Bi-quad Filter</td>
<td>Cross-fade 5 sounds and change from muffled to bright (fire sounds)</td>
<td>Perception of timbre and cultural semiotics</td>
</tr>
<tr>
<td>5</td>
<td>Phaser + Pitch Shift</td>
<td>Gradual pitch shift low to high, inverse tempo relationship (fast to slow)</td>
<td>Perception of pitch, tempo and associations</td>
</tr>
<tr>
<td>6</td>
<td>Layer Fader</td>
<td>A cross-fade of 5 sounds (calm forest – thunder)</td>
<td>Perception of timbre and associations</td>
</tr>
</tbody>
</table>

The final levels of the game, though represented simply by layering and cross-fading sound elements, had an added dimension to them, because as environmental sounds they held recognition and thus created narrative connotations for the game participants, helping them listen to and analyze the change in sound better.

For testing, we only implemented six of the seven conceptualized levels depicted in table 1, section 3.
Our approach to reward sounds, although perceived as "signals" and thus analogous to direct feedback systems, were still a result of a "long-term" composite group action, rather than immediate individual reward. The reward sound signaling the passing of a level is perhaps the only sound mapped one-to-one to a particular event – the completion of goal state.

5.3.2 Visual Display
In regard to the visual display, two techniques are used to provide a gradient response to the participants. Initially, the only feedback provided was based upon intensity of the overall lighting in the room. As participants came closer to achieving the goal, the lighting moved toward a condition of light or dark that corresponded to the system's estimation of closeness to the goal. A direct mapping between intensity and goal closeness is used. While this is a useful technique for providing feedback, it did not provide much room for creating an ambient environment.

A second technique, allowed for environmental ambience and feedback to the participant by using a relative gradient between two states (two lighting states such as different colors and lighting levels for fall and winter). Transitions in the lighting were based on the value of the intensity function (see section 5.2). This is used to signal the lighting to transition toward the goal state over a fixed period of time (10 seconds). If the participant moves away from the goal, the lighting moves toward the other environmental state (the start state). At any point if there is no progress, or there is negative progress, the system moves toward the non-goal environmental state. When the goal state is achieved, then the lighting moves to the goal environmental state.

It was found that both techniques worked for providing feedback. Surprisingly, it was found that the second technique that used relative direction toward the goal provided the participants with a more satisfying experience. We believe that this occurred because the absolute state feedback is provided at a lower resolution of five states. Relative feedback was provided from between 5 to 32 differential gradations.

6. AESTHETIC INTERACTION
We aimed to emphasize the qualities of interaction that result in play that facilitates discovery and therefore explored the embodied and situated aspects of interaction or aesthetic interaction as expressed by Djajadiningrat [9] and Petersen [21].

In the domain of tangible user interfaces (TUI), Djajadiningrat argues for a "perceptual-motor-centered" approach [9]. He is less sympathetic toward the cognitive view of interaction in which he terms the "semantic approach" where objects communicate action through metaphor. We believe this approach equally applies beyond the TUI context to include gesture and movement-based interaction. The direct approach is governed by a "sensory richness and action-potential" to carry expression through interaction. We took guidance in his description of three factors as having a role in aesthetic interaction: the interaction pattern of timing, rhythm, and flow between the user and actions (and/or tangible objects); the richness of motor actions found in the potential space of actions and skill development; and freedom of interaction in which a myriad of interaction paths coexist. In addition, Petersen and her colleagues description of aesthetic interaction shares the embodied aspects described above as well as the sense of aesthetic potential that is realized through the action or engagement [21].

Our approach in the interaction model, visual display and audio display was to create action-potential interaction integrated with interpretive feedback space allowing for aesthetic potential in meaning and expression. For example, our approach to the sound content was significantly different from other audio-based immersive display environments in that it was neither entirely musical, nor was it entirely computer-generated synthesized abstract sound. Instead, the base consisted of sampled (field-recorded) sounds with a varying degree of abstraction and connotation, ranging from water and fire, to processed vocals and transients. We arrived at the final content elements through experimentation with composition of different iterations of sounds and sound processes and transitions, thus drawing from electro-acoustic and musical composition to create an engaging and an aesthetically rich experience for our participants.

The same degree of rich ambiguity and interpretive space was attempted with the word puzzles and embodied solutions. As we discussed earlier, the visual display experimented with qualitative experiences of intensities and gradients (see section 5.3.2).

7. PRELIMINARY USER TESTING
Our user testing to date is preliminary. It includes two three-hour sessions with eight participants, and an additional two-hour session with four other participants. All the participants were new to the game. The group included three females and nine males ranging in ages from twenty-one to fifty-nine. Two of the three groups had a gender mix. Each session began with a warm-up session to introduce the concept of puzzles solved through physical action and support through implicit responses. Participants were also played the range of sonic cues and rewards in order to attune their perceptual hearing to our sound ecologies – we found that prior listening is sufficient in creating recognition of reward sounds when they are heard again later in the game.

Each team of four played two levels followed by questions and discussions. After all levels were achieved or a total of two hours of interaction (60 minutes in the shorter version), the game was stopped and a general open-ended interview and discussion took place. The first group participated in the environment for over two hours and completed four of the six operational levels. The second group completed all the levels in approximately ninety minutes. The third group completed three levels in sixty minutes. All groups were very engaged with the game and those who did not complete it wanted to continue.

Our post-play discussion was open-ended and focused on the overall experience and sense of game-play, collaboration and acquisition of game skills. We also pursued known issues or questions that were relevant to our stage of development. These included the perception of audio and visual thresholds, gradients and the role of abstraction or representation in the sound content; the appropriateness of our intensity function constructed through the reasoning engine (see section 5.2). Through observation and technical data we explored the applicability of our play types (see section 5.2), and known technical issues in relation to sensing and system performance. The results of these sessions serve the basis for our discussion in the following section.

8. DISCUSSION
Problems related to the sensing system did occur which in the end did not adversely affect the perceptions of the participants. While participants often occluded tags by walking out of the space
viewed by the cameras or through interactions where one person covered another's markers in each of the user testing sessions, these occlusions were brief and passed without disruption since the system utilized the last known position without consequence or the player's position was not relevant in progressing to a body state. Typical occlusions lasted less than two seconds but were observed for up to 30 seconds. Another issue with the system was the determination of divisions between categories, (threshold). Here the choice of a number required a calibration based on a subjective judgment of what is in one category or another (for instance what is slow and what is fast). It was found that careful choices of these parameters were adequate. However, in certain circumstances, slight adjustments were required to account for differences in abilities or body types.

Some of the known issues with the audio display include the mapping of game parameters to a number from 0 to 4 and its interpretation into sonic response. The almost complete lack of instant feedback as a departure from standard game design was a known issue and one we were exploring. Yet, the concept of intensity when using sound as representation is much less researched in the field of auditory perception than direct feedback-based response [15]. What we found in participatory design workshops and preliminary user testing was that this approach of gradual response was quite rewarding to the participants and encouraged their attentiveness towards the environment – both for light and sound display, interchangeably or complementary to each other. As well, our representation of intensity was largely successful as participants were able to identify their progress throughout the game based on the environmental cues.

As well, the audio content used in different levels had a varying degree of abstraction. In the course of our testing, it became clear that more abstract sound (such as a crackling of rocks) made it harder for participants to gauge the level of change in the gradient response, as compared to less abstract sound (such as fire or rain). Several of them suggested the feedback was not “crisp” enough, and our main extrapolation from this is that 1) they didn’t perceive enough change and/or 2) they didn’t recognize the sound and perceive its internal coherence or variance. The bi-quad filter turned out to be well received perceptually as an approach to gradual change in the sonic environment, while the phaser, depending of the content was subtler. What participants seemed to respond to most positively was a combination of environmental sound (whether intensifying fire or going deeper into the forest) and a multi-channel diffusion, rather than multiple stereo and abstract sounds. All participants commented on the immersion quality of the play space as a positive and rich experience.

As far as the intermediate and final reward sounds, the issues we were aware of had to do with masking, i.e. sound not heard over the ambience, as well as recognition of the sound, and appropriate content mapping between game state and reward. What we decided on was to use a random pitch variation of two soft abstract sounds (granulated tapping of glass). We also played the reward sounds to all participants before the game began to compensate for the fact that in our preparatory work it had been hard to distinguish the sound without a reference. This strategy seemed to be successful as people correctly identified the reward sound and used it effectively in their game play. On an aesthetic level it was hard to incorporate the seemingly intrusive, yet necessary, “reward” sound into the much more subtle and atmospheric soundscape environment. We considered the use of theme-appropriate rewards, that would be different for each level, but worried this will affect the rate of recognition of the reward, and ultimately – successful game-play.

Since we used a combination of approaches to representing intensity rather than a single approach per level, it is hard to generalize results about the perceptual effectiveness of individual approaches to sound processing. Yet we strongly believed that a combination of sound content and processing would work better than different approaches alone, and our preliminary tests supports that idea.

One of our intentions in this work was to explore the social interaction aspects of responsive environments. In this paper we have discussed the technical platform that we believe makes such an investigation possible. We plan targeted evaluation sessions focused on group collaboration, learning, communication and diverse approaches and goals within groups. In our discussion with testing participants three key observations arose: team collaboration represented by strategy-making and communication evolved considerably over a short period of time; communication was multi-modal including verbal, gestures and physical actions; leadership and decision-making typically rotated and all groups made an effort to share in the leading of the group to some degree, some groups more consciously than others.

9. CONCLUSION & FUTURE WORK

In this paper we have shown how our system builds on the theatrical, simple and physical interaction models. We have developed an ambient intelligence system that functions based on exploratory play with a conceptually structured interaction model. We feel the work contributes by providing a method for constructing group parameters from individual parameters with real-time motion capture data; and we provide a model for mapping the trajectory of participant’s actions in order to determine an intensity level used to manage the experience flow of the game; and design strategies for representing intensity via an audio and visual display. We discussed related work in the areas of ambient intelligent game spaces, play and learning, motion capture based systems, user modeling and auditory display. We provided a detailed account of the technical system. We provided accounts of technical issues related to customization of the motion capture system, a rules-based and composite inference for reasoning on groups, and display issues of gradient responses, audio spatialization and real-time sonic generation and processing. We discussed our movement-based interaction and display in the context of aesthetic interaction. We detailed how the success of the experience relied on collective responses that were real-time, gradient, provided rewards and were unique to different group user models.

Future work includes a series of evaluations of the system to better understand the influence of the game structured interaction model, the supporting user model, and the display. In particular, we aim to understand how our approach enables a better experience and more skilled interactors within an ambient intelligent environment. Other work includes exploring the range and types of information that can be extracted from sensing. The work in this area involves mainly two avenues of approach: exploring new sensor paradigms and exploring types of
information that can be extracted such as more parameters or gesture based information.

10. ACKNOWLEDGMENTS
We gratefully acknowledge the support of Canadian Heritage for the Am-l-able Network and socio-ec(h)os. We also acknowledge support of the Social Sciences and Humanities Research Council and the Natural Sciences and Engineering Research Council.

11. REFERENCES
socio-ec(h)o: Ambient Intelligence and Gameplay

Ron Wakkary, Marek Hatala, Robb Lovell, Milena Droumeva, Alissa Antle, Dale Evernden, Jim Bizzocchi
Simon Fraser University
2400 Central City
10153 King George Highway
604 268 7425
{rwakkary; mhatala; lovell; mvdroume; aantle; devernde; jimbiz}@sfu.ca

ABSTRACT
This paper describes the preliminary research of an ambient intelligent system known as socio-ec(h)o. socio-ec(h)o explores the design and implementation of an ambient intelligent system for sensing and display, user modeling, and interaction models based on game structures. Our interaction model is based on a game structure including levels, body states, goals and game skills. Body states are the body movements and positions that players must discover in order to complete a level and in turn represent a learned game skill. The paper provides an overview of background concepts and related research. We describe the game structure and prototype of our environment. We discuss games research concepts we utilized and our approach to group user models based on Richard Bartle's game types. We explain the role of embodied cognition within our design and elaborate on what we chose to encode as embodied actions, cognition and communication. We describe how we utilized selective responses that were real-time, gradient, provided rewards and were unique to different group user models. We introduce our approach to designing ambient intelligent systems that is ecologically inspired. We stress the empirical nature of the design work and the role of participatory design in developing our system.

Keywords
Ambient intelligence, responsive environment, user model, physical play, puzzles, embodied, audio, motion-capture.

INTRODUCTION
This paper describes the preliminary research of an ambient intelligent system known as socio-ec(h)o. socio-ec(h)o explores the design and implementation of an ambient intelligent system for sensing and display, user modeling, and interaction models based on game structures. Ambient intelligence computing is the embedding of computer technologies and sensors in architectural environments that combined with artificial intelligence, respond to and reason about human actions and behaviours within the environment.

Ambient intelligent spaces lend themselves extremely well to physical and group play. In this paper we describe our design of an interaction model and supporting system based on physical play. The overall research goal of this project is to understand to what degree physical play and game structures such as puzzles can support groups of participants as they learn to manipulate an
ambient intelligent space. Future evaluation of this project will allow us to more fully answer this question. To date we have designed and implemented the prototype and interaction model. We have incorporated formative feedback through a participatory design process. This approach allowed for the concurrent development of the concept, interaction model and prototype environment.

The paper provides an overview of background concepts and related research. We then describe the game structure and prototype of our environment, including a technical overview of the system. We discuss the utilized games research concepts and our approach to group user models based on Bartle’s game types [2], and our ecologically inspired design approach to socio-ec(h)o. Lastly we conclude with a discussion of our work to date and future research.

BACKGROUND
Key contextual issues in socio-ec(h)o include related research in the area of play and ambient intelligent spaces, and literature linking play and learning.

Björk and his colleagues have observed progress toward fully ubiquitous computing games yet they identify the need to develop past end-user devices such as mobile phones, personal digital assistants and game consoles. Accordingly, we need to better understand how “computational services” augment games situated in real environments [3]. Recent projects have investigated the play space of responsive environments and tangible computing utilizing sensors, audio, and visual displays. For example, Andersen [1], and Ferris and Bannon [9] engage children in exploratory play and emergent learning through sensor-augmented objects and audio display. Andersen’s work reveals how theatrical settings provide an emotional framework that scaffolds the qualitative experience of the interaction. Ferris and Bannon’s work make clear that a combination of simple feedback and control lead children to widely explore and discover a responsive environment.

In the Nautilus project, Stromberg and her colleagues employ bodily and spatial user interfaces as a way of allowing players to use their natural body movements and to interact with each other in a group game within a virtual game space [17]. Stromberg observed in physical and team games such as soccer or dodge ball that players coordinate their physical movements and rely heavily on communication to be successful. In their findings, participants reported that controlling a game through one’s body movement and position was “new and exhilarating.” In addition, playing as a team in an interactive virtual space was found to be engaging, natural and fun.

In relation to the above research, socio-ec(h)o builds on the theatrical, simple and physical interaction models in order to develop a game structure approach that lies between exploratory play and a structured game for adults within an ambient intelligent environment. In addition, we extend the notion of a game structure to an interaction model for the environment rather than a virtual game space. We also build on the idea that action, play and learning are linked in such physically-based environments.

In respect to the links between action, play and learning there is a substantial amount of literature. Dewey argued for the construction of knowledge based on learning dependant on action [8]. Piaget, through his child development theory believes in the development of cognitive
structures through action and spontaneous play [14]. According to Piaget, *constructivist* learning is rooted in experimentation, discovery and play among other factors. Papert extends Piaget’s notions by investigating the knowledge-construction process that emerges from learners actually creating and designing physical objects [13]. Malone and Lepper consider games as intrinsic motivators for learning [11]. Subjective motivations like challenge, curiosity, control and fantasy may occur in any learning situation; other motivations like competition, cooperation and recognition are considered to be inter-subjective, relying on the presence of other players/learners. Design related theories have placed activity at the center of design action as in Nardi and O’Day’s activity theory based information ecologies [12]. Schön argues that design is a series of actions involving experimentation and learning in the framing and re-framing of a design situation [16].

**GAME STRUCTURE AND PROTOTYPE**

**Description and Scenario**

The aim of our game is for a team of four players to progress through seven game levels. Each level is completed when the players achieve a certain combination of body movements and positions. At the beginning of each level, players are presented with a word puzzle as a clue in discovering the desired body states (see figure 1). The levels are represented by changes in the environment in light and audio. The levels are progressively more challenging in terms of body states and more complex in terms of the audio and visual ambient display. The physical environment currently consists of a circumscribed circular space (the area in which we can detect motion), surround sound audio, theatrical lighting, and two video projection surfaces.

![Figure 1. Depicted above are frames from a 33 second video introducing a level to players. The images at the beginning of the video provide a sense of the environmental change the players are trying to achieve for that level. This is followed by a clue in the form of a word puzzle aimed at helping players discover the desired body state. In this example, the puzzle is “Too sloe plum turtles – among trees.”](image)

Here is a short scenario of participants in the socio-ec(h)o environment:

*Madison, Corey, Elias and Trevor have just completed the first level of socio-ec(h)o. They discovered that each of them had to be low to the ground, still, practically on all fours. Once they had done that, the space became bathed in warm yellow light and filled with a wellspring sound of resonating cymbals. Minutes earlier, the space was very dim – almost pitch black until their eyes adjusted. A quiet soundscape of “electronic crickets” enveloped them. They discussed and tried out many possibilities to solving the word puzzle: “Opposites: Lo and behold.” They had circled the space in opposite directions.*
They stood in pairs on opposing sides of the space. At Corey’s urging, the four grouped together on the edge of the space and systematically sent a player at a time to the opposite side in order to gauge any change in the environment. Nothing changed. Madison, without communicating to anyone realized the obvious clue of “Lo” or “low”. While Corey was in mid-sentence thinking-out-loud about the puzzle with Trevor, and trying to direct the group into new body positions, Madison lowered herself to a crouching position. The space immediately glowed red and became brighter. The audio changed into a rising chorus of cymbals – not loud but progressively more pronounced. Corey and Trevor stopped talking and looked around at the changing space. Madison, after a pause began to say “Get down! Get down!” Elias stooped down immediately and the space became even brighter. Corey and Trevor dropped down in unison and the space soon became bathed in a warm yellow light like daylight. The audio reverberated in the space. A loud cheer of recognition came from the group, “Aaaaahh! We got it!” Corey asked everyone to get up. As soon as they were all standing, the space became pitched black again. They dropped down again and the space was full of light. They had learned how to “create daylight” in the space. They had completed level one.

Soon after, a new word puzzle was presented to them in a short video projected on two scrims hanging from the ceiling: “The opposite of another word for hello but never settles.” The lights have become very dim now and the audio has a slightly more menacing quality to it. Level two will clearly be more challenging...

Figure 2. Scenes from the final participatory design workshop in which the relationship between body states and word puzzles were explored. The system utilized is an early prototype with the substantive functions implemented. As players lower themselves, the environment becomes progressively darker and the ambient audio more pronounced. Note: the display response is not the same as depicted in the scenario. In this exercise, the goal was to create darkness rather than light.
We formalized our game structure into a schema of levels, body states and goals (see table 1). As earlier described, the game has seven levels. The body states are the body movements and positions that players must discover in order to complete a level. Goals are the change in environment players are aiming to achieve. The goals are implicit and not explicitly stated for the players. Each level has a beginning quality of light and audio. As the players progress toward achieving the right body state, the environment incrementally shifts toward the goal state of the environment. For example, as depicted in the scenario, when Madison lowered herself, the environment gradually shifted toward the goal of creating day. As each of the other three players followed Madison, the environment responded to movements of each player (see figure 2).

Table 1: This table describes the socio-ec(h)o game schema.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Levels</th>
<th>Body State</th>
<th>Goal</th>
<th>New Game Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery of light</td>
<td>1</td>
<td>&quot;high-low&quot;</td>
<td>create day</td>
<td>body position</td>
</tr>
<tr>
<td>Day for night</td>
<td>2</td>
<td>&quot;moving low&quot;</td>
<td>create night</td>
<td>movement/duration</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>&quot;loosely moving&quot;</td>
<td>create day</td>
<td>proximity</td>
</tr>
<tr>
<td>Rhizome</td>
<td>4</td>
<td>&quot;dense center - scattered edge&quot;</td>
<td>create spring</td>
<td>sequencing</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>&quot;this way slow - low to high&quot;</td>
<td>create winter</td>
<td>sequencing/duration</td>
</tr>
<tr>
<td>Biota</td>
<td>6</td>
<td>&quot;two low moving - two high&quot;</td>
<td>create summer</td>
<td>composition</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>&quot;ringing around the rosie&quot;</td>
<td>create fall</td>
<td>composition &amp; location</td>
</tr>
</tbody>
</table>

In addition, the schema includes new game skills and themes. We assigned each level a generic skill in relation to each body state and level. Despite the specific body state, the generic skill acquired at each level is required in order to discover the more complex body states at higher levels. Themes allowed us to design an implied progressive narrative based on natural evolution. Again, the specific themes and even the narrative are not known to the participants, rather they provide an underlying structure for body states, goal states and game skill acquisition. We intend for the progressive narrative to provide a sense of coherency across the levels, and to loosely map increased challenge to the reward of a more complex display.

