SOCIAL PRESENCE IN A CO-LOCATED NETWORKED ART INSTALLATION

by

Jack Nicholas Stockholm III
B.A. University of Detroit Mercy, 1996

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

In the
School of Interactive Arts & Technology

SIMON FRASER UNIVERSITY

Fall 2010

APPROVAL

Name: **Jack Stockholm**

Degree: **Master of Arts**

Title of Thesis: **Social Presence in a Co-located Networked Art Installation**

Examining Committee:

Chair:

Dr. Marek Hatala
Associate Professor,
School of Interactive Arts and Technology

Dr. Ron Wakkary
Senior Supervisor
Associate Professor,
School of Interactive Arts & Technology

Dr. Philippe Pasquier
Supervisor
Assistant Professor,
School of Interactive Arts & Technology

Dr. Thecla Schiphorst
External Examiner
Associate Professor,
School of Interactive Arts & Technology

Date Defended/Approved: 2010 - 11 - 15

Declaration of
Partial Copyright Licence

SFU SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Simon Fraser University
Burnaby, BC, Canada

Last update: Spring 2010

ABSTRACT

This thesis describes the exploratory art research project, Eavesdropping, which aims to increase social presence between individuals in shared public spaces. This internet-based system is designed to create an audio ecology in localized, networked environments like a cafe where several computer users are gathered by playing audio from each participant's laptop and capitalizing on the personal affinity and proximity between individuals and their computers by attracting the attention of others via audio.

Two versions of the system were created, the first passive, the second interactive which attempted to increase engagement and immersion to subsequently increase social presence by adding self-representation with the audio and meaningful interaction with the system. User studies involving an engagement and immersion questionnaire designed for this project and an established social presence questionnaire, showed that differences between the versions had a significant negative impact on engagement but did not create an overall change in social presence.

Keywords: *art research; acoustic ecology; interactive art; engagement; immersion; social presence; audio installation*

DEDICATION

To my wife, Veronica Zammitto.

ACKNOWLEDGEMENTS

I want to thank to the members of my committee: Ron Wakkary and Philippe Pasquier, for their time, their knowledge, their financial support, and fruitful discussion on this work.

This work has benefited from discussions, comments, and words of support from many colleagues and friends, notably: Veronica Zammitto, Josh Tanenbaum, Karen Tanenbaum, and Aaron Levisohn.

I also want to express my gratitude to Maia Engeli and Susan Kozel for all the encouragement and discussion at the beginning of this project that got the ball rolling, and to Jenifer Kent and Matthew Scheatzle for the holding the door open as I took the first steps.

This has been a very long project and I am surely missing several people who played important roles.

Thank you for financial support to Rhizome.org.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1: INTRODUCTION

1.1 Internet Connectedness

The pervasive reach of internet accessibility is enabling an increasing level of connectivity between inhabitants of both major metropolitan areas as well as more rural settings. With the growing evolution of social networking applications, internet users are able to maintain social connectedness with larger and more diverse circles of friends. While this network connectedness offers the potential of greater intimacy with remote acquaintances, it is often seen as having an isolating effect in local environments. People stand at a bus stop inwardly focused on their cell phones and gather in a cafe intently staring at their laptops. Even co-located individuals cannot be bothered to escape their focus on their computer systems and will send chat messages to fellows in the same room. Our mobile technologies increasingly offer us intimacy at a distance while sacrificing our attention to those in nearby.

A trend has begun to bring attention back to the local. GPS enabled phones and node-aware internet access allow web sites to share information about individuals in close proximity to us. Sites serve up the Twitter feeds of those in nearby latitude and longitude or alert us to events or gatherings or even traffic in our vicinity. There is still significant untapped potential for the network to operate as a sixth sense. How will we perceive others on the network who share our physical space – how will we represent them?