Technical Prototype
The technical system for socio-ec(h)o includes three key components, a sensing system, reasoning engine and display engine.

Sensing System
The sensing engine is comprised of a twelve-camera Vicon MX motion capture system (www.vicon.com) and a custom program written in Max/MSP. Each participant is differentiated by unique configuration of reflective markers worn on their backs. The system senses for discrete parameters such as velocity, position (x,y,z), orientation, proximity and movement. It measures across each unique player for participation and duration. Data is transmitted to the reasoning engine for high-level interpretation.

Reasoning Engine
The reasoning engine provides the intelligence for the system. It interprets the sensing data samples, identifies the level of body states completion, and manages the narrative flow of the socio-ec(h)o experience. The engine receives sensing data from the sensing system and interprets it in terms of high-level group behavior. For example the sensing system sends data on predefined parameters such as velocity and body positions and the reasoning engine synthesizes
the parameters to determine if a given body state is achieved. The characteristics and their combination, and in some cases their sequence determine the ‘intensity’ of the state. Another factor influencing the state intensity is the group user model that is dependent on the combination of user types as identified by Bartle classifications (see the section below, “Description of Game Concepts and User Model”). The role of the engine is to manage the flow in the game by sequencing of the states and managing the timing of the state transitions. The reasoning engine is rule based and allows seamless modification and extension for other applications. The reasoning engine feeds it’s output, state intensity and state transition to the display engine.

Display Engine
The display engine has two components, an audio and a lighting component. The audio display engine for socio-ec(h)o provides a sound ecology for each individual level of the system. It is custom software programmed in Max/MSP. We developed and structured the audio content on the principles of acoustic ecology and feedback-as-communication [18]. In addition, the audio display provides a gradient response to the participants, telling them how close they are to achieving their goal. The audio display system can alternate between stereo and multi-channel formats and localized and ubiquitous sound. The audio content follows the theme of evolution by utilizing sampled sound and several different sound processing techniques creating a shifting ambient soundscape that moves from simple, abstract sound to rich, environmental sound.

Lighting is manipulated with a DMX 512 controller via a Max/MSP patch. A small light grid and theatrical style lighting instruments and color scrollers are used. A lighting console was created to control multiple lights and color in concert through a cue list mechanism. Cues were written to simulate the various themes at each level.

Both the audio and the lighting systems take their cues from the reasoning engine, and respond to game aspects and configurations specified in the reasoning engine. Thus, the response of the display systems can potentially be used to provide feedback based on a variety of parameters such as how well participants are working together as a group.

Integration
The three components described above run on their own servers. The integration is achieved by lightweight communication protocol that is transferred over the User Datagram Protocol (UDP) communication channel. We consider uni-directional UDP communication appropriate for the real-time applications. Although the sensing system is capable of capturing data at the rate of 30 frames per second we are using slightly longer sampling rate of 200ms for data transferred between the servers. Considering the nature of the output (sound and video) this rate is sufficient for the required fine-grained response.

DESCRIPTION OF GAME CONCEPTS AND USER MODEL
We investigated the play and game aspects of socio-ec(h)o through short ethnography sessions, workshops among the researchers, and games research theory. We explored a range of game concepts including traditional game theory notions such as Nash’s Equilibrium to contemporary games research [4, 5, 15]. Our aim was to encode a form of play for groups that lie between a structured game and open-ended play. Each level acted as a kind of group puzzle - that is, a game with a single solution or winning state [10]. The alignment of increasing difficulty of level
solutions with the increasing skill of the participants as they proceed through the experience is consistent with Csikszentmihalyi's model for developing a state of flow [7].

In the end, we relied on our participatory design process to evolve our game structure. This was especially helpful since we were exploring social aspects of gaming such as communication, collaboration and shared cognition. In addition, our approach was highly physical which required an empirical approach to understanding our concepts. Theoretically, we utilized Bartle's concepts of collaborative play in Multi-User Dungeons (MUDs) and MUD Object Oriented (MOOs) to help us formulate a group user model to support the reasoning within our system [2]. Bartle identified four types of MUD player styles: *achievers*, *explorers*, *socializers*, and *killers*. Achievers seek in-game success, explorers satisfy their environmental curiosity, socializers value human interaction, and killers exercise their will at the expense of other players.

The group user model in socio-ec(h)o is constructed based on these Bartle's types. The group user model is used in the interpretation of the individual actions in the environment and the level with which individual actions contribute towards an overall group activity. The display response in socio-ec(h)o is adjusted with respect to the group composition. We are currently investigating our hypothesis that by considering the Bartle types participants have a better experience and more quickly become skilled interactors.

AN ECOLOGICAL AND PARTICIPATORY APPROACH

The key components of our ambient intelligent model are addressed as a systemic whole; they include interaction, reasoning, response and technology. For example, we investigated the balances between sensor technologies in relation to gesture, inference rules and display responses. Our design approach – inspired by an ecological frame – is centered on human activity and is participatory design driven, informed by observation and theory.

The concept of an ambient intelligence ecology emerged from findings in a previous research project known as ec(h)o. We discovered that ec(h)o had successfully balanced incongruent elements to form a dynamic and coherent system. Components such as interaction, reasoning, audio display and technology shaped the ambient intelligent environment around the purpose of a museum visit [20]. In ec(h)o we explicitly utilized an information ecology approach as an ethnographic analysis of the museum as well as a scaffold for our design decisions [19]. Nardi and O'Day describe an ecology to be a system of people, practices, values, and technologies in a local environment. They argue that the ecology metaphor shifts the focus to human activity rather than on technology [12].

The current research, socio-ec(h)o represents a preliminary exploration of the concept of an ambient intelligent ecology. The experimental nature of the project, its laboratory setting, and the fact that participants cannot be considered to be part of a relevant or definable ecology limit the degree to which this research fully explores the concept. Nevertheless, we feel this is a starting point in investigating an ambient intelligent ecology design approach. At this stage, we found the use of participatory design workshops to be a key component of an ambient intelligent ecology approach. The workshops simultaneously addressed issues of interaction, reasoning, response and technology. We ran five workshops that progressively explored open-ended concepts to more defined concepts. These workshops included investigations of the continuum between play and game, the physicality of our interaction model, the social aspect of play within our type of
space, puzzles and narrative and finally the relationship between our body states and word puzzles. It was evident to us that we required an empirical and qualitative understanding of our concepts, interaction, technology use and prototyped environment while we were designing them.

DISCUSSION
In designing an interaction model and supporting system for play and learning of a complex system, many issues relate to the communication and action between participants and between the group of participants and the system.

Embodied cognition and communication
In our design, a successful participant experience relies on a tightly coupled system emerging from real-time, goal-directed interactions between participants, and between participants and the responsive environment. The nature of these interactions influences the challenge, enjoyment and success within the environment. Communication is key to solving the puzzles and coordinating actions. While the verbal discussion among participants is frequent and active, the physical nature of the interaction model and game structure emphasizes explicit embodied action, cognition and communication. Players actively work out the puzzle physically, as well as communicate actions and ideas physically. In many respects, the interaction model is founded on an embodied cognition view of interaction [6, 21]. Success in the game requires a quality of interaction in which mind, body, and environment mutually interact and influence one another positively.

From the perspectives of the design of the interaction model and system, we realized it was important to decide where not to specify interaction and system functionality. In many respects, we learned to off-load formalized interaction among participants to the situated dynamics of a group of people working toward a shared goal, i.e. people will communicate together in whatever form possible given the resources in the environment without the need of formalizing communication modes. In addition, the system does not need to encode or sense actions or behaviours that are not relevant to the desired body-state at a given level, i.e. no response from the system is a perceived response. We feel we supported an embodied and inter-subjective approach through limited means such as design constraints. For example, limiting sensing actions to whole body positions and movements rather than gestures, opened gestures up to unique, specific and complex communication between participants. Ignoring or not encoding large parts of the embodied action supported a wide range of exploration of body movements.

Selective Real-Time Response
We were however selective as to when and how the system did respond to participant’s actions. The system responds when it appears that the group is on a trajectory toward the desired body state. The response is conceptually similar to someone telling another if they are close to a goal by stating if they are cool, warm or hot. Through our participatory design workshops we learned that four factors had to be met in our response in order to achieve this type of support through a changing environment. The response had to be in perceived real-time, the feedback is on a gradient related to the proximity to the goal, a reward is given for achieving a goal, and a response is mapped to the make-up of the group. The quality and nature of physical action requires an accepted real-time response. Given that we require a minimum duration before a body state is recognized, we had to find the threshold for what was understood as a required time
to hold a position versus a perceived lag or failure of the system. A gradient response is critical in aiding players in understanding they are on the right track. The response is coordinated between the audio and the lighting. The nature of the gradient response is well illustrated in the scenario and figure 2 above. An audio reward is given once a state is completed. This is needed since the continuum of coordinated actions and durations is not explicit to the participants. Based on the different groups of players determined as a composite of Bartle types, we modified response and time. While we have yet to fully evaluate this approach, we anticipate a group of achievers will expect a different response or precision of action than a group of primarily explorers. With this in mind, we also provided a range of word puzzles of differing difficulty and challenge.

Empirical nature of designing ambient intelligent systems

An ambient intelligent ecology, as stated above, is an ongoing investigation of how we might define a design process specific to the challenges of ambient intelligence. Given the centrality of situated human activity and the need to develop an interaction model and system as a systemic whole – reducing the process or system to discrete elements is not a reasonable approach. While a theoretical approach to the design and system helps frame the problem and support design decisions, ultimately it is an empirical process such as ethnography or in this case, participatory design that yields useful qualitative and quantitative understanding of how to design the interaction model and system. The application that arises, such as socio-ec(h)o becomes a specific ecology of constraints, affordances and system functions, that is situated and relies on a unique dynamic of embodied action between people and environment.

FUTURE WORK & CONCLUSION

In this paper we have reviewed related research and have shown how our system builds on the theatrical, simple and physical interaction models in order to develop a game-based approach to ambient intelligence that relies on exploratory play with a conceptually structured interaction model. We discussed the links between play, learning and action that we extended into an embodied cognition approach. We provided a description of our game structure and prototype from conceptual and technical perspectives, and we discussed how we use Bartle’s game types as the basis for our user types and group user model. We introduced our approach to designing ambient intelligent systems that is ecologically inspired. In our discussion, we explained the role of embodied cognition within our design, and elaborated on how we decided where and where not to formalize and encode embodied actions, cognition and communication. We detailed how the success of the experience relied on selective responses that were real-time, gradient, provided rewards and were unique to different group user models. Lastly, we stressed the empirical nature of the design work and the role of participatory design in developing our system.

Future work includes a series of evaluations of the system to better understand the influence of the game structured interaction model, the supporting user model, and the display. In particular, we aim to understand how our approach enables a better experience and more skilled interactors within an ambient intelligent environment.

ACKNOWLEDGEMENTS

We would like to gratefully acknowledge the support of Canadian Heritage for funding the Am-i-able Network for Mobile, Responsive Environments and this project. We’d also like to acknowledge funding support from the Social Sciences and Humanities Research Council and
REFERENCES

Museum as ecology: A case study analysis of an ambient intelligent museum guide

Ron Wakkary and Dale Evernden
Simon Fraser University, Canada

Abstract

This paper explores the usefulness of the ecology concept as an analytical framework for designing interactive technology in museums. We aim to describe and evaluate an ecological approach to understanding museums and to examine information and cultural ecologies as analytical tools for guiding the design of interactive systems. We focus on two related concepts of ecology, cultural ecology (Bell 2002) and information ecology (Nardi and O'Day 1999). Utilizing each of the two frameworks, we analyze observational and interview data we collected during the research of an ambient intelligent museum guide. We also discuss the design implications of our analysis. In this paper we found that an ecology framework is highly appropriate for representing the complexities of activities, relationships, technologies and people connected to museums. We also found the information ecologies framework to guide design decisions in the creation of our interactive prototype.

Keywords: ecology, information ecology, cultural ecology, design ethnography, ambient intelligence, design.

Introduction

This paper explores the usefulness of the ecology concept as an analytical framework for designing interactive technology in museums. We aim to describe and evaluate an ecological approach to understanding museums and to examine information and cultural ecologies as analytical tools for guiding the design of interactive systems. The paper will be of potential benefit for designers, researchers, developers, and museum administrators with an interest in interactive museum guides.

We focus on two related concepts of ecology, cultural ecology (Bell 2002) and information ecology (Nardi and O'Day 1999). We begin by providing an overview of the role of ecologies in helping designers better understand museums, and an overview of the case study, an ambient intelligence museum guide. The paper then discusses each ecology concept separately. In each case we provide an introduction to the concepts, the relevancy of the concepts to our case study, and an analysis of our design ethnography data.
or design process based on the ecology frameworks. This is followed by a discussion of design implications of the ecology analysis and its usefulness in describing museums for the design of technology.

Ambient intelligence computing is the embedding of sensor and display technologies in architectural environments that are supported by artificial intelligence such that the overall system responds and "reasons" about the human actions and behaviors within the environment. The aim is to distribute technology away from people into the surrounding environment in order to better support interactivity.

The idea of an ecology framework emerged from the research and development of the case study project. Our design was guided by an understanding of information ecologies as an approach to understand appropriate interventions through design in existing social systems such as museums. We had not yet encountered the concept of cultural ecologies but have since found it to be an applicable schema for our design ethnography. In this paper we found that an ecology framework is highly appropriate for representing the complexities of activities, relationships, technologies and people connected to museums. We also found the frameworks support appropriate and localized design decisions in creating interactive systems.

Ecologies, Museums and Design

Nardi and O'Day's notion of information ecologies describes a system of people, practices, values, and technologies in a local environment (Nardi and O'Day 1999). They argue that the ecology metaphor shifts the focus to human activity rather than on technology. For example, a library is an ecology for accessing information. It is a space with books, magazines, tapes, films, computers, databases and librarians to help find information. The technology components of the ecology are balanced to shape the environment around human action in accessing information. From the perspective of sociology, Bell has described museums in terms of cultural ecologies (Bell 2002). Both the museum space and museum visit experience are bound by interrelated ecological components.

Why is this relevant?

Our design choices were strongly influenced by the awareness of museums as deeply complex and dynamic spaces. Von Lehn, Heath and Hindmarsh have described museum experiences as multivariate, that is, they cannot be assessed by a single factor such as exhibit design, signage, or time spent in front of an artifact (Lehn, Heath et al. 2001). Instead, the museum experience is subject to multiple influences and results in multiple outcomes. Our design aimed to limit the degree to which we intervened in what is already a complex situation. We attempted to leverage existing actions, localized knowledge, and attributes of a museum. In support of these aims, we chose to develop an ambient intelligent system that utilized an audio display interface, a tangible interface, and a user-model and adaptive information retrieval system.

The audio display allowed us to provide a virtual context without having to explicitly impose a new visual interface on the museum visitor. We viewed the exhibition space to be made up of existing layers of visual interfaces from the objects themselves, to the didactics, the display cases, and the architecture of the space. The use of a tangible interface allowed us to transition between the physical space and the virtual information space. We wanted to minimize the degree to which visitors had to learn something new in an environment in which they were already learning interface conventions such as display cases, didactics and the exhibition itself. We aimed for the information the visitors received to be both responsive to them as well as specific to the knowledge and practices of museums – the locale of the experience.

The ecological model allows us to look further into the design process past the interface for guidance in our design intervention. Our aim is to examine the degree to which we as designers understand the museum as an existing ecology, thus capturing it in its multivariate state. Second, to examine how our design decisions were integral to the ecology or ecology inhabitants, thus developing appropriate design responses.
Case study

The case-study is of the design process for an ambient intelligent museum guide known as ec(h)o (Wakkary, Newby et al. 2003; Hatala, Kalantari et al. 2004). The prototype is an integrated audio, vision and location tracking system installed as an augmentation of an existing exhibition installation. ec(h)o is designed to create a museum experience that consists of a physical installation and an interactive layer of three-dimensional soundscapes physically mapped to museum displays and the exhibition installation. Each soundscape consists of zones of ambient sound and "soundmarks" generated by dynamic audio data that relates to the artifacts the visitor is experiencing. The soundscapes change based on the position of the visitor in the space, the visitor's history with viewing the artifacts, and their individual interests in relation to the museum collection. Through a tangible interface, a wooden cube, the visitor can interact with a single artifact or multiple artifacts in order to listen to related audio information. The audio delivery is dynamic and generated by agent-assisted searches inferred by past interactions, histories and individual interests (Hatala, Kalantari et al. 2004). The prototype was installed and tested in the Finders Keepers exhibition at the Canadian Museum of Nature. The exhibition theme was collecting natural history artifacts in Canada.

In this paper we will discuss aspects of the design process behind the development of the prototype. Many of the design choices for implementation were driven by a series of participatory design workshops and scenarios, details of which have been written in another paper (Wakkary 2005). Our particular attention here will be on our design ethnography data, and an analysis of our collaborative design with the staff members of the museum. We will examine the data we retrieved and documentation of the process through the ecological frameworks.

Figure 1. Museum visitor testing the ec(h)o system. The wooden cube held in her hand is the prototype's tangible interface.

Museum as Cultural Ecology

The anthropologist, Genevieve Bell has described museums in terms of cultural ecologies (Bell 2002). As an ecology, the museum space and experience are bound by interrelated components and attributes of the given museum ecology. For example, Bell identifies three significant components of all museum ecologies, liminality, engagement and sociality. Liminality defines museums as places that embody an experience apart from everyday life – as such the experience can be transforming. Engagement defines museums as places where people go to learn yet often in an entertaining and exploratory way. Sociality defines museums as social places for groups such as pairs and families:

- **Liminality**: A museum visit is liminal in that it is experientially set apart from everyday life. Positive museum experiences are transformative, spiritual, and moving. A museum visitor should be inclined to pause and reflect. Liminality permits a deeper engagement.

- **Sociality**: People attend museums in groups such as couples, families, classes, friends, and dates. Museums have become social spaces, for example supporting elements of this include cafes, gift shops, and lobby spaces. Museums are seen as safe places for families. They often include games and are designed to guide supported activities that increase the interaction between parents, children and artifacts.

- **Engagement**: People go to museums to learn. This engagement is often packaged in an entertaining way, for example museums are increasingly viewed as a tourist destination. Museums are a balance between learning and entertainment spaces.

Bell sees the museum visit as a ritual determined by space, people and design. She decomposes the visiting ritual into three observational categories: space, visitors, and interactions and rituals. Altogether the ritual has design implications in relation to technology. Among the different museum ecologies Bell has described, she outlines two museum ecologies: art museum ecologies and science museum ecologies. These ecologies are seen to be distinct and supportive of very different kinds of museum visits. In order to better understand the specific nature and differences among the ecologies we’ve compared Bell’s analysis of the two museum ecologies (see table 1).
Table 1. This table shows a comparison between art museum and science museum ecologies.

In our case study, a natural history museum, it became evident that it closely followed the science museum ecologies. For example the attributes of space were explicitly instructional, provoked play, and often demanded interaction.

**Relevant aspects of a cultural ecology for our case-study**

Not unlike Bell, we observed the Canadian Nature Museum visitors through traditional ethnographic techniques. While we had a different and more descriptive schema for analysis we found that all our data fit well in Bell's ecology components. The design implications of the science museum ecology are directly in line with our own assumptions coming in to the project. Yet, we had not systematically connected these assumptions to our observations as Bell's framework does.

**Analysis of our case utilizing cultural ecologies**

What follows is a summary analysis of some of the observational data we collected based on the cultural Bell's science museum ecologies' framework.

**Visitors**

Many examples of diversity in visitors and visit rituals reported by Bell were indeed observed in our own study. For example, the cultural diversity of visitors noted in the science museum ecologies was found in our observations, particularly across student and tourist visitors. Bell states that in contrast to art museums, "science museum ecologies seem to teem with activity and noise" (Bell 2002 p. 11). It is clear from both Bell's observations and our own that this can be attributed to the large number of children and teenage visitors to science museums. Bell observes a strong presence of middle school and high school students. This was in keeping with our own analysis, however visitors were typically on the younger end of the range (between 8-14 years of age), and often as groups on a school field trip. Though visiting as a large group, these visitors often fragmented into smaller groups and clearly enjoyed a level of autonomy—running, jumping, speaking loudly, laughing, and exploring the space as if it were a playground, despite the accompaniment of adult teachers, for example see figure 2 (below). This speaks to the safe social spaces that museums have become. Similarly aged visitors were observed as members of families. Despite parental supervision, the playful nature of the visit ritual was only slightly more tempered.

Most of the adults observed in the museum space were younger or middle aged, with few instances of elderly visitors. Most adults were observed as couples or parents with their typical visitation ritual being a quiet, semi-absorbed inspection of the space. This contrast between the adult ritual and the youth ritual was noted both by Bell and ourselves.

Like Bell, we observed very few instances of individual visitors. Although visitor groups dominated, we did note many instances of individual exploration. In these instances members of a group would venture off, explore the space for experiences of interest, and then report back to the group their discoveries. The younger visitors in particular seemed to enjoy this pattern of interaction. We made note of numerous instances where groups of students would disperse and come back together, or younger members of a family would explore and report back to their parents.

Figure 2. The Finders Keepers exhibition at the Canadian Nature Museum.

**Space**

Strong similarities between Bell's ecologies and our observations occur in other categories as well. For example, Bells discussions of space and displays matched our observations. Bell argues that displays and installations tend to be demanding of the visitors' attention. We found these characteristics in the four exhibitions we analyzed. For example in the mammals exhibit was a series of dioramas depicting animals in
their natural habitat. In addition to the visual scene, the dioramas were accompanied by ambient sound effects, a detailed didactic, and at least two tactile artifacts relevant to the mammal on display such as fur samples, hoof imprints, and horns busts.

Another key characteristic of Bell's ecologies is that installations in science museums are explicitly educational. Bell reminds us that this differs from the art museum ecologies where information tends to be sparse and the aim is to facilitate interpretation over education. Our observations noted numerous instances of explicit educational and instructional approaches to the information design of the displays. For example, in the Creepy Critters exhibit, living specimens in terrariums and aquariums were supported by graphical representations on didactics coupled with a short paragraph on each species. This information was further supported by a collection of nearby reading materials (books, magazines, species sheets) related to the display.

**Interactions and Rituals**

The displays and installations revealed diverse forms of interaction: microscopes with adjustable slide wheels that could be turned to explore different specimens; wooden puzzles which once completed would fall apart at the pull of a handle, creating a loud crashing sound that captured the attention of others; a collecting game called "The Rat Pack Challenge" which tasked visitors to search the room and discern collectable artifacts from non-collectable ones; discovery drawers filled with objects like, fossils, fur pelts, and minerals which visitors could touch and inspect at close range; push button audio and video installations; scale models and artist recreations of dinosaurs that people could walk up to and touch; terrariums and aquariums filled with living specimens; magazines, coloring books, and small library of natural history artifacts that were loaned to students.

Figure 3. Full-scale wooden puzzle that once completed would create loud crashing sounds at the pull of a lever (on the left).

Bell describes an attribute of the science museum ecologies is to support the fact that people learn in a variety of different ways. Alternative approaches to learning turned up throughout our observations. Examples include interactive puzzles, quizzes, and games that require visitors to explore and think about the artifacts being displayed; multimedia stations provided for deeper investigation into a topic or subject area; and dioramas and immersive displays allowed visitors to imagine through visual simulations of historical periods.

**Museum as Information Ecology**

Bell's ethnographic efforts focused on observational experiences of the museum visit. Anthropologists Bonnie Nardi and Vicki O'Day draw on activity theory (Vygotsky 1925/1982; Nardi 1996) and field studies of technology libraries, virtual worlds, an architectural firm, high schools and a teaching hospital in order to develop their concept information ecologies. The concept they describe strives for a more systematic view of organizations, based on the relationships among people, practices, technology, values and locale.

Nardi and O'Day use the concept of ecology in order to critique current understandings of technology as autonomous. The authors aim to reframe the metaphoric underpinnings of certain views of technology such as "technology as a tool", "technology as text", and "technology as system". They argue for a more complex understanding of interdependent elements and influences of which technology is only one part; a view that holds a sense of urgency due to opposing possible outcomes of entropy and failure or growth and dynamism. Constituent elements of information ecologies include a system, diversity, co-evolution, keystone species, and locality:

- **System**: Information ecologies are distinct through its strong relationships and dependencies among its different parts. The parts may be different yet they are complimentary and extend each other in capabilities. Change can also be seen to be systemic; when one part changes it affects all others.

- **Diversity**: As in a biological ecology, different species thrive in different niches in the ecology. The complex interdependencies ensure that many different kinds of roles and functions exist. An
information ecology contains a diverse set of people, activities and technologies. It in fact relies on diversity in order to successfully respond to change and potential chaos.

- **Co-evolution:** Parts of the information ecology adapt to each other or co-evolve as change occurs only to be reintegrated later. The ecology is always balancing the various toeholds of co-evolving parts.

- **Keystone Species:** Ecologies are marked by the presence of certain keystone species whose presence is crucial to the survival of the ecology itself. Often such species take the role of mediators who bridge institutional boundaries and translate across disciplines. Introduction of new technologies in an ecology is often reliant on mediators who shape tools to fit local circumstances.

- **Locality:** The participants within the ecology aid in giving identity and place to things. For example, the habitation of technology provides us with a set of relationships within the ecology, to whom a machine belongs determines the family of relationships connected to the technology. We all have special knowledge about our own local ecologies that is inaccessible to anyone outside thus giving us local influence on change. Only those who are integral to the ecology can provide a local habitation and name to technology.

**Relevant aspects of an information ecology for our case-study**

Can an information ecology understanding of a museum positively inform our design process and outcome? In our design approach we contextualized our efforts within the concept of information ecologies. That is we looked at the Canadian Nature Museum as an information ecology and even further, actively sought out keystone species for localized knowledge. One assumption of ours was that we were intervening at the level of the individual museum visitor at a scale and form that was appropriate. Nardi and O'Day argue that information ecologies are scaled to individuals, and through considering it at this scale we can find individual points of leverage, ways into the system, and avenues of intervention.

What follows is a summative reflection through the lens of information ecologies, of our collaboration in the design of our prototype with the Canadian Nature Museum and its staff.

**Analysis of our case utilizing an information ecology framework**

Our observations within the museum and the formal and informal sessions with museum staff were guided by our understanding of information ecologies. Our design goals lead us to this process and in many respects the experience validated for us the applicability of the framework to an organization like a museum. In table 2 (see below), we have provided a matrix analysis of our observations. In order to provide more detail to our findings we've summarized a few key dimensions of the ecology including diversity, co-evolution, and locality.

| Table 2: A matrix analysis of the Canadian Nature Museum as an information ecology |

**Diversity**

Visits with the museum staff, as well as observations of the exhibition spaces, indicated a high level of diversity of people involved with the museum. The visitor population was clearly diverse. We observed a range across groupings, ages, ethnicity and gender, including young students with teachers, college students, middle-aged tourists, large and small families, and young couples. Among the staff, the range of roles and functions at the museum is broad, including scientists, security personnel, office administrators, cafeteria workers, finance and administration personnel, archivists, media producers, 3D modelers, designers, and facilities personnel.

An outgrowth of the museum's staffing needs is that the institution demonstrated an equally diverse set of practices. This is exemplified in the list of service departments found within the museum: corporate services, collection services, communication services, community services, exhibition services, finance, human resources, information services, library services, and research services.
Further support of our use of an information ecology framework was the different values that emerged from interviews with museum staff. These differences appeared across and within departments. For example, among the values we found were scientific credibility, safety, security, entertainment, education, accessibility, affordability, preservation, longevity, bilingualism, community involvement, collaborative learning, and distance learning.

Finally, it was clear from our observations, interviews and tour of the collections and exhibition facilities that a wide range of technologies was employed. The types of technology varied from highly specialized to general solutions, and from technically sophisticated to simple. Observed technologies ranged from advanced 3D scanning equipment used to digitize artifacts for web exhibitions, to acid free collection boxes used to preserve beetle specimens, to motion detectors for artifact security within the exhibits, to a jar of water employed to maintain humidity in an elephant skull exhibit case.

Co-evolution

Nardi and O’day describe co-evolution as an adaptive response to change, “information ecologies evolve as new ideas, tools, activities, and forms of expertise arrive in them...parts of the system adapt to each other or co-evolve as newer, faster, and different tools, are integrated repeatedly” (Nardi and O’Day 1999 p. 52). In our work with the museum we observed numerous instances of co-evolution.

During one of our visits we were given a tour and demonstration of the museums recently purchased 3D scanning technology. It was explained that this technology and entirely new department was meant to support the museum’s mandate to provide Internet access to the collection for the general public and research community. Such an investment in technology, as well as the staff to support it, exemplified the sort of co-evolution described by Nardi and O’Day. Clearly the Internet and changing imaging technologies impinged upon the museum resulting in the decision to adopt new tools in order to respond to these changes.

During our formal and informal sessions with staff we encountered great excitement over major renovations planned for the museum’s exhibition facility. The museum’s exhibition design team discussed upgrades to exhibits and possible new technologies to support the upgrades. Such future possibilities and a new mandate of interactive technology use, seemed to genuinely increase the excitement and interest around the museum in new interactive technologies and in particular, our project. The significant focus on future spaces and interactive exhibits resulted in lessening interest or attention to existing displays and exhibits.

While parts of the system change and create new momentum other parts find themselves in stasis. For example we observed significant disparities in the adoption of new technologies across the ecology. Use and adoption of new technologies and supporting practices appeared to flourish behind the public face of the museum in areas such as the administration offices, research facilities, and collections/storage facilities. Conversely, there was a noticeable lack of new technology adoption in the public exhibition spaces. Similarly, the archives department of the museum had yet to respond through new practices or technologies to issues of access and media format. For example, it was very difficult to access in any indexical form the vast collection of field recordings the archive held. The outmoded and heterogeneous types of analog media of the archival objects further complicated things.

Locality

Nardi and O’Day argue that we all have special knowledge about our own local ecologies and that this knowledge tends to be inaccessible to anyone who exists outside of that ecology. In addition the habitation of a technology will always be an outgrowth of its location within a network of relationships. For example, understanding to whom an artifact is connected or belongs, sheds light on other relationships and aspects of local participation, local engagement, or local conditions unique to specific settings.

During a formal interview session with one of the collection managers, we found that that an absorbent flooring and sheet plastic we had observed under a collection of invertebrate jars was in fact a low budget conservation technique. The action guarded against potential spills caused by temporary construction in
an adjacent exhibit. In the same interview we discovered that a jar of water in an elephant skull case was a highly successful yet inexpensive solution for managing the humidity needs of that specimen.

We recorded numerous instances where a staff member’s localized knowledge transformed a collection of artifacts from opaque and dull to interesting. For example, one of the scientists we interviewed was a paleo-anthropologist. Her knowledge was highly specialized; the information and insights she passed on during the interview were valuable in that they gave life to what was a static display of bones. She provided engaging stories that connected the archaeological artifacts (animal bones) to the history and living practices of an ancient Inuit people. In addition, she provided details and explanation of the forensic process of sorting and analyzing the artifacts that she herself had performed.

Discussion

The cultural ecologies framework validated many of our design goals. The fact that our observations fit nearly completely with Bell’s description of science museum ecologies strongly validated our own assumptions. In addition, a key assumption mapped directly to Bell’s design implications which are formally linked to the ecologies: “technologies need to support existing interactions and create the possibilities for others conditioned by the visitor expectation that this interaction will be engaging and dynamic” (Bell 2002 p. 12). We therefore aimed to integrate within the wide range of low-technology interactive solutions, and to keep our prototype playful, and above all accessible to the museum visitor. Our observations that fall within Bell’s categorization of interaction and ritual emphasized that our system should be open to multiple forms of input such as movement and physical interaction with the displays, and responsive to different learning styles. In many respects, our prototype became a virtual extension of the exhibition space and acted as an augmentation to the didactics and other learning materials.

Information ecologies informed early conceptualizing of the project and specifically guided design decisions at later points. For example, our understanding of the co-evolution dynamics of the ecology led us to consider our prototype as a complementary influence. Co-evolutionary trends that we identified included a strong investment in visual technologies to support access to the collection and new approaches to exhibition design. We saw the opportunity to contribute an alternative approach to interactive display based on audio. In addition, we saw the chance to give form to the intellectual knowledge of the museum staff in addition to the embodied knowledge of the artifacts.

Another example of how the information ecology understanding guided a design decision arose out the issues of locality and keystone species. This led to a novel approach to content design and development. In our earlier discussion on locality, we commented on our numerous observations of informal yet engaging delivery of specialized knowledge on behalf of the museum researchers. The majority of these types of exchanges happened as we toured the collections and storage facility. Stories connected to artifacts ranged from anecdotes on where the artifact was found and how cold it was at the time or how difficult the terrain was, stories of the difficulties of mold-making on site or humorous tales of transportation and objects temporarily getting lost, to what the objects tell us or how their meaning has changed. Often these were first hand accounts and discussed in the most informal and wide-ranging manner. Factual or thesis driven accounts of artifacts were mixed with anecdotal and humorous tales related to the discovery, processing or research of the actual artifact. This experience deeply struck us since our shared perception of the public exhibition display space was quite the opposite. Not unlike many exhibitions, the artifacts and contextualizing information appeared static and lifeless. In locality terms, it was evident to us that once the artifacts were connected to people, the understanding of these artifacts became deeply connected to all aspects of the ecology and came out in the form of storytelling that covered activities related to the artifact, conservation, storage, research and display technologies, meaning and values associated with the artifacts – all situated in specific contexts of time and place.

The exhibition facility and the collections/storage facilities were geographically quite far apart. The collections/storage facilities also housed the research facilities. It was evident that on a large level, two distinct ecologies were forming. The ecologies of the research and collections facilities were more dynamic, involving present day practices and perhaps greater informational exchange around the artifacts and specimens. One goal in the design was to bridge the gap between the ecologies with the outcome of bringing more to life the artifacts on display. We aimed to model our information delivery and audio
experience on the informal storytelling manner we had experienced. We aimed to create a virtual cocktail party of natural history scientists that accompanied the visitor through the museum.

We identified the scientist collectors as our keystone species. Our content development process centered on formal interview sessions, a majority of which were scientist collectors, followed by a walkthrough of the exhibition space captured on video (Wakkary, Newby et al. 2004). In the first session, we asked that they speak generally about what they do and how it relates to artifacts at the museum. We aimed to elicit the range of storytelling we had encountered earlier. In the walkthrough sessions we asked them to speak to the artifacts as well as the exhibition display in the exhibition. These sessions became the basis for discrete audio objects that were categorized by topics and relationship to artifacts on display. In our prototype our system reasoned on possible objects the visitors might want to hear based on their previous interests, movement and interactions in the exhibition. In the end, it was very much like a virtual cocktail party, especially since we strived to maintain a diversity of delivery in language styles and voices.

We come to ecologies as designers who use ethnographic techniques and concepts to better our designs. Our aim is twofold, firstly to best understand the complexity of the people and context that make up our design situation; secondly to leverage the role of participant observer into a collaborative design relationship through direct collaboration or our informed representation of the stakeholders for whom we design. For our purposes, both ecological frameworks served our goals despite their strong differences. Bell’s descriptive framework, the cultural ecologies formally linked different actions and attributes of the museum visitor into a coherent system. As a descriptive tool it validated our assumptions and provided a clearer link between what we observed and the design implications. We could see it being useful at the early stages of the design process as an ethnographic schema and initial framing of the design situation. Nardi and O’Day’s information ecologies framework was generative as well as descriptive. It guided us in specific design decisions such as the aim to make localized knowledge accessible, as well as seeking a complementary influence to current technology interests and use.

More generally, our findings are that ecologies effectively capture in a formal and systematic manner the museum experience and the organization itself. The evidence for the information ecologies is tied to our design outcome and process of which we’ve detailed extensively in this paper. Bell’s cultural ecologies supported almost all of our observations and insights.

It is important to note however, that although Bell’s science museum ecologies mapped directly to our observations, some important characteristics did not. Despite Bell’s observations of a high adoption rate of interactive technology, we found with the exception of video kiosks and push button audio stations, very few technology driven interactive elements were found in our analysis of the museum. Interactivity not mediated through computer technology was common and took the form of puzzles, games, microscopes, and touch-me artifacts. The distinction between a science museum and natural history museum may be a result of differing disciplines and therefore different levels of receptivity to technology; yet, one might expect other differences in other attributes of the ecology to occur as well.

Our analysis of the ecologies is only evidenced in a single sample of one institution. Our visitations and interviews occurred over three separate trips each lasting between one to three days. We would like to pursue future research in other museums over slightly longer periods of time. In addition, we would eventually like to return to the Canadian Nature Museum after the renovations to the new space and perform another formal study to look for shifts in the ecologies.

**Conclusion**

At a time when museums are undergoing change and adopting new interactive technologies the ecological frameworks we’ve discussed in this paper are deserving of greater evaluation. Both cultural ecologies and information ecologies focus on technology use within the ecologies and therefore have a specialized function in understanding museums today. As designers, we found these two approaches supported design goals and contributed to the making of key design decisions. The ecologies provided us an in depth understanding of the museum visit experience and the organization.

The cultural ecologies framework validated many of our design goals. Our observations fit nearly completely with Bell’s description of science museum ecologies with the exception of the museum’s use
of computer technology for interactive displays. We could see using the framework as an ethnographic schema in future museum projects.

As designers, the information ecology provided us with deep and multiple views of the organization and the people connected to it. Like ethnography in design, ecologies have a role in the design process that can guide design decisions in addition to systematically providing a descriptive understanding of the design situation. In our particular case, our design was guided by our findings of locality, keystone species, and co-evolution.

Our case-study analysis suggests that ecological frameworks can be effective descriptive tools, and in certain instances a generative tool as well. Future research includes additional formal studies of museums as well as a study in the shift of ecologies over time.

References


Acknowledgements

Work presented in this paper was supported by a Canarie Inc. grant under the E-Content program. Authors would especially like to thank Mark Graham and his colleagues at the Canadian Nature Museum in Ottawa for their enthusiastic support of this project and the invaluable time they dedicated to it. We would also like to thank our colleagues and participants in this research, namely Marek Hatala, Kenneth Newby, Robb Lovell, Leila Kalantari, Doreen Leo, Gilly Mah, Mark Brady, Jordan Williams and the support of Simon Fraser University.
Consciousness Reframed 2004

Design and Complexity: Research into Practice

Ron Wakkary

Abstract: The paper proposes a rethinking of design practice through the lens of complexity. The paper analyzes design and human-computer-interaction (HCI) theories related to reflective practice and context as approaches complexity. The paper calls for the framing of a larger research agenda with the need to further work on issues of research into practice in design.

Keywords: design theory, complexity, reflective practice, design practice, human-computer-interaction

Introduction

A current challenge is that most design and human-computer-interaction theories implicitly claim that many design situations are virtually unapproachable, or at least better ignored. Further, design practice engages such situations in only a fragmented and tacit manner. This set of design situations lie between design problems as well-defined tasks and goals which characterizes most design today. For example, the designing of a mobile phone, and highly complicated problems for specialists such as designing a project management system for a mission to Mars are achievable outcomes. Yet what falls between these two examples on a continuum of design activities is a space that has been referred to as “wicked problems” (Rittel and Webber, 1973), and “ill-structured problems” (Newell and Simon, 1972), and as such are argued to be “unsolvable” or at best “coped” with by the designer. Examples of wicked problems included urban planning and environmentalism. These problems were seen to be complex due to the large scale and extraordinary uniqueness of the situations. Recently, ubiquitous computing has again raised the issue of complexity in design situations – situations that are beyond definable tasks and not of a specialist nature. Yet unlike the wicked problem, the very nature of ubiquitous computing challenges design in relatively small scale and in everyday design situations. Revisiting the role of complexity in design such that complexity is not of a special order represented in large scale and unique situations, but rather is seen as a ubiquitous attribute of design found across the continuum of design activities is a way forward. This research proposes the concurrent development of knowledge in theory and practice, involving the synthesis and augmentation of fragmented theories that approach complexity in design through ideas of reflective practice, context, embodied interaction, interactive cognition, and human activity. Like Qi, the immutability of material
and action is a framework for understanding design as a basic human activity. And while this paper argues from an exclusively theoretical point it will raise in the conclusion issues of our understanding of design and HCI practice that points to a broader understanding of design activity.

What Do We Know About Complexity In Design?

This research program connects an understanding of the practice of design with an understanding of complexity in design. The need for establishing this connection arises out of the author’s practice in interaction design, where design is approached socially and contextually, especially in the emerging applications of ubiquitous computing. Interaction design is understood to be an inter-disciplinary convergence of design and HCI (human-computer-interaction), inclusive of aspects of interactive art, performance, computing science, cognitive science, psychology and sociology (Sanders and Dandavate, 1999, Winograd, 1997, Löwgren, 2002, Precece et al., 2002). In the contexts of design and HCI, complexity has been discussed in isolated pockets along three dimensions, all of which assume a design process but rarely acknowledge it. Firstly, design outcomes are understood to be complex, as expressed in architecture, evolutionary theory, and human factors engineering (Dawkins, 1986, Norman, 1998, Venturi, 1966). Secondly, attention has been given to the definition of design problems as complex, from Rittel’s notion of the “wicked problem” to Simon’s “ill-structured problems” to Alexander’s “pattern language” (Rittel and Webber, 1973, Buchanan, 1995, Simon and Sliklössy, 1972, Alexander et al., 1977). Thirdly, there has been discussion on the role of HCI in supporting end users’ complex problem solving (Gay and Hembrooke, 2004, Albers and Mazur, 2003, Mirel, 2004). However, another trajectory is emerging, in which the term complexity is not explicitly used. This includes the ideas of reflective practice and contextual design. Design is seen to be boundless and dynamic rather than bounded and quantifiable (Buchanan, 1995, Barnard et al., 2000, Nardi and O’Day, 1999, Thackara, 2001, Dourish et al., 2004, Fischer, 2000, Schön, 1983). This latter trajectory is the starting focus of this research program.