Simultaneous internet access for people in a public space offers a means of interaction beyond physical proximity; we share the same bandwidth, our data comingles as we engage in private acts on our solitary screens (Manovich, 2002). This thesis explores tapping this digital connection to augment the social connection between individuals in shared public spaces via networked audio performance.

1.2 Motivations

Social spaces are alive with audio cues to the dispositions of their participants projected through culturally reinforced sounds such as coughing to indicate irritation or laughing loudly to attract attention. Even our body language makes sounds which help to communicate our state to those around us as we fidget with a pen or tap our feet. Most of this audio fades into a continuous drone of background noise that we tune-out, however, some audio events pique our interest and raise our awareness of the source of the sound.

The Eavesdropping project is motivated by the intention to produce an engaging web audio art environment that compliments the audible interactions which occur in public spaces in order to heighten awareness between individuals in the space and to increase social presence between those individuals. This is accomplished via a system which mimics the social acoustic ecology and auditory gesture encountered in public spaces (Davis, 1977; Truax, 1999, 2001) to increase shared experiences by creating an augmented reality where mood-based audio events occurring at participants' computers draw attention and create a sympathetic interaction.

This work stems from my artistic practice as a musician which is situated in two nearly opposite poles within the field of music. As an instrumentalist I am an improviser

with a keen ear to the sounds created by those around me; from the musicians who share my stage, to the audience members who shuffle, move and vocalize responses to the audio I am producing. Playing live music which drives as well as responds to others in the environment is a negotiation which has aspects of conversation and argumentation, leading and following, acceptance and resistance, and in the best of relationships, almost feels like a game. On the other hand, music can be structured into complex forms, can be numerically represented, and can be 'programmed' in advance. My interest in music composition has not followed traditional musicology but rather has been driven by the affordances of digital media and the construction of new interfaces for musical production and experience. When I received my first computer as a child, a Commodore 64, I learned to program basic code by exploring the potential of its 3-voice synthesizer in creating random music, by taking data that was meant for other purposes and seeing what it might sound like if I fed it through the synthesizer, and by creating my own generative patterns for the synthesizer to perform. In more recent work I have utilized data from the internet to as my explorative audio source. My WebMIDIfier converts web pages to audio by translating the text on the page to a listenable audio stream. A project I titled Market Sounds reads real-time stock ticker data and sends segments of a stock's chart to a synthesizer as a wave form. The Eavesdropping project extends the interface beyond a single user in front of a computer and is intended to explore an audio medium where several users are in front of several computers and they can all hear the sounds each other's computers are making.

1.3 Context and Goals

The research and design of this project are situated across the two disciplines of art practice and human-computer interaction (HCI). As an art performance, this project is intended to create an interesting aesthetic experience for participants which instigates new ideas about the relationship between individuals in public spaces and raises awareness of the social soundscapes in these spaces. As arts research, Eavesdropping explores networked public spaces as a venue for interactive performance. It provides a contemporary context to examine how people interact when participating in socially-aware performance; sitting behind their laptops and looking up and around the room, when they feel safe, or when their safety bubble has been broken, or when the events in the room pique a curiosity too strong to resist. This references traditions of social art like Happenings or Fluxus, where there is an interest in awakening the role of the participant-audience. Eavesdropping aims to situate a sound source at the location of each participant, and explores how much control people need to feel immersed and socially connected. From an arts practice perspective, this project challenges the artist to engage the participants while encouraging social interaction and through an aesthetic phenomenon. From a systems design perspective, a system must be created which is aware of each participant and can diffuse audio into the room to mimic a varied acoustic ecology which includes both continuous background sounds and attention grabbing sound events. Sets of music files must be created to cover a broad range of audio characteristics. An interface must be designed which invites participants to feel involved while not distracting from awareness of the social space. This thesis is written from the multiple perspectives of an artist, a programmer and an interaction designer. The

audience for this thesis will be those artists exploring interaction at the computer interface in localized performance.

The interface is the crossover between art practice and HCI. The interface designer develops the relationship between user and computer, between participant and art system. At the interface the participant is invited to become interested and invested in the performance that is about to occur. This investment leads to engagement where the participant is committed to fully involving himself in whatever may transpire. Throughout the performance the interface design will remind the participant that they are still involved by ensuring immersion into the sonic environment. The proximity to the sounds at each computer reinforces to the participant that the sounds are meant to be an extension of himself and that every other participant in the room shares that relationship. The computer enacts a cycle of requesting new audio from the server and returning to play it at the interface – each participant is connected to each other via the umbilical connection to the server as audio host. The model for this interaction is constructed through the participant's engagement and immersion with the audio art system and its impact on the social presence felt with other participants.

Three studies were run to accomplish these goals. The first study explored the audio and interaction in the context for this social art performance: public cafés. This study examined how café goers respond to audio events in the environment and how these responses caused interpersonal interactions. The results showed that audio events initiated increased awareness of others that were in close proximity to the audio source. The second set of studies involved pilot performances of each version of the art system in context. These studies showed that the performance system met the original goals and

raised several questions regarding interaction and control when user's laptop was co-opted for a performance and when participants were being associated with audio that is not of their choosing. The final study was a quantitative comparative analysis between the two versions of the performance system to assess differences between the two systems based on a model of engagement, immersion and social presence constructed to asses interactive social art performance. The model showed promise in evaluating this type of system and found that increasing interaction had a significant negative impact on engagement while other factors showed no statistically significant variance.

1.4 Thesis Organization

Chapter 2 discusses the foundation in art practice through which this research was conducted and situates the work in the context of similar networked audio performance. It goes on to develop the theoretical framework for a model of social presence enabled through engagement and immersion at the interface of the browser-based art installation, Eavesdropping.

Chapter 3 details the functioning of the server system and outlines the differences between the two versions which were developed, one passive, one interactive, one for performing mood compositions, one to perform audio representations of the moods of participants.

The methodology for this arts research path is fully discussed in Chapter 4. Three studies define the different stages in the development of this project: the first, a contextual participant observation to assess the installation environment; the second, pilot studies using participant observation and informal interview to assess the functionality of

the performance system in context; and third, isolated user studies utilizing quantitative self-report questionnaires to compare and evaluate the two versions of the system to assess the impact of interface design changes on social presence in the performance environment.

Chapter 5 reviews the results and data from the three studies. Chapter 6 explores these results with a wider lens and assesses the impact toward the further development of the project as well as broader implications for interactive audio art performance systems including a review of the effectiveness of the model of social presence and its value for interactive arts practitioners.

Finally chapter 7 reviews the work that has been done in this exploratory project and speculates on ideas which have been opened up as a result of this research.

2: APPROACH AND RELATED WORK

This chapter introduces the approach taken in the development of this artwork and the theoretical framework for the comparative evaluation of its success. It begins with a discussion of arts-based research and related creative pursuits in the field, and follows with a survey of presence work that constructs a model of engagement, immersion and social presence to form a structure for the development of this project and a lens to evaluate user experience through a perspective of HCI for interactive art performance.

2.1 Arts Research

Arts-based research incorporates several traditional and contemporary methodologies for generating, investigating and communicating knowledge in the academic forum. In the last ten years there have been a significant volume of publications and edited collections to define and validate arts practice as a context for academic research (Balkema & Slager, 2004; Holly & Smith, 2009; Leavy, 2009; Mäkelä & Routarinne, 2006; Sullivan, 2010). Traditionally, academic art research has been performed as art analysis and criticism where theorists situate a work within a lineage of related works and discuss art from a distance in its relationship to social and cultural movements. More recently there has been a broadening of accepted research practice in the arts, such as in communicating the critical and creative investigations that artists engage through their research practice as a systematic inquiry which achieves similar goals to methods of inquiry within the social sciences (Balkema & Slager, 2004; Holly & Smith, 2009).