Two key issues emerge from the current state of discourse on complexity and design. The fields of design and HCI are moving closer together and at times discussed interchangeably and at other times understood to be intertwined (Ehn, 1989, Norman, 1998, Fallman, 2003, Coyne, 1995, Gay and Hembrooke, 2004, Fischer, 2004, Zimmerman et al.,

---

1 Daniel Fallman contends that in HCI, there is little evidence of interest or concern for the designing of systems or prototypes. A syndrome he refers as “it just happens”, see Fallman, D (2003) Design-Oriented Human-Computer-Interaction, Proceedings of the conference on Human factors in computing systems, Ft Lauderdale FL, 225-232
Winograd was among the first to identify this trend (Winograd, 1997, Winograd, 1996). In large part, the motivating factor is the need to acknowledge the unique contextual aspects of interaction and the need to design in response to specific (typically complex) and not generic situations, a shift he coined as the move from machinery to habitat (Winograd, 1997). The second key issue is the lack of a coherent theory on complexity in design, especially inclusive of the practice of designing.

While many argue that design produces complex artifacts, and that design practice can be captured as complex formalisms, this research argues that we need to understand design as an activity that responds to situations of varying complexity. The key distinction is a question of understanding design as a prospective action, that is actively reflecting within a present moment on future action and contingency, as opposed to a retrospective event from which we view the design process or artifact as a stable past action with little attention to context. In the former, the relationship between activity and situation in design is integral and dynamic. For example, Schön views design as a conversation (Schön, 1983). Rittel understands design as argumentation (Rittel and Webber, 1973). In either case, each metaphor implies a dynamic act reliant on interpretation and multiple perspectives. The metaphors explicitly describe an activity in which the actions of speaking/listening, and the nature of what is being said/understood are intertwined and dynamically inform each other. In addition, like a conversation, design is quite ordinary and ubiquitous. And so, an alternate way to consider design is that it is an activity that is integrally related to complex, yet everyday situations.2

For example, a visit to a museum reveals an everyday, yet complex interaction situation. The factors within museum experiences are social, cultural, historical and psychological. The influences on the experience vary from the actions and previous knowledge of the visitor, visitor's learning style, the dynamics of others around them including friends, family and strangers. Naturally, the experience is affected by the presence of the artifacts and the relationships within collections as an outcome of institutional history, curatorship, exhibition design, and architecture. The time of day, duration of visit, room temperature and so on -- all have an impact. The experience can be characterized as multivariate, that is, it cannot be assessed by a single factor such as exhibit design, signage, or time spent in front of an artifact. Instead, the museum experience is subject to multiple

2 Dourish in discussing his related concept of embodied interaction has commented on the ordinairiness of conversation as discussed by Sacks in his analysis of conversation. Dourish argues that ordinairiness is a feature of context as it is understood within embodied interaction, see Dourish.

P (2004) What We Talk About When We Talk About Context, Personal and Ubiquitous Computing, forthcoming. This argument can equally be applied to complexity. What we then understand is that complexity is not a factor of scale (largeness) and extra-ordinairiness.
influences and results in multiple outcomes (Leinhardt and Crowley, 1998). Identifying a
design intervention that may have a direct and positive impact on experience is clearly not
easy! Many other complex situations have been discussed in design research such as how we
work (Ehn, 1989), seek information (Nardi and O'Day, 1999), learn (Gay and Hembrooke,
2004), and live in our homes (Tolmie et al., 2002, Bell and Kaye, 2002). Yet, few discussions
include how we might design for these situations.

In part this lack of exploration of design practice as a response to complex situations
is a result of the unwarranted focus given to viewing complexity in design as a quality of
artifact or process. Such views tend to abstract or isolate either the design object or process
from their context. To a large degree, this complexity in outcome is seen to be the result of a
complex design process. Such symmetrical linking of design process and outcomes is what
Gedenryd describes as an inaccuracy in today's understanding of design methodology; a
misconception that the structure of the process of design is necessarily reflected in the
outcome (Gedenryd, 1998).

Are design processes really complex or do we just assume that a complex outcome is
the result of a complex process? And therefore should we assume that a simple outcome is the
result of a simple process? Clearly we should not. I think we all understand that while many
design outcomes are complex artifacts and actions, many outcomes are extraordinarily simple.
And the reverse is true, simple processes can prove to be very effective. We have a tendency
to analyze design retrospectively as opposed to prospectively – and in the process over-
interpret for rational attributes such as logic and symmetry. If we did view design
prospectively, as in fact design practice demands, we would see that complexity precedes,
accompanies and follows design action. Complexity is contextual, situated and dynamic and
therefore cannot be isolated in processes or artifacts. That is design and designers' actions
respond to complex situations. What we find is that the process is not pre-determined as
complex, symmetrical or simple in structure, rather it is a dynamic process that is
improvisational and responsive to the changing design situation. An active stance is required
in design. Such design strategies have come to be understood as reflective, embodied, or
contextual in practice.

Another reason for the lack of research in design processes is that in many ways the
issues discussed in this research are part of a call for a new research agenda in viewing design
as contextual, social and diverse in strategies (Nardi, 1996, Dourish, 2004, Louridas, 1999). In
achieving that goal it is important to underscore the creative as well as reflective aspects of
such an undertaking. There is a need to generate a theory of the complex through the making
in design and there is a need to create or make concrete, the possibilities of the complex in
design.

A Call For Research Into Practice

Schön makes the argument against what he refers to as “technical rationality,” whose
basis lies in logic and reasoning in theory outside of practice. The rational in his view is the
logical abstraction of thought and action (Schön, 1983). Gedenryd, in analyzing design
methodologies, traces the rational view of design practice to math, logic and cognitive
science. In his arguments, design methodologies adopt a rational approach to design methods
characterized by the common principles of separation and sequence (Gedenryd, 1998). Design,
based on traditional cognitive science has to date, adopted the linear model of analysis,
synthesis and evaluation. To a large degree this approach has been inadequate in addressing
the ordinary complexity of design situations.

In contradistinction, is the practice we know only tacitly; an approach that guards
against abstracting views of practice, and rather grounds interpretation and reflection on
practice -- in practice? In Gedenryd’s terms, the approach is interactive and in the world,
rather than intra-mental. In other words, the movement from design decision to design
decision is like dead reckoning in navigation where one determines the next destination in
absolute reference to the last. And the movement between design choices is like rotating
puzzle pieces or Tetris pieces to find the right fit -- it can only be done through action. In
relation to Dourish, complexity is seen as an “interactional” issue where it is understood
through and by action, as opposed to a “representational” issue -- it simply cannot be mapped
out beforehand (Dourish, 2004). Situated in context of place, people, and technical systems,
the design approach encourages action to manipulate the context through playtypes such as
“lo-fi” prototypes, “Wizard of Oz” systems, and of course, directly through dialogue with
participants and designers (Wakkary, 2004).

However, certain limitations are evident in further research in design and HCI
practice. This includes further research in the analysis and observation of design practice.³
This would naturally entail work on the methodologies for documentation and analysis of
design practice, including first-person methodologies, ethnography in action, protocol
analysis, and other methods. It is by far more common to record interactions and real-time
observations of “users’ but not practitioners.

Discussion

Limitations to current theories related to complexity are readily admitted. The contextual approaches based on activity theory are quick to point out that the theory is a research framework for understanding individual consciousness that has strong implications for better understanding the nature of human activity and interactive technologies but says little about practice (Nardi, 1996, Gay and Hembrooke, 2004). The related ecological approaches have only begun to examine how to bridge the sociological methods of ethnography and participatory design with emergent systems that characterize information ecologies. Reflective practice has mostly remained a descriptive theory in design. In design, the nature of addressing complexity can be characterized by Alexander’s idea of “building over time” and the embodiment of design knowledge in pattern languages over time and through social use. What is not clear in Alexander’s pioneering work is again the method by which patterns occur (Alexander, 1979).

Coyne acknowledges the need to make operational the alternative methods to rational models, such as reflective practice (Coyne, 1995 p. 226). There is a clear need to move the theory of reflective practice in design beyond critique and into practice where it can be evaluated and considered fairly in relation to other approaches. Schön’s account of reflective practice in design rests on a single account of a conversation between an architecture teacher and his student (Schön, 1983 p. 226)! Needless to say, more accounts, case studies and examples of reflective practice in design are required.

Reflective practice relies on knowing-in-action as an embodied state of practitioner knowledge and theory. In effect it is a descriptor of tacit knowledge or know-how related to practice – all experienced practitioners carry this within them (Schön, 1983 pp. 50-51). Reflection-in-action arises through surprise when intuitive and spontaneous performance lead to unexpected results: “in such processes, reflection tends to focus interactively on the outcomes of action, the action itself, and the intuitive knowing implicit in the action” (Schön, 1983 p. 56). Generated theories, knowledge and actions as a response to the surprise and result of the reflection become new and unique to that design situation. States of knowing-in-action and reflection-in-action could be understood as attributes of many designers and only tell us part of the story of design. Excellent research in the area of design and reflective practice in the domain of design education have focused on reflection-in-action (Adams et al., 2003). However, the critical need for the mobilizing of the theory of reflective practice is in the domain of design practice and in pursuing the state of reflection-in-practice.
According to Schön, the role of reflection-in-practice is to ask: what is practice? Understanding the new and unique theories arising from reflection-in-action within a design situation is understood as part of a creation of practice - an epistemology of practice, based on divergence and complexity (Schön, 1983 p. 60). It is only on the level of reflection-in-practice that the theory can be compared with an understanding of design in the design science sense – as an alternative epistemology of design, and the methodological approach – as an alternative epistemology of design practice.

By way of a conceptual underpinning, the idea of complexity is proposed as a *framing experiment* for a translation, synthesis and augmentation of design practice and the role of context, human activity, and experience in HCI. Complexity is used as a descriptive term and not in the mathematical sense. The aim is not to quantify or create an abstraction of design practice based on the epistemologies of mathematics or science. I do not believe that is possible. Rather, complexity is proposed as a descriptive term based on the general understanding of the complex and the emergence of the term in design and HCI theory and practice.

References


Ron Wakkary is Associate Professor in Information Technology & Interactive Arts at Simon Fraser University in British Columbia. His research includes interaction design theory, and alternative design methods and practice for designing interactive systems. He has presented and published widely, including Computer Human Interaction ACM, Siggraph, Interact and Consciousness Reframed conferences. He currently serves on the Canadian Culture Online Advisory Board to advise the Minister of Canadian Heritage, and is on the Banff New Media Institute Research Advisory Board. He has lead design and technology research projects for the Canadian Nature Museum, Nokia Research, Museum of Modern Art, the Guggenheim Museum, and Electronic Arts Intermix.
Ontology and Rule based Retrieval of Sound Objects in Augmented Audio Reality System for Museum Visitors

Marek Hatala, Leila Kalantari, Ron Wakkary and Kenneth Newby
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada, V3T 2W1
Phone: +1-604-268-7431
{mhalala, lkalanta, rwakkary, newby}@sfu.ca

ABSTRACT
ce(h)o is an "augmented reality interface" utilizing spatialized soundscapes and a semantic web approach to information. The initial prototype is designed for a natural history and science museum. The platform is designed to create a museum experience that consists of a physical installation and an interactive virtual layer of three-dimensional soundscapes that are physically mapped to the museum displays. The source for the audio data is digital sound objects. The digital objects originate in a network of object repositories that connect digital content from one museum with other museums collections. The interface enables people to interact with the system by movement and object manipulation-based gestures without the direct use of a computer device. The focus of this paper is the retrieval mechanism for the sound objects for the museum visitor. The retrieval mechanism is built on the user model and conceptual descriptions of the sound object and museum artifacts in the form of ontologies for sound and psychoacoustics, topic ontology and Conceptual Reference Model for museum information. The retrieval criteria are represented as inference rules that represent knowledge from psychoacoustics, cognitive domain and composition aspects of interaction. The system will be demonstrated in exhibition space in Nature Museum in Ottawa in January 2003.

Categories and Subject Descriptors
H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval – information filtering, search process, selection process. H.5.1 [Information Interfaces and Presentations]: Multimedia Information – augmented reality, audio output

General Terms
Algorithms, Design, Experimentation, Human Factors

Keywords
Augmented audio reality, user model, ontologies, inference rules

1. INTRODUCTION
Audio museum guides have existed for some time as a means of overcoming the scheduling inflexibility of group tours by museum docents. While beneficial in many respects, the audio guides are limited by their linear sequence and non-interactive structure. Bederson [3] developed a prototype utilizing portable mini-disc players and an infra-red system to allow museum visitors to explore at their own pace and sequence. As museum visitors approached artifacts on display, relevant audio information would be triggered on the mini-disc player and heard through headphones. Hyperaudio [16] provided visitors with palmtop computers and developed specific user models for adaptive systems within a museum setting. MEG [2] is a portable digital museum guide for the Experience Music Project in Seattle that allows visitors 20 hours of audio and video on demand. Visitors make their selections either by use of the keyboard within the PDA device or by pointing the device at transmitters located adjacent to artifacts.

In the previous works, the relationship of the digital content to the artifacts is either pre-planned and fixed, or the digital content is not networked and limited to the local device, in some cases both limits are true. ec(h)o employs a semantic web approach to the museum's digital content thus it is networked, dynamic and user-driven. The interface of ec(h)o does not rely on portable computing devices, rather it utilizes a combination of gesture and object manipulation recognized by a vision system.

The dynamic and user-driven nature of ec(h)o requires a highly responsive retrieval mechanism with a criteria defined by psychoacoustics, content and composition domains. The retrieval mechanism is based on user model that is continually updated as user moves through the exhibition and listens to the audio objects. The criteria are represented by rules operating on the ontological descriptions of sound objects, museum artifacts and user interests.

The capturing of the user interests is in the center of the research of several disciplines such as information retrieval, information filtering and user modeling [21]. Most of the systems were developed for retrieval of documents where document content is analyzed and explicit user feedback is solicited to learn or infer the user interests. In the context of ec(h)o there is no direct feedback from the user. Ec(h)o can be categorized as a personalized system as observes user's behavior and makes generalizations and predictions about the individual user based on their interactions [11][18]. Our is an unobtrusive approach to
observation of user behavior, similar to the certain approaches to monitoring user browsing patterns [12][14] or user mouse movement and scrolling behavior [8].

The paper is organized as follows. We first describe how echo works and present an overall echo architecture. Next we provide details of the semantic object descriptions, retrieval criteria, user model and describe retrieval mechanism. Finally, we provide details on the current stage of the system implementation and conclude by highlighting our contributions.

2. EC(HO) ARCHITECTURE

The platform for ec(h)o is an integrated audio, vision and location tracking system installed as an augmentation of an existing exhibition installation. The platform is designed to create a museum experience that consists of a physical installation and an interactive layer of three-dimensional soundscapes that are physically mapped to museum displays and the overall exhibition installation.

Each soundscape consists of zones of ambient sound and "soundmarks" generated by dynamic audio data that relates to the artifacts the visitor is experiencing. The soundscapes change based on the position of the visitor in the space, their past history with viewing the artifacts, and their individual interests in relation to the museum collection. To achieve this type of audio experience the overall system must be integrated with a position tracking system that has a frequent update cycle and a high level of spatial resolution. A pattern of the user movement can indicate the type of the museum visitor [20] as well as user intentions [17].

When the user stops in front of an artifact, she is presented with three sound objects spatially positioned to the left, center and right. By way of a gesture-based interaction, the visitor can interact with a single artifact or multiple artifacts in order to listen to related audio information. The audio delivery is dynamic and generated by agent-assisted searches inferred by past interactions, histories and her individual interests. The source for the audio data is digital objects. In the case of ec(h)o, we developed a large sample set of digital objects that originated from the partner museums. These digital objects were used to populate the network of object repositories.

The ec(h)o architecture (Figure 1) consists of four independently functioning modules: position tracking module, vision module, sound delivery module, and reasoning module. Two main types of events trigger the communication between the modules: user's movement through the exhibition space and user's explicit selection of the sound objects.

3. SOUND OBJECT RETRIEVAL MECHANISM

One of the main goals of ec(h)o is to achieve an enhanced experience for the museum visitors without inserting an extra layer of technology between the visitor and the museum exhibit. Two mechanisms contribute to an accurate retrieval of sound objects in ec(h)o: the user model and ontology descriptions of objects. As mentioned above user's interaction space is limited to three sound objects. This poses extreme requirements on the retrieval mechanism as there is no recourse once the 'bad' choices are made.

3.1 Semantic Description of Objects

We have identified the following information as essential for ec(h)o:

- the content description of the user interests (user model), sound objects and museum artifacts
- psychoacoustics and sound characteristics of the sound objects
- sequencing models of an interaction

3.1.1 Ontologies for Describing Content

The interaction model is based on the semantic description of the content of the objects. We have developed an ontology where a sound object is described using several properties. As an ability to link to other museums is an important feature of ec(h)o our ontology builds significantly on the standard Conceptual Reference Model (CRM) for heritage content developed by CIDOC [5]. The CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation. To describe sound objects we use CRM TemporalEntity concept for modeling periods and events and Place for modeling locations. We describe museum artifacts using the full CRM model.

The content of the sound object is not described directly but annotated with three entities: concepts, topics, and themes. The concepts describe the domains that are expressed by the sound object such as evolution, behaviour, lifestyle, diversity, habitat, etc. Since the collections in individual museums are different so are the concept maps describing these collections. A topic is a more abstract entity that is represented by several concepts, such as botany, invertebrates, marine biology, etc. To facilitate the mappings between topic ontologies in individual museums we have mapped the topics to the Dewey Decimal Classification [7] whenever possible. Finally, themes are defined as entities supported by one or more topics. For example, the theme of bigness in invertebrates and marine biology.

Table 1 shows content related properties with their domains and ranges.

Table 1 shows content related properties with their domains and ranges.

1 To enable the system to relate sound objects to exhibition artifacts exhibition ontology defines exhibition artifacts as a subclass of an content object. Effectively this provides an exhibition object with the same content descriptive properties as sound objects.
Table 1 Content related properties

<table>
<thead>
<tr>
<th>Property</th>
<th>Domain</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasTheme</td>
<td>SoundObject</td>
<td>Theme</td>
</tr>
<tr>
<td>hasTopic</td>
<td>SoundObject</td>
<td>Topic</td>
</tr>
<tr>
<td>hasPrimaryConcept</td>
<td>SoundObject</td>
<td>Concept</td>
</tr>
<tr>
<td>hasSecondaryConcept</td>
<td>SoundObject</td>
<td>Concept</td>
</tr>
<tr>
<td>relatesToTemporalEntity</td>
<td>SoundObject</td>
<td>CRM.TemporalEntity</td>
</tr>
<tr>
<td>relatesToPlace</td>
<td>SoundObject</td>
<td>CRM.Place</td>
</tr>
<tr>
<td>describesArtifact</td>
<td>SoundObject</td>
<td>MuseumArtifact</td>
</tr>
</tbody>
</table>

3.1.2 Psychoacoustics and Sound Characteristics

The auditory interface of ech(h)o follows an ecological approach to the sound composition. It provides the basic mechanisms of navigation and orientation within the information space. Three areas are taken into account: psychoacoustic, cognitive, and compositional problems in the construction of a meaningful and engaging interactive audible display. Psychoacoustic characteristics of the ecological balance include spectral balancing of audible layers. Cognitive aspects of listening are represented by content-based criteria. Compositional aspects are addressed in the form of the orchestration of an ambient informational soundscape of immersion and flow that allows for the interactive involvement of the visitor.

Table 2 shows the psychoacoustics ontology that defines the characteristic of the sound objects that are used by the composition rules.

Table 2 Psychoacoustic properties for the Sound Object

<table>
<thead>
<tr>
<th>Property</th>
<th>Range Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasSpectralDensityCenter</td>
<td>&lt;number&gt;</td>
</tr>
<tr>
<td>hasSpectralDensityWidth</td>
<td>&lt;number&gt;</td>
</tr>
<tr>
<td>hasBandwidth</td>
<td>&lt;number&gt;</td>
</tr>
<tr>
<td>relatesToEnvironment</td>
<td>Physical_ENVIRONMENT</td>
</tr>
<tr>
<td>relatesToEvent</td>
<td>CRM_Event</td>
</tr>
<tr>
<td>hasSource</td>
<td>SourceTypeValue (i.e. AnimalSound, HumanENVIRONMENTSound)</td>
</tr>
</tbody>
</table>

3.2 The User Model

In the core of the ech(h)o’s reasoning module is a user model [21] that is continually updated as user moves through the exhibition and listens to the audio objects.

Figure 2 shows an interaction schema of the user model with other modules. There are two main update sources in the system. First, as the user moves through the exhibition the speed of the movement and stops or slowing down at different artifacts provide updates to the user model. The user behaviors are computed based on the speed and homogeneity of the user movement. The stops and slowing down in front of an artifact are interpreted as an interest in topics represented by the artifact. The user interests and intentions influence the presentation of soundmarks. For example, soundmark radius and volume is increased for those artifacts that correspond with current user interests. Another example can be reducing the number of soundmarks in the exhibition if user’s recognized intention is to quickly cross the room.