Arts research in the early 21st century is being redefined, though many suggest that one of the tenets of art research is that as a 'post-disciplinary' practice, it cannot be defined in the same way that methodologies are defined in the social sciences (Sullivan, 2010). Validation in artistic inquiry is not based on the probable likelihood of occurrences, nor on the plausibility or relevance of outcomes but on identifying possibilities or transforming awareness. Leavy suggests that arts-based practices are particularly useful for research projects which aim to describe, explore or discover. "The arts provide qualitative researchers a broader palette of investigative and communication tools with which to garner and relay a range of social meanings (Leavy, 2009)." Knowles and Cole, in a broad review of methods and issues for arts in qualitative research, explores the potential of arts practice and its place in the academic community.

> *Arts-informed research provides a context for promoting innovative research that infuses processes and forms of the arts into scholarly work for the purposes of advancing knowledge and bridging the connection between academy and community.*
> *(Knowles, 2008)*

As the number of programs offering a PhD in arts research continues to grow, the academy has come to accept a broad array of different contexts for research with the arts. Elkins, in examining the results of several colloquia discussing PhD-level scholarship in creative works has identified eight configurations for arts research: (1) the thesis as art history, (2) the thesis as philosophy, (3) the thesis as art criticism, (4) the thesis on art from an outside perspective (his examples are natural history, or economics, or engineering), and (5) the thesis about the techniques of artmaking (each of these first five is intended to inform and direct art practice), (6) the thesis and artwork as a combined

interdisciplinary configuration (this relates more directly to arts practice as research where the studio process is evaluated as part of the thesis), (7) the thesis and the creative artwork understood as wholly separate projects, and (8) the thesis itself is an artwork (Elkins, 2004). While Elkins is most intrigued by the last two and their potential for grounded yet radical reinterpretations of thesis criteria, the configuration that is becoming increasingly common in interactive arts is the thesis from an outside perspective. Since interactive arts as a field is often situated at an interdisciplinary boundary between computer science, human computer interaction (HCI), psychology, and sociology, these lenses are often used for the exploration and evaluation of interactive art research. There are many academic conferences each year, some in their third decade, focused specifically on this boundary, such as Ars Electronica, Transmediale, Digital Interactive Media Entertainment and Arts (DIMEA), and New Interfaces for Musical Expression (NIME), to name a few, hosting a vast array of artists and scientists involved in a cross-disciplinary exploration and analysis of aesthetic, entertainment, transformative and productive works. It is in this context that the research project Eavesdropping will be scoped. This is a project defined as an interactive art installation which was conceived and developed through arts practice research but which attempts to achieve a sociological transformation in a public space. The outside lens used to evaluate this transformation is situated in the methodological frameworks of HCI and sociology, in interactive engagement and social presence discussed in 2.4 of this chapter.

> *With its emphasis on constructivism, critical interpretation, and contextualism, arts-based research has spawned a rapidly expanding literature that attempts to draw on wide theoretical and cultural support to further build its status as a viable methodology. (Sullivan, 2010)*

Graeme Sullivan, in a recent comprehensive review of literature on arts research methods, identifies several different foci for research within the arts (Sullivan, 2010). He notes that research can focus on the artist, on the art object, on the art making process, into the creative process, or into the cultural / political role of art. Missing from this list is a focus on the art experience, the long moment in which the audience encounters an artwork and the immediate effect of the artwork on the audience. This seems to be due in part to Sullivan's attention to physical media art rather than performance art or interactive works. In the case of the latter, the 'art object' is less important than the immediate phenomenon created by the process which the audience goes through in the art environment (Potts, 2009). Likewise, research into the immediate experience, while affected by the cultural and political context of the work, does not examine the role of art but looks at the effectiveness of the artist at transforming the audience through a momentary experience.

> *"Contemporary arts practice-led research strategies, especially in interactive art, involve creation and evaluation first in the studio environment and then to a wider audience in an exhibition space. Often this process involves several iterations with different audiences, friends and associates in the studio providing feedback to affect the final production prior to a public showing. These interactive audience experiences are recorded and analyzed and evaluated with respect to audience experience. The artist(s) are then able to reflect upon the results to determine whether the work meets their intentions and expectations"*
> *(L. Candy, Amitani, & Z. Bilda, 2006)*

The research artist frequently goes through several iterations in the development and evaluation of a work of art which may incorporate multiple strategies and methodologies to define the context, the questions and the results of an art project. Edmonds and Candy have outlined a set of trajectories which assist in understanding the

flow of different types of arts and practitioner research. In a detailed study of several arts research projects they have broken down practitioner trajectories into three main elements: Practice, Theory and Evaluation. The main outcome from Practice is the artwork itself; the outcomes from Theory are criteria (design strategies) and frameworks, consisting of conceptual structures used to influence practice, inform theory and shape evaluation; the outcomes from Evaluation are the results. These phases are not typically worked out in advance and may occur in any order, often iteratively repeating a phase multiple times based on knowledge, events or needs identified in earlier phases. Utilizing this system the following model has been identified as the trajectory of the Eavesdropping research project.

1. Theory: Theoretical knowledge drawn from the fields of sonification and acoustic ecology was used to derive a framework consisting of methods for understanding social audio interaction in a public context, particularly as it relates to networked technologies in these contexts.
2. Evaluation: The framework was assessed in field studies involving contextual participant observations and results were analyzed.
3. Theory: The results of the field studies informed the generation of an initial design criteria for an interactive artwork which augmented the social acoustic ecology in a space via participants' networked computers. This initial conceptualization was validated by the arts community in two significant acts of recognition. First, the project design received funding through a commission from Rhizome.org, one of the leading organizations for the promotion, recognition and archiving of internet art (Rhizome.org, 2008).

Second, the project concept and design were published as an article in a Leonardo Music Journal special issue on live interactive performance (Stockholm, 2008).

4. Practice: The initial framework and design criteria were used to develop a first iteration of the interactive artwork.

5. Evaluation: The first version of Eavesdropping was presented in several contexts and user feedback was gathered through participant observation and informal interviews. The work also received some recognition from the larger art community by being selected for presentation in an international art festival, the 404 International Festival of Interactive Art (404 Festival, 2008).

6. Theory: Results of the participant interviews were analyzed resulting in a revised framework for the both the participant's relationship to the artwork (immersion) as well as their interaction with the project during the performance (engagement). A second set of design criteria was devised based on this new framework.

7. Practice: The second version of the project was developed incorporating the design changes. New technologies were used, particularly an artificial intelligence (AI) engine, to implement the design criteria as well as to maintain the interest of the artist in continuing the project. This version of the project received recognition through two conference publications documenting: the updated concept and design in ACM Multimedia 2008 (Stockholm & Pasquier, 2008), and the AI for audience interaction in IEEE Social Communications 2009 (Stockholm & Pasquier, 2009).

8. Evaluation: The second version of Eavesdropping was presented in several contexts and further user feedback was gathered through participant observation and informal interviews. The changes between the two versions had met the criteria raised by participants of the first version but raised new issues.

9. Evaluation: A final evaluation was performed as a comparison between the two versions of the project based on the artist's original criteria for the project, the ability to generate increased social presence between participants in the space. This evaluation consisted of isolated performances and a quantitative questionnaire based on a widely accepted instrument for the measurement of social presence and a questionnaire constructed for this project meant to measure engagement and immersion.

The following section, 2.2, outlines the original theoretical basis for this work as a public networked audio art project. Section 2.3 surveys social presence research, the area of intended effect for the Eavesdropping project.

2.2 Interactive Audio Art

This project takes advantage of the use of laptops in public places like cafés to create an interactive networked audio installation. Participants log into a website and audio plays from each of their computers. Participants are connected via their laptops to the network and to each other via the audio performance.