Second type of information stored in Interaction History is user's selections in the form of URLs of sound objects selected by the user.

This information is essential for several tasks ranging from simple avoidance of the delivery of redundant narratives to updating user interests.

3.2.2 User Behavior

The user behavior in the museum context is well studied in curatorial science [6] and used in several systems personalizing the user experience [19][20]. Several categorizations were used, for example one user may go through almost every artifact that is on his/her way, and another user may be more selective and chooses artifacts that have certain concepts. Our categorization of user types is based on Sparacino’s work [20] and it classifies users to three main categories. These categories were validated by our own research of site studies and interviews with staff at our partner museums:

- The greedy type wants to know and see as much as possible. He is almost sequential, and does not rush.
The selective type explores artifacts which represent certain concepts, and wants to dig into those concepts only.

The busy type does not want to spend much time on a single artifact and wants to stroll through the museum to get a general idea.

In cc(h)o, the user behavior is not static but is updated every minute. The rules consider the location data from user history accumulated within 3 minute interval and topics of previously selected sound objects.

### 3.2.3 User Interests

The interests for the user are represented as a set of facts where each fact represents a single interest and its relative level

\[
\text{(user-interest (user-id john) (concept evolution) (level strong))}
\]

As described in the previous sections, each artifact/exhibition is associated with a set of concepts. The sound objects address a set of particular concepts as well. The interaction of the user and artifacts and sound objects is stored in the Interaction History that together with the user behavior type are used to infer the user's interests. The following principles for the user interest inference are implemented using the reinforcement learning approach [13]:

- If a greedy type user slows/stops in front of an artifact, we can infer that the user is interested in any of general concepts represented by the artifact. If the user continues with his greedy behavior in front of that artifact, his interests are updated with related concepts from sound objects selected (not necessarily closely related).

- Interests of a selective user do not get easily overwritten. If a selective user is moving slowly in front of an artifact he is not interested in, one of his previous interests is overwritten by a concept that is "close" to his previous interests. If a selective user slows in front of an artifact he is not interested in, one of his previous interests is overwritten by a concept that is represented strongest by the artifact.

- If a busy user slows/stops in front of an artifact, several of his interests are overwritten by general concepts that are also represented strongly by the artifact.

- If a user's behavior is not categorized yet, User Interests can be any general concepts that are strongly represented by the artifact the user slows/stops in front of.

We limit the number of concepts represented in the user model as can • If a user's behavior is not categorized yet.

\[
\text{User Interests}
\]

\[
\text{The interests for the user are represented as a set of facts where each fact represents a single interest and its relative level}
\]

\[
\text{(user-interest (user-id john) (concept evolution) (level strong))}
\]

As described in the previous sections, each artifact/exhibition is associated with a set of concepts. The sound objects address a set of particular concepts as well. The interaction of the user and artifacts and sound objects is stored in the Interaction History that together with the user behavior type are used to infer the user's interests. The following principles for the user interest inference are implemented using the reinforcement learning approach [13]:

- If a greedy type user slows/stops in front of an artifact, we can infer that the user is interested in any of general concepts represented by the artifact. If the user continues with his greedy behavior in front of that artifact, his interests are updated with related concepts from sound objects selected (not necessarily closely related).

- Interests of a selective user do not get easily overwritten. If a selective user is moving slowly in front of an artifact he is not interested in, one of his previous interests is overwritten by a concept that is "close" to his previous interests. If a selective user slows in front of an artifact he is not interested in, one of his previous interests is overwritten by a concept that is represented strongest by the artifact.

- If a busy user slows/stops in front of an artifact, several of his interests are overwritten by general concepts that are also represented strongly by the artifact.

- If a user's behavior is not categorized yet, User Interests can be any general concepts that are strongly represented by the artifact the user slows/stops in front of.

We limit the number of concepts represented in the user model as can • If a user's behavior is not categorized yet. User Interests can be any general concepts that are strongly represented by the artifact the user slows/stops in front of.

\[
\text{We limit the number of concepts represented in the user model as can • If a user's behavior is not categorized yet. User Interests can be any general concepts that are strongly represented by the artifact the user slows/stops in front of.}
\]

This is an example of a rule that computes interests of a greedy user who just stopped in front of an artifact:

\[
\text{(defrule get-greedy-user-interest}
\text{ (time ?current-time)}
\text{ (user-behavior (user-id ?user)}
\text{ (behavior greedy))}
\text{ (or (is-slow ?user ?current-time)}
\text{ (stopped ?user ?current-time))}
\text{ (just-came-in-front ?user ?a)}
\text{ (object-model (object-name \(a) (x-position ?x)}
\text{ (y-position ?y) (radius \(r)))
\text{ (assert}
\text{ (user-interest (user-id ?user) (concept \(i)))
\text{ (has-concept ?object \(i) strong))}
\text{ =>)}}
\]

### 3.3 Sound Object Retrieval

We have identified the following requirements for the retrieval of appropriate sound objects:

1. Content-relevant to the viewed artifact
2. Content-relevant to the user interests
3. Inviting to explore other areas
4. Plausible from the psychoacoustics perspective

In addition to the criteria for an individual objects the following criteria apply to the sequence of the objects offered to the user:

5. Provide for exploration of a subject in depth
6. Provide for the fluidity in experience both in content and sound experience
7. Provide a mix of informational and entertaining objects

#### 3.3.1 Retrieval Process

The retrieval process in cc(h)o can be broken into four steps as illustrated in Figure 3.

First, the system determines the candidate concepts as an overlap between user interests and concepts represented by the museum artifacts. The candidate concepts are ranked by a combination of the level of the interest of the user and how strongly they are represented by the artifact.

In the second step the candidate concepts are used in the simple pattern matching algorithm to retrieve semantic descriptions of the information sound objects. The temporal and location properties of the artifact are used to narrow the search to sound objects that are closely related to the presented artifacts.

While the first two steps considered objects as independent acts, the rules in the next step, the content related composition criteria are applied. The criteria consider the next object in the context of the previous objects the user listened to before. The selection is based on theme, topic, concepts, and described artifacts. The relative weight of each type of composition criteria depends on the user type. For example, for the "greedy" user concepts/topics, and 'described artifacts' are of equal criterion to enable the system to offer a wide range of audio objects. For the selective user, the artifact criterion is dominant since the user is very selective among the artifacts on display.

Finally, the background sound objects for each information objects are retrieved using psychoacoustics criteria and the psychoacoustics rules are applied to finalize the choice of the sound objects. For example, if neither event nor environment is specified then both place and temporal information to infer

2 There is no overlap between artifact concepts and user interests.
environment type and use it for selecting the background sound object.

3.3.2 Implementation

The reasoning module is fully implemented with all features described in the previous section and embedded into the Tomcat environment. The Figure 4 shows the implementation schema with Jess Inference engine in the center of the reasoning module. DAMLjessKB converts DAML+OIL ontologies to Jess facts. Reasoning module is connected with other modules through the UDP socket connections and communicates with other museums via SOAP based protocol [10]. We have developed a web-based tool for editing of ontological descriptions of sound objects that generates forms based on the ontology definition by direct querying of the ontologies loaded into the Jess inference engine.

Figure 4 Implementation schema of the reasoning module

The use of a forward chaining inference engine has proved itself to be an efficient mechanism for responding to the dynamic nature of the user input. The system loading time is relatively long as a lot of initial inference is performed on the ontologies and object descriptions. After the startup phase the amount of the inference is limited to the updates from user input resulting in fast responses. Although more extensive testing still needs to be done the pattern of this behavior makes us optimistic with regard to the scalability of the system.

4. NETWORK OF MUSEUMS

One of the main features of the ec(h)o system is that it enables the user to experience the richness of the museum collections located not only in the visited museums but also from the other linked museums. For example, a visitor standing in front of a bear specimen in Nature Museum in Ottawa can listen to the sound object about the role of the bear in the mythology of aboriginal tribes on the West Coast retrieved from the Museum of Anthropology in Vancouver.

Two aspects are critical for fluid retrieval and access of sound objects from other museums: protocol compatibility and semantic mapping between conceptual structures. We addressed the protocol compatibility issue by reusing the infrastructure and protocols our group developed for connecting learning object repositories [9][10]. The only difference is that instead of learning object metadata we share the sound object semantic descriptions.

As different museums can have different conceptualizations of the topics covered by their stories the problem of mapping between these conceptualizations need to be addressed. First, we use the standard Conceptual Reference Model for describing temporal and spatial entities which allows us to relate sound objects to time and space. Second, we use Dewey Decimal Classification as an intermediary for mapping between museum specific topic maps. Although this does not provide for an exact mapping our solution is acceptable in the museum setting where the exploration aspect of the user experience dominates the in-depth learning aspect.

5. CONCLUSIONS

In this paper we presented retrieval mechanism used in an augmented audio reality system for museum visitors named ec(h)o. Each visitor’s experience is tailored to the user interests. The user interests are inferred from the user movement through the exhibition as well as from the visitor’s interaction with the sound objects. The sound objects are retrieved based on their relevance to the user interests, narrative criteria and psychoacoustic criteria. Ec(h)o uses ontologies to describe concepts, temporal and spatial characteristics, psychoacoustic and sound characteristics of sound objects. The retrieval mechanism is represented in form of the rules that capture contextual, sound, psychoacoustic and composition criteria for plausible user experience.

The system is a result of convergent research streams from research in object repositories, interaction design, auditory display, knowledge representation, and information retrieval. The ontologies combined with the rule based inference proved to be a powerful implementation platform well suited for this type of the systems. We believe this has enabled us to extend works cited through the paper in several directions. First, it extends the work of the Alfaro et al. work [1] by building rich model of the concepts represented by the sound objects. In ec(h)o, the content presented to the user is not pre-processed for possible linkages as in the systems using Rhetorical Structure Theory [22]. Our approach replaces pre-processed linkages with a retrieval mechanism based on composition and interaction criteria formulated in the form of the rules and applied to semantically described independent objects. This allows ec(h)o to create a network of museums sharing objects and providing richer user experience.

6. ACKNOWLEDGEMENTS

Work presented in this paper is supported by Canarie Inc. grant under E-Content program. Authors would especially thank to the Mark Graham and his colleagues in the Nature Museum in Ottawa for the enthusiastic support to this project. We would also like to provide that information is relevant to the temporal and spatial aspects of a museum artifact.
thank our colleagues and participants in several workshops that contributed to the development of the project, namely Dale Everdén, Doreen Leo, Gilly Mah, Robb Lowell, Mark Brady, Jordan Williams and Phil Thomson.

7. REFERENCES


[17] Schliedcr, Ch., Vogele, l., Werner, A. Location Modeling for Intentional Behavior in Spatial Partonomies Ubicomp 2001, Workshop "Location Modeling for Ubiquitous Computing", Atlanta, Georgia, USA, pp.63-70


Cross-Dressing And Border Crossing: Exploring Experience Methods Across Disciplines

Ron Wakkary, Thecla Schiphorst, Jim Budd
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada, V3T 2V1
{rwakkary, thecla, jbudd}@sfu.ca

Abstract
As designers of interactive systems (spaces, process and products for people), we find ourselves stretching the limits of methodological structures that enable us to explore, build, communicate, and prototype experience. This workshop aims to investigate divergent disciplines that each contains rich knowledge and rigorous methodologies for addressing human experience in interactive systems.

Categories & Subject Descriptors: H.5.2 [Information Interfaces and Presentation]: User Interfaces — Theory and methods; H.53 [Information Interfaces and Presentation]: Group and Organization Interfaces — Theory and models; H.1.2 [Information Systems]: User/Machine Systems — Human factors.

General Terms: Design, Human factors, Theory

Keywords: experience design, interaction design, interdisciplinary methods, prototyping.

INTRODUCTION
As designers of interactive systems (spaces, process and products for people), we find ourselves stretching the limits of methodological structures that enable us to explore, build, communicate, and prototype experience. We argue that designing experience requires a 're-dressing' of methodological practice, and that HCI can benefit from drawing on methodological frameworks that traditionally fall outside of its purview. Domains such as performance, theatre, dance, architecture, conceptual design, industrial design, and visual art each contain rich knowledge and rigorous methodologies for constructing experience. Each of these domains defines experience, experience qualities and attributes, and defines affordances for enacting [and re-enacting] experience as a fundamental methodological tool in the respective discipline.

We invite participants from multiple disciplines across and within HCI, including kinesiology, performance, visual art, architecture, anthropology, organizational research, computing science, visualization and engineering. Participants are expected to be practitioners exploring unique methodological frameworks for designing technologically mediated experiences that live in technologically mediated environments. Participants will be expected to share, explore their methodologies for constructing and designing experience. Our fundamental assumption is that experience matters. We assume that an understanding, exploration and sharing of experience design is central to HCI. Building experience is an interdisciplinary practice, we invite participants to share and explore the diverse practices that contribute to the evolution of methodologies for designing experience.

GOALS OF THE WORKSHOP
The focus of this workshop is to cross boundaries, assume other roles in order to experiment methodologically and to establish a new common knowledgebase aimed at design and human experience. We see this as a step toward establishing a community of practice within HCI. We propose the following key issues as points of departure and exploration during the workshop:

- In today's HCI landscape, experience is felt, defined and modeled across multiple media and disciplinary domains, and environments. This provides a scope challenge that requires creative solutions derived by a diverse community of practice.
- Members of this community can engage each other in a cross-disciplinary dialogue around the task of creating positive "user experiences".
- In doing so each practitioner sits at the experience design table with a slightly different set of assumptions, knowledge, methodology and context around what it means to consider user experience.
- The considerations related to user experience in each discipline are unique and valuable in their own right. It is important to recognize this and embrace alternate perspectives.

THE WORKSHOP ACTIVITIES
The workshop will be divided into three main parts with the key goals of finding a more common language around problem setting, hybridizing practices for the development of criteria for new methods, and reflecting on the cross-disciplinary practices of each team.

Part 1. Problem setting: Organizers and participants will present and review several of the experience scenarios. Activity and discussions will center on developing a set of shared analysis and language for defining and problem
setting interaction experiences. In addition to discussions, organizers expect group activities in the form of role-playing, re-enactments and re-articulations as a form of analysis.

Part 2. Practice and play: Teams will brainstorm and “prototype” new methods that could address the understanding of the problem articulations that emerged in part 1. The activities will shift from structured “brainstorming” to open ended development of a method within a condensed period of time. The activity will end with a “swapping” of methods to be used by another team to address the problem situations form part 1.

Part 3. Reflection and mirror-gazing: A key goal of the workshop is to identify criteria for new methods while also identifying the rich and diverse set of practices that can be pulled in within HCI in order to respond to experience interaction situations. Teams will be asked to discuss and report out on three key items: criteria for methods, identification of the intertwining of practices within their methods and methods from other teams, identify key disciplinary and non-disciplinary connections within the teams and in other teams. The workshop in plenary will discuss the reports as a possible group report that identifies issues of methods, cross-disciplinary knowledge sets, and key relationships and connections that could form the basis of a community of practice centered on human experience.

RELATED LITERATURE

Terry Winograd was among the first to identify a design practice whose outcome and focus was a qualitative process rather than a “thing” or an object [14]. He labeled this new practice as “interaction design”. Winograd identified the need to focus on the perceptual and psychological aspects of human experience by rooting interaction design equally in graphic design, psychology, communication, linguistics and computing science. A key genesis point in the evolution of “experience” as a design concept is the work in the 1930s of the industrial designer Henry Dreyfuss [3]. Dreyfuss’ work in ergonomics lead to the publication of the “Measure of Man”, an extensive database of human measurement to facilitate the design of products tailored to a “standardized” human body. In the late 1960’s ergonomics split into the related science and kinesiology based field of human factors, the political and social movements in Scandinavia that became known as participatory design [4, 8], and the cognitive science and design methodology of user-centered design [11]. Design experience was seen in surprisingly different lights, one functional the other social and political. Enabling the audience experience was also a key goal of theorists and practitioners of the fields of performance and theater, namely the Russian, Vsevolod Meyerhold [1], and later the work of theorist and theater director Jerzy Grotowski [6]. This tradition directly informed the concepts of interactive design from the early work of Norman Bel Geddes [9] to today’s interactive technology experiences and environments [2, 10]. In the field of computing science, particularly in the field of HCI (Human Computing Interaction), experience design is viewed as an extension of user-centered design methods [7, 13]. This approach has a particular focus on the “User Experience” aspect of design, in particular, quantifying the interactive experience as a means to determining standards for interface and interaction design. On a methodological note, some of the framework of this workshop is indebted to the work Donald Schön and Henrik Gednryd [5, 12].

REFERENCES

7. IBM. What is user experience design?, IBM, 2002.
Interactive Audio Content: An Approach to Audio Content for a Dynamic Museum Experience through Augmented Audio Reality and Adaptive Information Retrieval

Ron Wakkary, Kenneth Newby, Marek Halala, Dale Evernden and Milena Droumeva
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada, V3T 2W1
Email: rwakkary@sfu.ca

Abstract

ece(h)o is an "audio augmented reality interface" utilizing spatialized soundscapes and a semantic web approach to information. The paper discusses our approach to conceptualizing museum content and its creation as audio objects in order to satisfy the requirements of the ece(h)o system. This includes, the conceptualizing of information relevant to an existing exhibition design (an exhibition from the Canadian Museum of Nature in Ottawa). We will discuss the process of acquiring, designing and developing information relevant to the exhibition and its mapping to the requirements of adaptive information retrieval and the interaction model. The development of the audio objects is based on an audio display model that addresses issues of psychoacoustics, composition and cognition. The paper will outline the challenges and identify the limitations of our approach.

Keywords: augmented reality, audio, adaptive information retrieval, museum guide

1. Introduction

Museums are a natural laboratory in which to examine the complex nature of constructing meaning, learning from and enjoying objects and environments through interaction. Museum visits have been described as interactive, situational, social, subjective and inter-connected with the physical environment (Leinhardt and Crowley 1998; Lehn, Heath et al. 2001). Research and commercial practice in the development of electronic museum guides have typically focused on the use of portable computing devices for interaction, data storage and audio delivery (Proctor and Tellis 2003). Despite the growing use of such systems there are known limits to this approach. These include cognitive and learning difficulties of using a new graphical interface, competition for attention between the device and its surroundings, and the ergonomic problems of weight and operation (Woodruff, Aoki et al. 2001; Proctor and Tellis 2003). Often portable computing based systems deliver content in ways familiar to computing but not familiar to museum visitors (Leinhardt and Crowley 1998; Lehn, Heath et al. 2001; Woodruff, Szymanski et al. 2001). An arguably more important limitation in current practice is the approach to digital content. Typically content for museum guides has been developed much like CD-ROM content, interactive but finite and limited structurally in terms of associations and linkages.

Our goal is to design a system that fits with the interactions and everyday competencies of the museum visitor, such that it amplifies and strengthens the visitor's ability to explore, learn from and construct the meaning of exhibitions.
The paper will discuss our approach to conceptualizing the content and its creation as audio objects in order to satisfy the requirements of ec(h)o. This includes, the conceptualizing of information relevant to an existing exhibition design (an exhibition from the Canadian Museum of Nature in Ottawa). Firstly, we will provide an overview of the ec(h)o system and interaction model, followed by a discussion of challenges in relation to adaptive information retrieval and interactive audio. Next, we will discuss the process of acquiring, designing and developing information relevant to the exhibition and its mapping to the requirements of adaptive information retrieval and the interaction model. The development of the audio objects is based on an audio display model that addresses issues of psychoacoustics, composition and cognition. The paper will outline the challenges and identify the limitations of our approach.

2. Context

2.1 Overview of ec(h)o

The platform for ec(h)o is an integrated audio, vision and location tracking system installed as an augmentation of an existing exhibition installation. The platform is designed to create a museum experience that consists of a physical installation and an interactive layer of three-dimensional soundscapes that are physically mapped to museum displays and the overall exhibition installation. Each soundscape consists of zones of ambient sound and "soundmarks" generated by dynamic audio data that relates to the artifacts the visitor is experiencing. The soundscapes change based on the position of the visitor in the space, their past history with viewing the artifacts, and their individual interests in relation to the museum collection. By way of a gesture-based interaction, the visitor can interact with a single artifact or multiple artifacts in order to listen to related audio information. The audio delivery is dynamic and generated by agent-assisted searches inferred by past interactions, histories and individual interests.

The source for the audio-data is digital objects. Our original sample set of digital objects was developed using content that originated from our partner museum, the Canadian Museum of Nature. In the ec(h)o context digital objects populate a network of repositories linked across different museums. The networked nature of these repositories makes it possible for visitors in the context of one museum to access data from another. For example, a visitor at the Canadian museum of Nature could access content from the local repository as well as repositories of other museums or online resources.

The ec(h)o architecture consists of four main system components: position tracking, vision system, wireless audio delivery, and reasoning. Two main types of events trigger the communication between the components: user's movement through the exhibition space, and user's explicit selection of the sound objects. A more detailed description and analysis of the technical and information retrieval aspects can be found in our previous writing (Hatala, Kalantari et al.).