Gil Weinberg has constructed a compelling categorization of interconnected music networks (ICN) based what he sees as the central innovative concept of the

medium: “the level of interconnectivity among players and the role of the computer in enhancing the interdependent social relations” (Weinberg, 2002, 2005). Weinberg has defined four approaches to the use of networks with music, the *Server*, the *Bridge*, the *Shaper*, and the *Construction Kit*. The *Server* merely provides a network-accessible music making tool to several networked users individually, their experiences are distinct and they can only hear the music they themselves are creating. The *Bridge* provides connectivity between distally located musicians to allow them to play as if in the same room with each other. The system in the *Bridge* approach does not enhance or enrich collaboration but merely provides technical capacity to communicate musically through the network. The *Shaper* approach relies on a generative system to actively provide musical output which can be shaped and adjusted by the input of several networked users. Users’ contributions to this system are not interdependent though they can individually hear and manipulate the output that is modified by all participants. The *Construction Kit* approach offers a high degree of interconnectivity but the interactions are typically asynchronous and sequential. Users evaluate the contributions of others in order to best add their own contribution.

Weinberg’s categorization is situated in the production of music by musicians and non-musicians through the network. While the first version of Eavesdropping lack of any direct interaction with the network by the user, and the second version only allows for the user to provide basic information to the system, the project fits squarely into the model defined by a Shaper approach. The system provides audio to the participants based on how many are in the room, the time they joined the performance, and in the case of Version 2, their contributions to the system. Version 2 users input their moods and

provide feedback to whether the audio matches their moods allowing the system to adjust as it provides additional audio samples. Input from users it not directly interdependent but all contribute to a corpus of knowledge concerning the audio file-to-mood relationships which are utilized throughout the performance and in subsequent performances.

Barbosa looks at location, media, timing, physicality and other factors in constructing a model of network use in music performance (Barbosa, 2003). In Barbosa's map, location is along one axis and interaction is along the other. Within location are co-located and remote and within interaction are synchronous and asynchronous. Like Weinberg's ICNs, Barbosa's examples are primarily focused on systems for making music rather than systems for the intelligent diffusion of audio events in a localized environment via the network.

2.3 Auditory Display and Acoustic Ecology

This project is situated across a boundary of performance and interaction. The initial project concept relies on the relationship participants have with their laptops to connect them to the audio being performed from their computers. This relationship becomes additionally important in that the audio from each individual's computer can be heard by other participants in the room, thus activating a social awareness between participants in the performance (Gaver, 1991). Sounds heard from each computer are associated with the participant sitting at that computer due to the proximity of the participant to the sound. Since sounds have the ability to carry information or communicate an idea, this idea is then also associated with the participant sitting at the computer. One aspect of indirect interaction participants have with the performance is to

acknowledge the sounds coming from their computer. Once they acknowledge a sound, they then have the option to accept or reject the association being made between themselves and the sound. This is typically exhibited through social cues such as making faces or laughing. This is the social connection intended for this project, on one level a sympathy between participants who are all part of the same sound association performance, and on another an empathy where participants can come to understand each other on a deeper level by seeing how each other react to the different types of sounds (Dey & Ed de Guzman, 2006; Edward S. De Guzman, Yau, Gagliano, Park, & Dey, 2004).

There are several auditory factors at play in this interaction. One is the utilization of the laptops as a medium for auditory display. Nearly all audio signal elicits some sort of response in a participant, whether it be the perceptual response of merely acknowledging the presence of the sound, its timing, rhythm or timbre, or a more affective response, for instance triggering an emotion or referencing a familiar sound that has cognitive associations for the listener (Cohen, 1998; Cumming, 2000; Hermann & Ritter, 2004; Kivy, 1989; Kramer, 1994). Hearing a sound that piques one's interest draws their attention from visually focusing on the computer in front of them to look to the source of the sound and thus to make the social connection to the person sitting at its source. Utilizing sound to communicate in this way frees up the eyes to look around the room and make these social connections (Ballass, 1994).

As a requirement for this system to work, participants need to clearly hear individual sounds coming from each computer rather than all of the sounds blending into general room sound. Truax identifies three characteristics essential to developing a

functional acoustic ecology: there must be a *variety* of sounds, a *complexity* of sounds in relation to the types and levels of information they contain, and there must be a *balance* within the environment based on spatial and temporal as well as social and cultural constraints (Truax, 2001). This creates a challenge for the system designer to ensure that sounds are categorized effectively to allow the system to distribute varied sounds in the environment, as well as for the musician in creating the sounds to ensure that a broad possibility of sounds are available to the system. These requirements will be addressed in Chapter 3 where the system is described in detail.

2.4 Social Presence Research

In order to evaluate social presence we must first establish workable definitions of both presence and social presence. While there is a significant amount of literature on presence as it relates to telepresence and virtual environments, the newer fields of ubiquitous computing and mobile devices which allow presence to play a role in co-located interactions supported by laptops and portable networked electronics have not been as well explored. In this section the author brings together research in presence which has often been defined by factors involving immersion and engagement to discuss a model of social presence that can be applied to low interactivity, co-located, mediated environments.

2.4.1 Presence

Presence refers to a psychological state, specifically the subjective feeling of being transparently connected to a media experience (Dow, Mehta, Harmon, MacIntyre, & Mateas, 2007). Presence is often defined loosely as the sense of "being there" in a

mediated environment. Lombard and Ditton (Lombard & Ditton, 1997) bring together six different conceptualizations or definitions of presence from a broad review of literature. The first they describe is 'Presence as Social Richness'. In this definition, which unites social presence theory (Short, 1976) and media richness theory (Rice, 1992) in relation to communication, presence is described as the extent to which a medium is perceived as sociable, warm, sensitive, personal or intimate when it is used to interact with other people. This relates to two key concepts originally applied to non-mediated communications, intimacy and immediacy. Mediums which support higher adjustability of proximity, eye contact, facial features, gesture, and other behaviours offer optimization of the overall level of intimacy (Argyle & Dean, 1965). For social presence in mediated, localized environments, the importance is in ensuring that the system allows this adjustability and does not overly interfere with physical access to cues between participants, such as physical gestures, eye contact, or audible communication, and can provide an information-rich sensory space to amplify or expose subtle social cues.

Another relevant conceptualization Lombard and Ditton describe is 'Presence as Immersion' which emphasizes the idea of perceptual and psychological immersion. Immersion is a concept frequently raised and evaluated as a factor in presence (Bystrom, Barfield, & Hendrix, 1999; Short, 1976; Witmer & Singer, 1998). Perceptual immersion, described as when the senses are immersed to the extent that the outside world fades from awareness, is a key factor in analysis of both virtual environments as well as mediated reality. While in virtual environments the intent is that awareness of the system fades from perception leaving the participant immersed in a replacement reality. In augmented reality and mediated environments the intent is often to integrate the system into the

physical surroundings leaving participants in an enriched environment where they lose a sense of the system as a contrived tool for the augmentation of the sensory ecology. On the other hand, psychological immersion is described as a state where users feel involved, engaged, absorbed and engrossed. In the following section, 2.4.1.2, engagement and engrossment are further explored as factors in immersion.

Lombard and Ditton attempt to unify their different conceptualizations of presence in mediated environments into a unified definition, identifying presence as the "perceptual illusion of non-mediation". They note that the term "perceptual" indicates a phenomenon involving continuous responses of the human sensory, cognitive and affective processing systems. An "illusion of non-mediation" occurs when a person fails to perceive or acknowledge the existence of the communication medium in their environment.

Witmer and Singer (Witmer & Singer, 1998) explored presence and immersion in the context of virtual environments and defined presence as "the subjective experience of being in one place or environment, even when one is physically situated in another." They identified presence as "a normal awareness phenomenon that requires directed attention and is based in the interaction between sensory stimulation, environmental factors that encourage involvement and enable immersion, and internal tendencies to become involved." This notion of "involvement", described as a "psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events", further develops Lombard and Ditton's discussion of engagement and engrossment as factors of psychological

immersion. They suggest that both involvement and immersion are necessary for experiencing presence.