2.2 Interaction Model

2.2.1 Conversation structure

Similar to the work of Woodruff, Aoki, Hurst and Szymanski (Woodruff, Aoki et al. 2001; Woodruff, Szymanski et al. 2001; Aoki, Grinter et al. 2002), we have adopted the storytelling structure based on Sacks' conversation analysis theory (Sacks 1974). In our case, we modeled the system and interaction on this conversation structure. ec(h)o offers the visitor three short audio pieces that we refer to as audible icons. These audible icons serve as préfaces. They are in effect offering three turn-taking possibilities to the visitor. The visitor selects one and the system delivers the related audio object. This turn-taking represents the telling phase. After the delivery of the object, the system again offers three audible icons. It is at this stage that the response phase occurs. The visitor's response is expressed through the gesture selection with the interaction object. Additionally, the system may be met by no response, because the visitor does not wish to engage the system. It will then enter into a silent mode. The visitor may also have moved away and the system will initiate a soundscape and prepare for the next conversational encounter.
2.2.2 Navigational model

It is important to explain the navigational model for both its novelty and simplicity, and of course for its support of the interaction. The audio objects are semantically tagged to a range of topics. At the beginning of each interaction cycle or "conversation", three topics are inferred to be more relevant than others to the visitor based on their user model, location and interaction history. Audio objects are cued representing each of the three chosen topics. Audible icons or prefaces related to the objects are presented to the visitor (each audible icon is differently spatialized in the audio display for differentiation), a visitor chooses one of the prefaces and listens to an object representative of the topic chosen. The topics are not explicit to the visitor, rather the consistency and content logic is kept in the background. After listening to the object, the visitor is offered a new preface based on their previous topic selection. The two previous prefaces that were not selected are offered once again. If after three offerings of the same preface and topic have transpired without selection, that topic is replaced. A more detailed description and analysis of the interaction model and design process can be found in a previous writing (Wakkary, Hatala et al.).

Figure 1. 1-2-4 navigation model

3. Challenges

In related works (Bederson 1995; Sarini and Strapparava 1998; Andolesk and Freedman, 2001), the relationship of the digital content to the artifacts is either pre-planned and fixed, or the digital content is not networked and limited to the local device, in some cases both limits are true. ec(h)o employs a semantic web approach to the museum's digital content thus it is networked, dynamic and user-driven. The interface of ec(h)o does not rely on portable computing devices, rather it utilizes a combination of gesture and object manipulation recognized by a vision system.

The dynamic and user-driven nature of ec(h)o requires a highly responsive retrieval mechanism with a criteria defined by psychoacoustics, content and composition domains. The retrieval mechanism is based on a user model that is continually updated as the user moves through the exhibition and listens to the audio objects. The criteria are represented by rules operating on the ontological descriptions of sound objects, museum artifacts and user interests.

The capturing of the user interests is in the center of the research of several disciplines such as information retrieval, information filtering and user modeling (Wahlster and Kobsa 1989). Most of the systems were developed for retrieval of documents where document content is analyzed and explicit user feedback is solicited to learn or infer the user interests. In the context of ec(h)o there is no direct feedback from the user. ec(h)o can be categorized as a personalized system as it observes user's behavior and makes generalizations and predictions about the individual user based on their interactions (Goecks and Shavlik 2000; Kobsa and Fink 2002).

Particular challenges in relation to the use of audio in ec(h)o include: the designing and preparing of the audio objects for dynamic and personal delivery; the information management aspects of developing classifications and relationships; and the ultimate need to create an audio display and user experience that is coherent, consistent and pleasurably exploratory in relation to an existing exhibition. The following section focuses on how we addressed these challenges.

4. Audio Content Design and Development

We will describe the design and production of the audio content in fours stages: 1) Our expert-based system approach to data collection, describing how we acquired the raw information related to the exhibition and artifacts; 2) Concept mapping and audio object design, discussing the initial knowledge management design of the information and the design and development of objects; 3) Design of audio objects, description of the audio display and acoustical experience issues related to the objects as audio; 4) User scenarios and inference rules, a discussion of our development of user scenarios as a design approach to developing the inference rules. This set of descriptions outlines the entire process of the design and development of the audio objects, the ontologies, and inference rules.
4.1 Stage One - Expert Based System Approach for Data Collection

In order to acquire the relevant information for the audio objects we devised two modes of interviews with researchers at the Nature Museum of Canada. The interview sessions took place in the museum over the course of several days. Our goal was to develop an information gathering process that paralleled our conversation approach to the interaction model. The aim was that the interview processes would provide us with audio material that could be used directly and would create the experience for the museum visitor of an interactive guide to the museum with a group of different 'experts' (in the end we used the interview texts to create a script and so we did not use the recordings directly). In keeping with the our conversation model we hoped to emulate the experience of experts conversing (both with themselves and the visitor), each taking turns contributing bits of information based on their particular interests and area of expertise.

We organized interviews with members of the museum research staff. These individuals were chosen based on their expertise in a number of different knowledge domains related to the exhibition -- Zoology, Ichthyology, Botany, Vascular Plants, Invertebrates, conservation, etc. The interviews were conducted in two parts: part one, introduced the interviewee to the ec(h)o project and asked them to comment or provide contextual information from their perspective and area of expertise related to the exhibit; phase two, involved a video walk-through of the exhibit space in which the interviewer and expert engaged in a discussion of the artifacts and collections on display. Here interviewee's were asked to provide discipline specific information about the exhibits themes and sub-themes, as well as relevant information about specific artifacts with in each of the exhibits. Here is a sample set of questions:

- Each display tells a story, what is that story?
- Can you discuss the different groupings of the artifacts and explain how and why they are clustered?
- Can you describe the significance of each artifact or group of artifacts?
- What makes these particular artifacts best suited to the their task?
- Can you describe type of sounds that you think would supplement this exhibit?
- How might these sound effects work to enhance the visitor experience?
- Can you speak to the potential of linking content to other museums?

The results of the interviews were largely successful, however there were problems and gaps in our information set. Some interviewee’s limited their discussion to high-level explanations of the exhibit that were difficult to integrate into a museum visit, while others provided interesting anecdotal information about artifacts. While we wanted to avoid an "encyclopedia" approach to the information, we supplemented the interview information gathered with research of the museums archives and research collections. We met with archival experts to filter through potential source material that already existed. Source material was, for the most part, limited to audio tracks taken from studies conducted in the field, as well as video productions that the museum had collected or produced over the years.

4.2 Stage Two - Concept Mapping & Ontologies

4.2.1 Concept map development

In order to translate the information gathered in the interview process for adaptive retrieval, we needed to conceptualize the information within a loose taxonomy or concept map that could eventually be developed into semantic web ontologies. The concept map would guide us in the design and relationships of the information in the form of digital objects. As part of the information management related research, the strength of a semantic web approach is the interoperability of generic and specific topic ontologies. We wanted to test the ability to develop specific ontologies that could function with generic ontologies. In addition, we were very aware that the existing curatorship and exhibition design represented a knowledge map in its own right relevant to the objects and collections on display,
nevertheless our goal was to insert another level of knowledge mapping that could be productively “superimposed” over the existing exhibition.

In order to develop the concept map for the ec(h)o version of the exhibition we analysed the recorded video and audio from the expert interviews. This analysis entailed watching and listening to video and audio followed by a mapping process. This was undertaken by the entire interdisciplinary research team which helped ensure that the design of the concept map could function in the different contexts of adaptive retrieval, audio display and user experience. The concepts and themes that the team clustered were organized into a relational map. These concepts and themes became classifiers that were used during the meta-tagging stage of audio object development. Conceptual and thematic classifiers evolved out of the concept mapping exercise, where as the topical classifiers were taken from the established Dewey Decimal Classification system. The conceptual map served as an important visualization tool that helped the team understand the topical and conceptual links between artifacts, and exhibit sections. The map also served as a point of departure for helping the team recognize potential openings for bringing in content from other museums. The concept map was also the starting point for the development and adoption of different ontologies.

Figure 2. Preliminary concept mapping

4.2.3 Ontologies

The interaction model is based on the semantic description of the content of the objects. We have developed an ontology where a sound object is described using several properties. As an ability to link to other museums is an important feature of ec(h)o our ontology builds significantly on the standard Conceptual Reference Model (CRM) for heritage content developed by CIDOC (Crofts, Dionissiadou et al. 2001). The CRM provides definitions and a formal structure for describing the implicit and explicit concepts and relationships used in cultural heritage documentation. To describe sound objects we use CRM TemporalEntity concept for modeling periods and events and Place for modeling locations. We describe museum artifacts using the full CRM model.

The content of the sound object is not described directly but annotated with three entities: concepts, topics, and themes. The concepts describe the domains that are expressed by the sound object such as evolution, behaviour, lifestyle, diversity, habitat, etc. Since the collections in individual museums are different so are the concept maps describing these collections. A topic is a more abstract entity that is represented by several concepts, such as botany, invertebrates, marine biology, etc. To facilitate the mappings between topic ontologies in individual museums we have mapped the topics to the Dewey Decimal Classification whenever possible. Finally, themes are defined as entities supported by one or more topics, e.g. the theme of “bigness: in invertebrates and marine biology.

4.3 Stage Three - Audio Object Design

In this stage, the aim was to design and develop the audio objects that supported the interaction and audio display model, and that could be classified and meta-tagged based on the concept mapping. In support of the interaction model, the audio objects would need to be different types of audio elements—prefaces, audio objects, sound-marks, and keynote sounds. In the early stages of this work we focused on developing audio objects, and their corresponding prefaces—sound mark, and keynote sound production would come later. The production of the audio objects started with dividing up the interviews into manageable information objects. In doing so it was clear that each object needed to be cognitively manageable for the user, as well as manageable for the system, which meant it needed to be classifiable. Embedded references to artifacts were either made explicit or removed all together, and the scripts for all objects were edited to be suitable, as well as interesting, to as broad an audience as possible.

Once refined, each discrete audio object was then entered into the repository database where they were meta-tagged for retrieval purposes. Meta-tagged information includes location and associative information such as the exhibit an object belonged to, as well as the specific artifact it was most relevant to. Objects were also meta-tagged based on their topical, conceptual, and thematic qualities. For example, if we were classifying an object that spoke about the collecting tools used by early plant
collector Catherine Parr Trail, its topical classifier would be “botany”, its conceptual classifier “tools and techniques,” and its thematic classifier, “early collectors”.

4.3.2 Audio Display
In order to deliver a seamless integrated audio display experience, ec(h)o works on several levels—the first mode of interaction involves a movement-related immersion in a dynamic soundscape, related thematically to different parts of the exhibit; another, second mode of interaction, the visitors engage with the audio display installation via a manipulation object, responding to spatially displayed audio prefaces; and a third level is knowledge acquisition or learning, by listening to the audio knowledge objects. It is important that all levels work together, physically, cognitively and psychoacoustically in order to deliver a worthwhile immersive experience. Issues of sound amplitude and frequency range must be considered for all elements of the audio display system.

In addition, we felt it was important to provide the visitor a variety of voices with a spectrum spanning the serious and authoritative to the playful and whimsical. Before recording the audio objects, consideration was given in choosing the voice talents to perform the scripted content. Issues of gender, voice quality, timbre, clarity and other psychoacoustic sound markers came into play. For ec(h)o, an even gender split between the voices is used with care in order to develop differences in both timbre and performance and to facilitate easy discrimination due to variations in range, frequency and timbre. The voices consist of one deeper, broad-range strong male voice, one warmer timbre, softer male voice, one mid-high pure female voice, and one deeper, richer timbre female voice. For an initial database of just over 200 short sound objects, four different voices (two male and two female) appeared to be sufficient to provide the diverse, yet consistent and recognizable audio web of information.

In order to create an atmosphere of ‘engaged and fun learning,’ the aural design attempts to stay away from a highly accented ‘authoritative’ presentation of the museum information. For this reason, voice talents used are not professionals, but ‘real’ people. The style of narration determined during recording is natural pace, moderate inflection, with an even dynamic speech envelope in distinction to the emphasized polished performance typical of professional voice talents.

Preliminary testing of different approaches in the presentation of informational audio options—options that are effective in pointing to thematically or conceptually different information objects—suggests a conversational approach is appropriate to maintaining a level of playful engagement and dialogue with ec(h)o. Since this approach is based on a style of presenting artifact information that has a ‘teasing,’ humorous quality, the vocal approach taken is appropriately different. Of the four voice talents delivering the audio objects, two are used to present the prefaces—one male and one female voice. Again, the objective is to have a natural, spontaneous voice, but with a greater emphasis on ‘character’—a more upbeat accent and inflection. This enhances the immediate playful engagement of the museum visitor and, as a consequence of this engagement, successfully provokes a greater interest in selecting a particular audio object.

4.3.3 Audio Production
Once the scripted objects are transferred to audio, the files are compressed in high-resolution mp3 format for the purpose of quick retrieval over an Internet connection. Given the slight loss of fidelity due to this compression technique it is essential that the source recordings be clear and of optimal amplitude to from the outset, in order to be clearly heard through the transmitted wireless audio format.

Another important production process to be considered is the storing and categorizing of the database of audio objects as a basis for a cross-institutional adaptive information retrieval and interaction model system. One option for a naming convention involves using a semantic signifier in combination with a numerical index: [botany]_00001.mp3, where this signifier could be derived from any of the subject tags applied to each individual record. The final ec(h)o audio object database design omits this signifier due to possible future inconsistencies with the collections of other museums that might wish to participate in the development and sharing of knowledge object repositories.
4.4 Stage Four - Inference Rules

In order to develop the inference rules we developed three models to conceptualize and test the rules: 1) visitor model; 2) narrative model; 3) soundscape model. In addition to the content and content mapping process outlined above, we relied on our initial observational and site studies of museums and museum visitors, discussions with museum administrators, exhibition designers and curators, and the research literature in museum studies (Lehn, Heath et al. 2001; Sparacino 2002).

4.4.1 visitor models

Our visitor model is comprised of three classifications of users:

- A busy visitor does not want to spend much time in each exhibit; instead he/she wants to stroll through the museum to get a general idea;
- An avaricious visitor wants to know as much as possible; He does not rush and moves from one exhibit to another in near sequent order;
- A selective visitor mostly chooses sound objects that represent certain concepts.

There are three levels of interest: -1 (indicates disinterest), 0 (indicates some interest), and 1 (indicates more interest), but they can be extended. Visitor's interest are computed as follows:

- When an avaricious visitor enters an exhibit, and is slow, his/her interests will be asserted to the primary concept of any narration that describes an artifact in that exhibit. This makes sense, because we do not need to be picky about interests and we can assume that he is interested in almost any concept.
- Interests of a selective visitor does not get easily overwritten. The rules engine should infer new interests only after he repeatedly choose narrations with certain concepts.
- For each exhibit, we need to calculate what is primary concept of most narrations that are about that exhibit. The interests of a busy visitor can only be overwritten with those, when he enters an exhibit.
- For any visitor, when he repeatedly refuses to listen to narrations with certain primary concepts, we can infer his/her disinterest to those concepts.

4.4.2 Narrative models

In addition, to our goal of linking repositories and ontologies across different museums, we also faced the task of linking content across different exhibit sections. In order to maintain coherency in an ec(h)o visitors experience we saw it as being necessary to provide meaningful links between audio objects. To facilitate this, it would be important to avoid situations where a clear disconnect existed between two audio objects. In defining the notion of a clear connection we identified the following categories of linkage types:

- Artifact to artifact: This occurs when the content of two audio objects makes reference to, or explicitly speaks to, the same artifact. For example: audio object A and audio object B fall into this category when they both reference the same moose antlers.
- Concept to concept: This link occurs when two audio objects are conceptually linked—for example, audio object A and audio object B might both talk about adaptation, and could therefore be linked with out being considered discontinuous. Note: it is our assumption that concept to-concept links are less tangible then artifact-to-artifact links. It is also worth noting that an audio object will often speak to more then one concept. When multiple concepts are present in an audio object, it is usually possible to discern one that is more prominent then the others present. Therefore, a classification hierarchy of sorts can exist when we consider the notion of an audio objects conceptual make up—that is, we might have an audio object with a primary concept, and a secondary concept. Here secondary concepts are defined as being less explicit then primary.
• **Localized links:** The notion of the localized link comes from the observation that visitors like to explore when they are taking in an exhibit. The idea here is that disconnects are not always a bad thing, and that visitors find inherent satisfaction in the experience of re-orienting themselves. To provide for this, we have made room for supporting discontinuous links between objects, as long as they are at least partially contextually localized—that is, in the same exhibit space.

Based on the above explanation, linkage classifiers were formalized and used to create rules that the system could then manage. In total two types of linkage classifiers were developed (primary, and secondary) and each classifier was given a point value. Point values reflected the concept of linkage tangibility. It was our assumption that, in general, conceptually linked objects are less tangible (unless, of course, the concept is made very explicit) than artifact linked objects.

The primary link classifiers were those described in the discussion above (artifact, concept, and localized) Secondary link classifiers deal with the presence of contextual information embedded in an audio object itself. Context information is defined as that which makes explicit reference to an artifact—i.e: “the shell marked number 5”...or, “the moose antlers in the center of the display”. Contextual information helps to facilitate the visitor’s reference, and is thus important when dealing with artifact changes, and objects that are linked based on the localization classifier. Two kinds of secondary classifiers exist—contextualized, and non-contextualized.

In evaluating the linkage potential between two objects, sameness and difference across the primary link classifiers is considered. Contextualized content with in the objects is also considered. To be linked, the sum of the primary and secondary scores must achieve a certain value. An artifact-to-artifact link is the most tangible, and therefore it is always classified as being linkable, regardless of it’s conceptual, and contextual information score. Note that contextualized objects that are not localized are prone to creating strong disconnects, therefore any objects that fall into this category is never allowed to be linked.

### 4.4.3 Soundscape model

The soundscape model is composed of zones of ambient sounds that are modulated when compared to a user’s interactions and interests. In addition, proximity to soundmarks effects the overall soundscape. The sounds are generally abstract in nature.

### 5. Evaluation

Given the complex nature of the system and user evaluation we tested our design and development of the audio objects as we went along. User tests were performed to evaluate the interaction model, the use and style of prefaces and audio objects, the inference rules and narrative models, and a series technical and integration tests that allowed for limited user testing of the overall system. The final prototype will be installed at the Nature Museum of Canada in March 2004, and we will then perform extensive user testing. The series of progressive testing allowed us to modify our current design and inform subsequent designs.

To date, users have found the interaction experience to be coherent and that design of prefaces and audio objects has been effective. Participants reported no significant issues around poor flow, or clunky content presentation. A consensus emerged in support of the style and flavor of the audio object prefaces, which were viewed as being entertaining and effective based on their ability to pique curiosity and motivate further interaction. For the most part, topical links between objects were better observed then conceptual links. Two characteristic behavior patterns emerged to indicate that our original concern over avoiding disconnects across linked objects may have been unwarranted.

First, participants tended to jump across topically, and in doing so often encountered disparate content in their turn taking. Second, participants admitted that their impetus for choices was more in keeping with a need to satisfy their curiosity (curiosity created by the prefaces, that is). This partly countered our assumption that participants would be exercising choice based on a need to hear more information about a specific topic or concept. Both of these insights indicated that users were more inclined to approach the experience from a position of play, rather then structured, focused exploration. A welcomed result!
6. Future Work and limitations

Current limitation of our process is the timeliness by which audio objects are designed, meta-tagged and then tested. This is mitigates open development of audio objects available for use within the network by other producers. In a related issue, the current system has a very limited implementation of the networked potential of the system.

In the areas of audio display and interaction, we will need further testing to evaluate if our minimal intervention in terms of contextual guidance is successful or not. We may find that visitors will need more explicit instructions either through audio or text. In addition, we have some concern about issues of selection and integration of the various modes of audio display and their combination as determined by an inference system. It should be ensured that the frequency range, amplitude and ambient elements from one sound layer are not interfering with the bandwidth and clarity of the other sound layer, for optimal auditory satisfaction.

Future work will lead us to researching further the complex roles of the design of audio objects, inference rules, audio display and the interaction model in creating an engaging and playfully exploratory interaction.

References


**Acknowledgements**

Work presented in this paper is supported by Canarie Inc. grant under E-Content program. Authors would especially thank to the Mark Graham and his colleagues in the Nature Museum in Ottawa for the enthusiastic support to this project. We would also like to thank our colleagues and participants in several workshops that contributed to the development of the project, namely Doreen Leo, Gilly Mah, Robb Lovell, Mark Brady, Jordan Williams and Leila Kalantari.

17. Socio-Ec(h)o: Ambient Intelligence and Gameplay

Ron Wakkary, Marek Hatala, Robb Lovell, Milena Droumeva, Alissa Antle, Dale Evernden, and Jim Bizzocchi

Introduction

This chapter describes the preliminary research of an ambient intelligent system known as socio-ec(h)o. Socio-ec(h)o explores the design and implementation of an ambient intelligent system for sensing and display, user modeling, and interaction models based on game structures. Ambient intelligence computing is the embedding of computer technologies and sensors in architectural environments that combined with artificial intelligence, respond to and reason about human actions and behaviours within the environment.

Ambient intelligent spaces lend themselves extremely well to physical and group play, and here describe our design of an interaction model and supporting system based on physical play. The overall research goal of this project is to understand to what degree physical play and game structures such as puzzles can support groups of participants as they learn to manipulate an ambient intelligent space. Future evaluation of this project will allow us to more fully answer this question. To date we have designed and implemented the prototype and interaction model. We have incorporated formative feedback through a participatory design process. This approach allowed for the concurrent development of the concept, interaction model and prototype environment.

In this chapter, we provide an overview of background concepts and related research. We then describe the game structure and prototype of our environment, including a technical overview of the system. We discuss the utilized game research concepts and our approach to group user models...
based on Bartle’s game types [1], and our ecologically inspired design approach to socio-cc(h)o. We conclude with a discussion of our work to date and future research.

**Background**

Key contextual issues in socio-cc(h)o include related research in the area of play and ambient intelligent spaces, and literature linking play and learning.

Björk and his colleagues have observed progress toward fully ubiquitous computing games, yet they identify the need to develop past end-user devices such as mobile phones, personal digital assistants and game consoles. Accordingly, we need to better understand how “computational services” augment games situated in real environments [2]. Recent projects have investigated the play space of responsive environments and tangible computing utilizing sensors, audio, and visual displays. For example, Andersen [3], and Ferris and Bannon [4] engage children in exploratory play and emergent learning through sensor-augmented objects and audio display. Andersen’s work reveals how theatrical settings provide an emotional framework that scaffolds the qualitative experience of the interaction. Ferris and Bannon’s work makes clear that a combination of simple feedback and control leads children to widely explore and discover a responsive environment.

In the Nautilus project, Strömberg and her colleagues employ bodily and spatial user interfaces as a way of allowing players to use their natural body movements and to interact with each other in a group game within a virtual game space [5]. Strömberg observes that in physical and team games such as soccer or dodge ball, players coordinate their physical movements and rely heavily on communication to be successful. In their findings, participants reported that controlling a game through one’s body movement and position was “new and exhilarating.” In addition, playing as a team in an interactive virtual space was found to be engaging, natural and fun.

In relation to the above research, socio-cc(h)o builds on the theatrical, simple and physical interaction models in order to develop a game structure approach that lies between exploratory play and a structured game for adults within an ambient intelligent environment. In addition, we extend the notion of a game structure to an interaction model for the environment rather than a virtual game space. We also build on the idea that action, play and learning are linked in such physically based environments.

In respect to the links between action, play and learning there is a substantial amount of literature. Dewey argued for the construction of knowledge based on learning dependent on action [6]. Piaget, through his child development theory, believed in the development of cognitive structures through action and spontaneous play [7]. According to Piaget, constructivist learning
is rooted in experimentation, discovery and play, among other factors. Papert extends Piaget's notions by investigating the knowledge-construction process that emerges from learners actually creating and designing physical objects [8]. Malone and Lepper consider games as intrinsic motivators for learning [9]. Subjective motivations like challenge, curiosity, control and fantasy may occur in any learning situation; other motivations like competition, cooperation and recognition are considered to be inter-subjective, relying on the presence of other players/learners. Design related theories have placed activity at the center of design action as in Nardi and O'Day's activity theory based on information ecologies [10]. Schön argues that design is a series of actions involving experimentation and learning in the framing and re-framing of a design situation [11].

**Game Structure and Prototype**

**Description and Scenario**

The aim of our game is for a team of four players to progress through seven game levels. Each level is completed when the players achieve a certain combination of body movements and positions. At the beginning of each level, players are presented with a word puzzle as a clue to discovering the desired body states (see Figure 17.1). The levels are represented by changes in the environment in light and audio. The levels are progressively more challenging in terms of body states and more complex in terms of the audio and visual ambient display. The physical environment currently consists of a circumscribed circular space (the area in which we can detect motion), surround sound audio, theatrical lighting, and two video projection surfaces.

![Figure 17.1: Depicted above are frames from a thirty-three second video introducing a level to players.](image)

The images at the beginning of the video provide a sense of the environmental change the players are trying to achieve for that level. This is followed by a clue in the form of a word puzzle aimed at helping players discover the desired body state. In this example, the puzzle is "two sloe plum berries among trees."
Table 17.1: This table describes the socio-ec(h)o game schema

<table>
<thead>
<tr>
<th>Theme</th>
<th>Levels</th>
<th>Body state</th>
<th>Goal</th>
<th>New game skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery of light</td>
<td>1</td>
<td>&quot;high-low&quot;</td>
<td>create day</td>
<td>body position</td>
</tr>
<tr>
<td>Day for night</td>
<td>2, 3</td>
<td>&quot;moving low&quot;</td>
<td>create night</td>
<td>movement/duration</td>
</tr>
<tr>
<td>Rhyme</td>
<td>4, 5</td>
<td>&quot;loose moving&quot;</td>
<td>create winter</td>
<td>sequencing</td>
</tr>
<tr>
<td>Rota</td>
<td>6, 7</td>
<td>&quot;two low moving—two high&quot;</td>
<td>create summer</td>
<td>composition &amp; location</td>
</tr>
</tbody>
</table>

Here is a short scenario of participants in the socio-ec(h)o environment:

Madison, Corey, Elias and Trevor have just completed the first level of socio-ec(h)o. They discovered that each of them had to be low to the ground, still, practically on all fours. Once they had done that, the space became bathed in warm yellow light and filled with a wellspring sound of resonating cymbals. Minutes earlier, the space was very dim—almost pitch black until their eyes adjusted. A quiet soundscape of “electronic crickets” enveloped them. They discussed and tried out many possibilities for solving the word puzzle, “Opposites: Lo and behold.” They circled the space in opposite directions. They stood in pairs on opposing sides of the space. At Corey’s urging, the four grouped together on the edge of the space and systematically sent a player at a time to the opposite side in order to gauge any change in the environment. Nothing changed. Madison, without communicating to anyone realized the obvious clue of “Lo” or “low.” While Corey was in mid-sentence thinking-out-loud about the puzzle with Trevor, and trying to direct the group into new body positions, Madison lowered herself to a crouching position. The space immediately glowed red and became brighter. The audio changed into a rising chorus of cymbals—not loud but progressively more pronounced. Corey and Trevor stopped talking and looked around at the changing space. Madison, after a pause began to say “Get down! Get down!” Elias stood immediately and the space became even brighter. Corey and Trevor dropped down in unison and the space soon became bathed in a warm yellow light like daylight. The audio reverberated in the space. A loud cheer of recognition came from the group, “Aaaaahhh! We got it!” Corey asked everyone to get up. As soon as they were all standing, the space became pitched black again. They dropped down again and the space was full of light. They had learned how to “create daylight” in the space. They had completed level one.

Soon after, a new word puzzle was presented to them in a short video projected on two screens hanging from the ceiling: “The opposite of another word for hello but never settles.” The lights have 247 become very dim and the audio has a slightly more menacing quality to it. Level two will...


Figure 17.2: Scenes from the final participatory design workshop in which the relationship between body states and word puzzles were explored. The system utilized is an early prototype with the substantive functions implemented. As players lower themselves, the environment becomes progressively darker and the ambient audio more pronounced. Note: the display response is not the same as depicted in the scenario. In this exercise, the goal was to create darkness rather than light.

We formalized our game structure into a schema of levels, body states and goals (see Table 17.1). As earlier described, the game has seven levels. The body states are the body movements and positions that players must discover in order to complete a level. Goals are the change in environment players are aiming to achieve. The goals are implicit and not explicitly stated for the players. Each level has a beginning quality of light and audio. As the players progress toward achieving the right body state, the environment incrementally shifts toward the goal state of the environment. For example, as depicted in the scenario, when Madison lowered herself, the environment gradually shifted toward the goal of creating day. As each of the other three players followed Madison, the environment responded to movements of each player (see Figure 17.2).

In addition, the schema includes new game skills and themes. We assigned each level a generic skill in relation to each body state and level. Despite the specific body state, the generic skill acquired at each level is required in order to discover the more complex body states at higher levels. Themes allowed us to design an implied progressive narrative based on natural evolution. Again, the specific themes and even the narrative are not known to the participants, rather they provide an underlying structure for
body states, goal states and game skill acquisition. We intend for the progressive narrative to provide a sense of coherency across the levels, and to loosely map increased challenge to the reward of a more complex display.

**Technical Prototype**

The technical system for socio-ec(h)o includes three key components, a sensing system, reasoning engine, and display engine.

**Sensing System**

The sensing engine comprised a twelve-camera Vicon MX motion capture system (www.vicon.com) and a custom program written in Max/MSP. Each participant is differentiated by unique configuration of reflective markers worn on their backs. The system senses for discrete parameters such as velocity, position (x,y,z), orientation, proximity, and movement. It measures across each unique player for participation and duration. Data is transmitted to the reasoning engine for high-level interpretation.

**Reasoning Engine**

The reasoning engine provides the intelligence for the system. It interprets the sensing data samples, identifies the level of body states completion, and manages the narrative flow of the socio-ec(h)o experience. The engine receives sensing data from the sensing system and interprets it in terms of high-level group behavior. For example, the sensing system sends data on predefined parameters such as velocity and body positions and the reasoning engine synthesizes the parameters to determine if a given body state is achieved. The characteristics and their combination, and in some cases their sequence determine the 'intensity' of the state. Another factor influencing the intensity of a state is the group user model that is dependent on the combination of user types as identified by Bartle classifications (see "Description of Game Concepts and User Model"). The role of the engine is to manage the flow in the game by sequencing of the states and managing the timing of the state transitions. The reasoning engine is rule based and allows seamless modification and extension for other applications. The reasoning engine feeds its output, state intensity, and state transition to the display engine.

**Display Engine**

The display engine has two components, an audio and a lighting component. The audio display engine for socio-ec(h)o provides a sound ecology for each individual level of the system; it is custom software programmed in Max/MSP. We developed and structured the audio content on the principles of acoustic ecology and feedback-as-communication [12]. In addition, the audio display provides a gradient response to the participants, telling them
how close they are to achieving their goal. The audio display system can alternate between stereo and multi-channel formats and localized and ubiquitous sound. The audio content follows the theme of evolution by utilizing sampled sound and several different sound processing techniques creating a shifting ambient soundscape that moves from simple, abstract sound to rich, environmental sound.

Lighting is manipulated with a DMX 512 controller via a Max/MSP patch. A small light grid and theatrical style lighting instruments and color scrollers are used. A lighting console was created to control multiple lights and color in concert through a cue list mechanism. Cues were written to simulate the various themes at each level.

Both the audio and the lighting systems take their cues from the reasoning engine, and respond to game aspects and configurations specified in the reasoning engine. Thus, the response of the display systems can potentially be used to provide feedback based on a variety of parameters such as how well participants are working together as a group.

**Integration**

The three components described above run on their own servers. The integration is achieved by lightweight communication protocol that is transferred over the User Datagram Protocol (UDP) communication channel. We consider uni-directional UDP communication appropriate for the real-time applications. Although the sensing system is capable of capturing data at the rate of thirty frames per second we are using slightly longer sampling rate of 200 ms for data transferred between the servers. Considering the nature of the output (sound and video) this rate is sufficient for the required fine-grained response.

**Description of Game Concepts and User Model**

We investigated the play and game aspects of socio-ec(h)o through short ethnographic sessions, workshops among the researchers, and games research theory. We explored a range of game concepts including traditional game theory notions such as Nash's Equilibrium to contemporary games research [13,14,15]. Our aim was to encode a form of play for groups that lays between a structured game and open-ended play. Each level acted as a kind of group puzzle—that is, a game with a single solution or winning state [16]. The alignment of increasing difficulty of level solutions with the increasing skill of the participants as they proceed through the experience is consistent with Csikszentmihalyi's model for developing a state of flow [17].

In the end, we relied on our participatory design process to evolve our game structure. This was especially helpful since we were exploring social
aspects of gaming such as communication, collaboration and shared cognition. In addition, our approach was highly physical which required an empirical approach to understanding our concepts. Theoretically, we utilized Bartle's concepts of collaborative play in Multi-User Dungeons (MUDs) and MUD Object Oriented systems (MOOs) to help us formulate a group user model to support the reasoning within our system [1]. Bartle identified four types of MUD player styles: achievers, explorers, socializers, and killers. Achievers seek in-game success, explorers satisfy their environmental curiosity, socializers value human interaction, and killers exercise their will at the expense of other players.

The group user model in socio-ec(h)o is constructed based on Bartle's typology. The group user model is applied in the interpretation of the individual actions in the environment and the level with which individual actions contribute towards an overall group activity. The display response in socio-ec(h)o is adjusted with respect to the group composition. We are currently investigating our hypothesis that by considering Bartle’s types participants have a better experience and more quickly become skilled interactors.

**An Ecological and Participatory Approach**

The key components of our ambient intelligent model are addressed as a systemic whole; they include interaction, reasoning, response and technology. For example, we investigated the balances between sensor technologies in relation to gesture, inference rules and display responses. Our design approach—inspired by an ecological frame—is centered on human activity and is participatory design driven, informed by observation and theory.

The concept of an ambient intelligence ecology emerged from findings in a previous research project known as ec(h)o. We discovered that ec(h)o had successfully balanced incongruent elements to form a dynamic and coherent system. Components such as interaction, reasoning, audio display and technology shaped the ambient intelligent environment around the purpose of a museum visit [18]. In ec(h)o we explicitly utilized an information ecology approach as an ethnographic analysis of the museum as well as a scaffold for our design decisions [19]. Nardi and O’Day describe ecology as a system of people, practices, values, and technologies in a local environment. They argue that the ecology metaphor shifts the focus to human activity rather than on technology [10].

The current research, socio-ec(h)o represents a preliminary exploration of the concept of an ambient intelligent ecology. The experimental nature of the project, its laboratory setting, and the fact that participatory cannot be considered to be part of a relevant definable ecology limit the degree to which this research fully explores the concept. Nevertheless, we feel this is a
starting point in investigating an ambient intelligent ecology design approach. At this stage, we found the use of participatory design workshops to be a key component of an ambient intelligent ecology approach. The workshops simultaneously addressed issues of interaction, reasoning, response and technology. We ran five workshops that progressively explored open-ended concepts to more defined concepts. These workshops included investigations of the continuum between play and game, the physicality of our interaction model, the social aspect of play within our type of space, puzzles and narrative and finally the relationship between our body states and word puzzles. It was evident to us that we required an empirical and qualitative understanding of our concepts, interaction, technology use and prototyped environment while we were designing them.

Discussion

In designing an interaction model and supporting system for the play and learning of a complex system, many issues relate to the communication and action between participants and between the group of participants and the system.

Embodied Cognition and Communication

In our design, a successful participant experience relies on a tightly coupled system emerging out of real-time, goal-directed interactions between participants, and between participants and the responsive environment. The nature of these interactions influences the challenge, enjoyment and success within the environment. Communication is key to solving the puzzles and coordinating actions. While the verbal discussion among participants is frequent and active, the physical nature of the interaction model and game structure emphasizes explicit embodied action, cognition, and communication. Players actively work out the puzzle physically, as well as communicate actions and ideas physically. In many respects, the interaction model is founded on an embodied cognition view of interaction [20,21]. Success in the game requires a quality of interaction in which mind, body, and environment mutually interact and influence one another positively.

From the perspectives of the design of the interaction model and system, we realized it was important to decide where not to specify interaction and system functionality. In many respects, we learned to off-load formalized interaction among participants to the situated dynamics of a group of people working toward a shared goal, that is, we will communicate together in whatever form possible given the resources in the environment without the need of formalizing communication modes. In addition, the system does not
need to encode or sense actions or behaviours that are not relevant to the
desired body state at a given level, that is, no response from the system is a
perceived response. We feel we supported an embodied and inter-subjective
approach through limited means such as design constraints. For example,
limiting sensing actions to whole body positions and movements rather than
gestures opened gestures up to unique, specific and complex communication
between participants. Ignoring or not encoding large parts of the embodied
action supported a wide range of exploration of body movements.

Selective Real-Time Response

We were, however, selective as to when and how the system did respond to
participants' actions. The system responds when it appears that the group is
on a trajectory toward the desired body state. The response is conceptually
similar to someone telling another if they are close to a goal by stating if they
are cool, warm, or hot. Through our participatory design workshops we
learned that four factors had to be met in our response in order to achieve
this type of support through a changing environment. The response had to be
in perceived real-time, the feedback is on a gradient related to the prox­
imity to the goal, a reward is given for achieving a goal, and a response is
mapped to the make-up of the group. The quality and nature of physical
action requires an accepted real-time response. Given that we require a min­
imum duration before a body state is recognized, we had to find the thresh­
old for what was understood as a required time to hold a position versus a
perceived lag or failure of the system. A gradient response is critical in aiding
players in understanding they are on the right track. The response is coordi­
nated between the audio and the lighting. The nature of the gradient
response is well illustrated in the scenario and Figure 17.2. An audio reward
is given once a state is completed. This is needed since the continuum of
coordinated actions and durations is not explicit to the participants. Based
on the different groups of players determined as a composite of Bartle types,
we modified response and time. While we have yet to fully evaluate this
approach, we anticipate a group of achievers will expect a different response
or precision of action than a group of primarily explorers. With this in mind,
we also provided a range of word puzzles of differing difficulty and challenge.

Empirical Nature of Designing Ambient Intelligent Systems

An ambient intelligent ecology, as stated above, is an ongoing investigation of
how we might define a design process specific to the challenges of ambient
intelligence. Given the centrality of situated human activity and the need to
develop an interaction model and system as a whole, reducing the
process or system to discrete elements is not a reasonable approach. While a theoretical approach to the design and system helps frame the problem and support design decisions, ultimately it is an empirical process such as ethnography or, in this case, participatory design that yields useful qualitative and quantitative understanding of how to design the interaction model and system. The application that arises, such as socio-ec(h)o, becomes a specific ecology of constraints, affordances and system functions that is situated and relies on a unique dynamic of embodied action between people and environment.

**Conclusion and Future Work**

In this chapter we have reviewed related research and have shown how our system builds on theatrical, simple, and physical interaction models in order to develop a game-based approach to ambient intelligence that relies on exploratory play with a conceptually structured interaction model. We discussed the links between play, learning, and action that we extended into an embodied cognition approach. We provided a description of our game structure and prototype from conceptual and technical perspectives, and we discussed how we use Bartle’s game types as the basis for our user types and group user model. We introduced our approach to designing ambient intelligent systems that is ecologically inspired. In our discussion, we explained the role of embodied cognition within our design, and elaborated on how we decided where and where not to formalize and encode embodied actions, cognition, and communication. We detailed how the success of the experience relied on selective responses that were real-time, gradient, provided rewards and were unique to different group user models. Lastly, we stressed the empirical nature of the design work and the role of participatory design in developing our system.

Future work includes a series of evaluations of the system to better understand the influence of the game structured interaction model, the supporting user model, and the display. In particular, we aim to understand how our approach enables a better experience and more skilled interactors within an ambient intelligent environment.

**Acknowledgments**

We would like to gratefully acknowledge the support of Canadian Heritage for funding the Am-able Network for Mobile, Responsive Environments and this project. We would also like to acknowledge funding support from the Social Sciences and Humanities Research Council and the Natural Sciences and Engineering Research Council.
References


Do we play in museums? Anthropologist Genevieve Bell identifies the notion of play together with learning in museums. She describes museums as different cultural ecologies in which the museum visit has the qualities of liminal (a space and time set apart from everyday life) and engagement (where visitors interact to both learn and play) (Bell). Guided by the notion of play in a museum experience we have considered playfulness equally with functionality and learning in the design of an adaptive museum guide. Our approach includes a tangible user interface (TUI), spatial audio, and an integrated user modeling technique combined with semantic technologies that support exploration and discovery. We understood our interface as playful action along the lines of aesthetic interaction. By this we do not mean the type of structured play that is found in a software game on a mobile device, rather we refer to the less structured and open play that is always possible and often can be subtle and implicit like toying with a ball.

Here, we will describe our case study of an augmented reality museum guide, known as Ec(h)o, which we installed and tested at the Canadian Museum of Nature in Ottawa. We discuss our use of a set of ecological approaches from the domains of regquel, ethnography, and information technology. Aesthetically, we
based much of our work on the ideas of sound ecologies. We also adopted the idea of museums as ecology informed by Bell's cultural ecologies and Nardi and O'Day's information ecology. This approach led to us being inspired by simple physical displays and puzzles we observed in our ethnographic sessions. These observations encouraged the playful tangible object and use of puzzles in our audio content. We were also motivated by the storytelling of the museum staff and researchers that was often humorous as well as informative. In this project, we found that learning effectiveness and functionality can be balanced productively with playful interaction through an adaptive audio and tangible user interface approach and that diverse ecology models help define the imaginative space and contextual aspects of play.

**Previous Work**

Ec(h)o shares many characteristics with the adaptive systems of HyperAudio, HIPS, and Hippie (Petrelli et al.; Benelli et al.; Oppermann and Specht). Similar to ec(h)o, the systems respond to user location and explicit user actions. All systems adapt content to the user model, location, and interaction history. Among the key differences with ec(h)o is that these systems depend on a personal digital assistant (PDA) and graphical user interface (GUI); ec(h)o uses audio display as the only delivery channel and a tangible object as an input device. In addition, ec(h)o uses inference at the level of semantic descriptions of independent audio objects and exhibits, and ec(h)o treats user interests as dynamic.

Prior to the evolution of adaptive and user modeling approaches in museum guide systems, there has been a strong trajectory of use of the PDA graphical user interface. Typically, hypertext is combined with images, video, and audio (Semper and Spasojevic; Proctor and Tellis). Yet, a PDA is essentially a productivity tool for business, not a device that lends itself easily to playful interaction. Given that PDAs use miniaturized desktop-based GUIs, we wondered if we should metaphorically carry around our desks in order to experience such things as museums—in what might be described as a *world-behind-a-desk* approach to mobile computing. Museum systems have mostly maintained the PDA graphical user interface approach despite the shifts in other domains to other approaches that better address the experience design issues most prominent in social, cultural, and leisure activities. The play constraints of these devices are too great for the level of interaction that goes beyond playing a software game on a mobile device. For example, in the area of games and ubiquitous computing, Bjork and his colleagues have identified the need to develop post-end-user devices such as mobile phones, per-
sonal digital assistants, and game consoles (Björk et al.). They argue that we need to better understand how “computational services” augment games situated in real environments. The same can be said for museum visits.

Audio is seen as an immersive display that can enrich the physical world and human activity while being more integrated with the surrounding environment (Brewster et al., Pirhonen et al.). In addition, audio tends to create interpretive space or room for imagination as many have claimed radio affords over television. In the HIPS project, different voices and delivery styles were used to create an “empathetic effect” between the user and the artifacts they engaged (Marti). We have adopted a similar approach to our use of audio content.

**Description of Ec(h)o**

Our approach includes a tangible user interface (TUI) for its inherent playfulness, spatial audio display for its imaginative qualities, and an integrated user modeling technique combined with semantic technologies that supported exploration. Our aim is to improve the visitor engagement by considering playfulness equally with functionality and learning.

The implementation went as follows: visitors are given a pair of headphones connected to a small, light pouch to be slung over their shoulder. The pouch contains a wireless receiver for audio, and a digital tag for position tracking. When in front of an exhibition display, ec(h)o offers the visitor three short audio pieces as prefaces that are acoustically to the visitor’s left, center, and right respectively. This spatial structure allows the three prefaces to be distinguishable. The spatialization is mapped to the TUI, a wooden cube (see figure 22.1), for selection. The visitor responds by rotating the wooden cube in his hand and thus selecting a preface. The system delivers the audio object related to the preface. After the delivery of the object, the system again offers three prefaces. The visitor’s response is expressed through the gesture selection with the wooden cube. Additionally, the system may be met by no response, because the visitor does not wish to engage the system. The system will then enter into a silent mode. The visitor may also have moved away and the system will then initiate a soundscape that continuously plays while the visitor is in the spaces between exhibition displays.

The audio objects are semantically tagged to a range of topics for possible integration with networks of information across the Internet. At the beginning of each interaction cycle, three audio objects are selected based on ranking, using...
several criteria such as current levels of user interest, location, and interaction history reasoned through our user model (Hatala and Wakkary). The topics of objects are not explicit to the visitor; rather, the content logic is kept in the background.

In regard to the design process, many of the design choices were made through a series of participatory design workshops and scenarios (Wakkary). For example, the tangible user interface and its implementation as an asymmetrically shaped wooden cube resulted from these workshops, as did the use of the conversation metaphor, navigation, and audio interface. In addition, we prototyped the exhibition environment and system in our labs in order to design the interactive zones, audio display, and interaction with the exhibit displays.

**Tangible Object**

The tangible user interface is an asymmetrically shaped wooden cube with three adjacent colored sides (see figure 22.1). The cube was carefully designed to ensure proper orientation and ease of use. The "bottom" of the cube has a convex curve to fit comfortably in the palm of the visitor's hand and a wrist leash is attached to an adjacent side to the curved bottom suggesting the default position of being...
upright in the palm and at a specified orientation to the visitor’s body. The leash allows visitors to dangle the cube when not in use and frees the use of their hand. The opposite side of the bottom of the cube is colored and shows an icon denoting a pair of headphones with both channels active. The sides to the left and right are each uniquely colored and display icons showing active left and right channels of the headphones, respectively. The cube is made of balsa wood in order to mitigate tiredness from carrying the object and is therefore very light (approximately 100 grams or 3.5 ounces).

The input of the selection is done through video sensing. The ergonomic design of the cube and biomechanics of arm and wrist movement form a physical constraint that ensures the selected cube face is almost always held up parallel to the camera lens above and so highly readable. We experienced no difficulties with this approach.

**Audio Display**

The audio display has two components, a soundscape and paired prefaces and audio objects. In the latter component, we used a simple spatial audio structure to cognitively differentiate the objects. Switching between the stereo channels created localization: we used the left channel audio for the left, right channel audio for the right, and both channels for the center. In addition, we provided simple chimes to confirm that a selection had been made.

The prefaces were written to create a sense of surprise, discovery, and above all play, especially in contrast to the informational audio objects. In order to create this sense we used diverse forms of puns, riddles, and word play, such as ambiguous word play used in the preface “Sea urchins for sand dollars”; turn of phrase like the preface “An inch or two give or take a foot”; and riddles like “What is always naked and thinks on its feet?” In addition, the audio recordings of the prefaces and audio objects used a diverse set of voices that were informal in tonality and style. This added to the conversational feel and created an imaginary scene of a virtual cocktail party of natural historians and scientists that followed you through the museum.

Visitors’ movements through the exhibition space in between artifact displays generated the soundscape. Visitor movement was tracked using a combined Radio Frequency Identification (RFID) and optical position tracking system developed by Precision Systems (www.precision-sys.com). We divided the exhibition space into interactive zones and mapped concepts of interest to each zone and display.
The concepts are translated into environmental sounds such as the sound of an animal habitat, and sounds of animals such as the flapping of cranes' wings. The visitor navigates the exhibit exploring it on a thematic level through the ambient sounds that are dynamically created. If a set of concepts strongly matches the visitor's interest the related audio is acoustically more prominent (see figure 22.2).

Ecological Approaches

We endeavored to consider how our design both intervenes and integrates within a complex museum experience. The ecological models of cultural ecologies and information ecologies provided us with frameworks for contextual analysis (Bell, Nardi, and O'Day). This approach allowed us to look further into the design process, past the interface for guidance, into how our design decisions were integral to the ecology or its inhabitants, thus supporting us in developing more appropriate design responses. Bell's cultural ecologies formally linked different actions and attributes of the museum visitor into a coherent description. As a descriptive tool it validated our assumptions and provided a clearer link between what we observed and the design implications. It was therefore generative, much like Nardi and O'Day's information ecologies framework. Both guided us in specific design decisions, namely the high degree of physical interaction that suggested a tangible user interface; the wide use of puzzles, riddles, and games as modes of learning which led to our use of a riddle-like approach to our audio content; and the localized and informal storytelling on behalf of the museum staff and researchers that inspired us to structure our audio experience like a virtual cocktail party. (For further discussion of the role of ecologies in museums we refer readers to Wakkary and Evernden.)

The auditory interface is another form of ecology. It provides the basic mechanisms of navigation and orientation within the information space. This entails investigations of mechanisms for navigating space-time modulations within the
narrative information space. These mechanisms form the key components of a modeled acoustic ecology that takes into account the variety, complexity, and balance of the informational soundscape. The research takes into account both psychoacoustic and cognitive characteristics of the ecology as well as compositional problems in the construction of a meaningful and engaging interactive audible display. Psychoacoustic characteristics of the ecological balance include spectral balancing of audible layers. Cognitive aspects of listening contribute to the design and effective use of streaming mechanisms allowing segmentation, selection, and switching among audible semantic objects within the soundscape. Compositional problems were addressed in the form of the orchestration of an informational soundscape of immersion and flow that allowed for the interactive involvement of the visitor. Techniques were drawn from sound design for cinema in developing relationships among soundscape, speech, and musical elements of the audible scene. The interface display takes into account transitions, beginnings, endings, and, perhaps most importantly, interruptions in the narrative audio-informational flow situating the awareness of the visitor.

Summary of User Evaluation

User experience was evaluated through observation, a questionnaire, and a semi-structured interview. The evaluation group included two men and four women, from twenty-five to fifty-three years old. We also performed expert reviews by a senior researcher and senior interaction designer from the museum. Participants found the system enjoyable and stimulating, perhaps in part due in its novelty. The results were quite clear that play was a critical experiential factor in using the system. It was often remarked how the experience was similar to a game:

"The whole system to me felt a lot like a game. I mean I got lost in it, I found myself spending a lot more time in a particular area than I normally would. And just the challenge of waiting to hear what was next, what the little choice of three was going to be. Yeah... So I found it overall engaging, it was fun, and it was very game-like." (Participant 5)

The playfulness did in most instances suggest a quality of engagement that led to learning, even though diverse types of museum visitors were involved, from the visitor who browses through quickly but is still looking to be engaged, to the repeat visitor who experiences the audio information differently each time.

The evaluation did point out challenges and areas for further research. Some things we expected, such as the headphones were uncomfortable, yet to such a
degree that we are currently rethinking the use of personalized spatial audio and headphones. Other results point to a threshold in the balance between levels of abstraction and local information, since visitors had difficulties at times connecting what they were listening to and what was in front of them (in part this was an inherent challenge in the exhibition since the display cases had from dozens to over a hundred artifacts). In many respects this points to the finding that the semantic technologies approach did not always provide a clear enough contextual link between the artifacts and the audio information. In addition, we see both a threshold point between play and focused attention on the exhibit. For example, one user's enthusiasm for the game-like quality led her to at times pay more attention to the interaction with the system than the exhibition. This raises the issue of balance in play and the possibility to shift attention away from the environment rather than play as a means of further exploring the environment.

Conclusion

Echo is an augmented audio reality system for museum visitors that uses a tangible interface. We developed and tested the prototype for the Canadian Museum of Nature in Ottawa. The findings of this project are positive, while also calling for more research in several areas. We conclude that learning effectiveness and functionality can be balanced productively with playful interaction through an adaptive audio and tangible user interface approach and that diverse ecology models help define the imaginative space and contextual aspects of play.

References


APPENDIX 6 - ABSTRACTS OF OTHER PUBLICATIONS

Journals


This paper discusses how families appropriate artifacts and surroundings that lead to the design of everyday household systems, such as combining a chalkboard, a door frame, a hanging basket with paper and sticky notes to manage lists and messages. Such systems continually evolve through the catalytic pressures of individual actions and design-in-use. The paper reports on a study of four families in which we were researching the concept of everyday design in the home. Presented are in-depth descriptions and discussions of our observations and patterns. The design implications of our study are also discussed. The research contributions are an explanation of everyday design as a novel way to understand interactions and routines in the home, descriptions of the key actions and process in everyday design, and the need to reconstruct the user in the sense of an everyday designer.


This paper presents an overview of the main features of sound design for games, and argues for a new conceptualization of it, beginning with a closer look at the role of sound as feedback for gameplay. The paper then proposes and details a new approach for sound feedback in games, which provides ambient, intensity-based sonic display that not only responds to, but also guides the player towards the solution of the game. A pilot study and leading outcomes from it are presented, in the hopes of laying a foundation for future investigations into this type of sonic feedback.

Conference Proceedings


In this paper we explore sustainability in interaction design by reframing concepts of user identity and use in a domestic setting. Building on our own work on everyday design and Blevis's Sustainable Interaction Design principles, we present examples from an ethnographic study of families in their homes which illustrate design-in-use: the creative and sustainable ways people appropriate and adapt designed artifacts. We claim that adopting a conception of the user as a
creative everyday designer generates a new set of design principles that promote sustainable interaction design.


We discuss three design strategies for improving the quality of social interaction and learning with interactive museum guides: 1) embodied interaction; 2) game-learning; 3) a hybrid system. We used these strategies in our prototype Kurio, which is aimed at supporting families visiting museums. The results of our evaluation show positive implications of implementing the design strategies: closing the social gap, naturalizing technology, and supporting exploration and discovery in learning.


Juxtaposing differences in methodologies and epistemologies is a method for provoking thought and identifying connections, which can lead to the development of new knowledge [1]. The goal of the panel is to catalyze discussions that build a foundation of mutual understanding and respect, and from this foundation, to build new ideas and human relationships that can lead to fruitful collaborations in the cycles to come.


The topic of interactive narrative has been under research for many years. While there has been much research exploring the development of new algorithms that enable and enhance interactive narratives, there has been little research focusing on the question of how players understand and internalize their interactive narrative experiences. This paper addresses this problem through conducting a phenomenological study on participants playing Fa9ade; we specifically chose a phenomenological methodology due to its emphasis on the participants' lived experience from the participants' viewpoint. We chose Fa9ade, because it is the only accessible example of an experience that revolves around social relationships, conflict, and drama as its core mechanics. In this paper, we discuss sixteen themes that resulted from the analysis of the data gathered through the study. In addition, we reflect on these themes discussing their relationship to participants' backgrounds, and project implications on the design of future interactive narratives.


This paper explores an intensity-based approach to sound feedback in systems for embodied learning. The theoretical framework, design guidelines, and the
implementation of and results from an informant workshop are here described. The specific context of embodied activity is considered in light of the challenges of designing meaningful sound feedback, and a design-based research approach is shown to be a generative way of uncovering significant sound design patterns. This exploratory workshop offers preliminary directions and design guidelines for using intensity-based ambient sound display in interactive learning environments. The value of this research is in its contribution towards the development of a cohesive and ecologically valid model for using audio feedback in systems, which can guide embodied interaction. The approach presented here suggests ways that multi-modal auditory feedback can powerfully support interactive and collaborative learning and problem solving.


We discuss our study that looks at family members as everyday designers. We explain the design actions of family members to be creative, as evidenced by the resourceful appropriation of artifacts and surroundings, the ongoing adaptation of systems and routines through design-in-use that allows emergent properties to arise and addresses individual needs, and how implicit understanding and explicit tests occur for judging quality. We present a preliminary analysis of design implications in the area of interaction design in the home. Our findings are based on a five-month ethnographic study of three families.


This paper describes the theoretical framework, design, implementation and results from an exploratory informant workshop that examines an alternative approach to sound feedback in the design of responsive environments for children. This workshop offers preliminary directions and models for using intensity-based ambient sound display in the design of interactive learning environments for children that offer assistance in task-oriented activities. We see the value of this research in developing a more cohesive and ecological model for use of audio feedback in the design of embedded interactions for children. The approach presented here takes the design of multi-modal feedback beyond being experiential, to one that supports learning and problem solving.


This paper discusses two studies, which investigate how designer and non-designer participants discussed and engaged in design activity. We found that close to half of non-designers can be mapped to “designer” profiles in analyzing design, and a distinct non-designer profile can also be described. We found that non-designers engage in a reflective practice approach to design that can be a
substantial base of design skills. This research contributes by providing a
description of an everyday designer that serves as a starting point for recasting the
role of those we design for as everyday designers.

Budd, J., Wakkary, R., A New Educational Model for Interactive Product Design:
Validating Utility, Performance and Experience (2005), IDSA 2005 National Education

New technologies are fundamentally changing the way we learn, work, and play. Technical
knowledge and understanding alone are inadequate to deal effectively
with many of the implications of new technology. This raises questions
concerning both what technology can do and what technology should do. In either
case, the products, systems, and services we create with new technology are of
little value if we can’t readily understand what they are, what they do and how to
use them.

Hosseini, M., Wakkary, R., Influences of Concepts and Structure of Documentary
Published at: http://www.archimuse.com/mw2004/papers/hosseini/hosseini.html [last
accessed: March 27, 2009]

This paper investigates the emergence of documentary practices on the Internet. It
is potentially beneficial to researchers and practitioners in the emerging area of
documentary web work, and for museums or cultural institutions who utilize web
sites for exhibition programming and thematic presentations. The goal of the
paper is to develop a methodological framework for analyzing web documentaries
based on current film theory, since to date, documentary practice has been rooted
in filmmaking. A secondary goal is to describe the current state of web
documentaries and potential benefits and hurdles in its future development. The
paper provides a review of potentially relevant film theories from which a
framework for analyzing documentary practices on the Internet can be built upon.
The framework proposed is based on Bill Nichol’s theory of documentary film.
The study includes the analysis of five documentary web sites that include a range
from independently to institutionally produced sites. We aim for the paper to
contribute to further investigations of the relationship between documentary
cinema and documentary practices on the Internet.

Budd, J., Taylor, R., Wakkary, R., & Everenden, D., Industrial Design to Experience
Design: Searching for New Common Ground (2003), ICSID 2nd Educational Conference
2003 Hanover, Hanover, Germany. (eds. Birgit Weller and Gunnar Spellmeyer) Cologne,

Given the rapid growth in internet commerce the promise of experience design
has attracted a significant following, but do we really have a sufficiently clear
understanding of the field to capitalize on the concept? In the past few years it has
become increasingly apparent that we are undergoing a major shift in consumer
preference. Mass production has given way to mass customization. In response
design has a more critical role to play in a more complex scenario. This shift from
Industrial Design to “Experience Design” has had a profound effect on the way we work. The scope of experience design has precipitated the need for business, design, cognitive science and computing science to work in collaboration. The history of experience design is one of converging fields and can generally be located among five knowledge domains including: industrial design, graphic design, architecture, computing science and management. It is therefore not surprising that a variance in terminology units has generated confusion surrounding the concept of experience design. This paper proposes steps to create a more precise interdisciplinary and integrated terminology that advances our understanding and our ability to interpret the results of research that could contribute to more effective use of Experience Design.


This paper explores the idea of Pattern Language as a method for interaction design in a social network environment. It is the belief of the author that Pattern Language can enable interaction design to address group interaction within an “unbounded” interaction event. The aim is to define a theoretical starting point for adapting Pattern Language that addresses criteria informed by the concept of social network analysis. The criteria are comprised of three properties: community definition, reciprocity and scale, and rich methodology.


The aim of this paper is to demonstrate how an evaluation framework was developed and implemented in order to support goals of an e-learning university. The paper will discuss the aim of assessment to support institution-wide e-learning goals, the benefits and challenges of this approach to assessing the development of content for e-learning, a description of the evaluation framework and the assessment tool GUIDE (Guided Instructional Design Evaluation), and the results based on implementation for the development of graduate and undergraduate programs.


This paper examines the benefits and challenges of for an institution to support cooperative learning through assessment. The paper will focus on the key issues behind the design and management of teamwork assessment, and the development of guidelines to undertake the implementation of an institution-wide approach to cooperative learning.
Informance is a design technique that calls on designers to role-play and perform design ideas. It combines performance, role-playing and improvisation, which enables designers to enact situations physically as well as conceptually [3]. We found that informance design used in design ethnography bridges ethnographic observations with the generation of design ideas. This approach emphasizes physical and embodied observations that are often overlooked, and guides formal analysis in the later stages of the process. These aspects are important; firstly because of the typical length of ethnography studies, designers need to be ready to deliver preliminary outcomes during ongoing studies. Secondly, ethnographic analysis tends to be textual rather than embodied and perceptual. This paper illustrates the value of informance in design ethnography by describing the process by which our observations led to design patterns and scenarios. In our case, we discuss a scenario based on a six-month ethnography study of everyday design.

Laura is the main protagonist in Janet Cardiff’s “Eyes of Laura” web project for the Vancouver Art Gallery (http://www.eyesoflaura.org). We never see Laura. Her presence is felt in the opening lines excerpted above. She provides us with a minimal physical description and her occupation (a security guard) through a disembodied voice, heard on the home page of the web site. However, the desire in her life is made clear: for something to happen.

This paper proposes a new design method. Key aspects of the method are the concept of “framing the problem”, and the need to design “in-the-world” throughout the design process. The method is based on Henrik Gedenryd’s notion of “interactive cognition” and Donald Schon’s ideas as embodied in his concept of “reflective practice”. The paper discusses the related elements from Gedenryd. It describes preliminary principles for an interactive design method, and presents examples from a research project on an augmented reality interface for museums, known as ec(h)o.

The aim of this paper is to demonstrate how evaluation models for e-learning were designed through the use of pattern language in order to support a related set of scalable goals including institutional, course, and learning activity level goals. The paper will discuss the aim of assessment to support this range of inter-related learning goals, the benefits and challenges of this approach to assessment in e-learning, the use of pattern language to develop evaluation models, and the results and examples based on implementation at the graduate and undergraduate levels of an interdisciplinary technology and arts program at the Technical University of British Columbia and Simon Fraser University.


This paper examines Christopher Alexander’s pattern language as an alternative model for design that is both non-rational and non-linear. The paper suggests that pattern language is a response to the development of rational models of design. The paper provides an overview of the critiques of “design science” by Richard Coyne, Donald Schöen and Henrik Gedenryd. It argues against the current emphasis of stand-alone patterns over the structural relationships of pattern language as embodied in Alexander’s network concept of language. Alexander asserts that pattern language is a network, and it becomes evident that the further development of pattern language as a non-rational model for design will require a potential rethinking of pattern language in the context of current network theories.
EXPERIENCING INTERACTION DESIGN: A PRAGMATIC THEORY

Volume 2

Ronald L. Wakkary