Slater and Wilbur (Slater & Wilbur, 1997), also working on virtual environments, discuss presence as both subjective and objective. Subjective presence relates to the degree of "being there" a person feels in a virtual environment, whereas objective presence relates to the extent to which a person behaves in the virtual environment the same as they would in a non-mediated environment. They suggest that the fundamental idea is that participants who are highly present should experience a virtual environment as the more engaging reality than their surrounding environment. Higher levels of subjective presence correlate to higher levels of immersion.

Each of these foundational definitions of presence offers contexts to build a model which informs how we affect and measure social presence. In an augmented reality environment with fairly low interactivity, the key relevant factor in developing a sense of presence is immersion.

2.4.1.1 Immersion

Immersion as defined by Slater and Wilbur is a quantifiable phenomenon describing the extent to which a mediated environment is *inclusive* (I), *extensive* (e), *surrounding* (s) and *vivid* (v). While these factors were originally identified for use with virtual reality environments, only one, *inclusive*, as defined, is not well suited to describe a related factor of immersion in augmented reality or mediated environments. Slater and Wilbur discuss *inclusive* as the extent to which physical reality is shut out, however augmented reality typically involves a meshing of communicative information with the

physical world to create an ecology where both come together as the mediated environment.

The remainder of Slater and Wilbur's factors fit well within an analysis of mediated environments. *Extensive* identifies the range of sensory modalities displayed. *Surrounding* identifies how panoramic is the scene created rather than limited to a narrow field. *Vivid* identifies the resolution or sensory fidelity of the representation.

Slater and Wilbur go on to bring up two additional factors relevant to immersion. The first is *proprioceptive matching*, more relevant to challenge-based immersion in virtual reality where the avatar's movement matches that of the player. The second is that immersion requires an effective *self-representation*, that the user has an avatar in the virtual environment that matches their psychological notion of how they should be represented in that environment. This concept can be extended in multi-user environments to suggest that each player has a self-representation that is both a matching representation of user expectations and recognizable by the other users. In virtual reality these avatars are replacements or stand-ins for the actual users. In augmented reality the avatars map onto physical reality to increase the information available to the users about the physical object or person that the avatar represents. Augmented reality, unlike virtual reality, doesn't intend to increase the realism of the mediated environment, but to increase the relevance of the representations (Hoorn, Konijn, & Van der Veer, 2003).

Immersion in game design is often related to Csikszentmihalyi's concept of flow, defined as a "deep but effortless involvement, reduced concern for self and sense of time" (Csikszentmihalyi, 1991; Sweetser & Wyeth, 2005). Ermi and Mäyrä propose a model of immersion in game worlds which is differentiated in three forms: sensory immersion,

challenge-based immersion, and imaginative immersion (SCI) (Ermi & Mäyrä, 2005; W. IJsselsteijn, de Kort, Poels, Jugelionis, & Bellotti, 2007). Sensory immersion refers to the extent to which the surface features of a game have a perceptual impact on the user. Challenge-based immersion relates to the cognitive and motor skills required to overcome the game's objectives. Imaginative immersion identifies the success of the richness of the mediated environment to create an imaginary world that the player inhabits. McMahan, in a similar mode, suggests that presence is the result of perceptual (sensory) and psychological (imaginative) immersion (McMahan, 2003). While these models apply well to highly interactive environments like video games, the does not work ideally for the low interactivity of a more passive art installation.

Brown and Cairns describe a progression of immersion, also in game environments, based on increasing levels of involvement which related directly to Lombard and Ditton's earlier described definition of psychological immersion in presence (E. Brown & Cairns, 2004). Engagement, the first stage of immersion that must occur before any other, is described as the willingness of the gamer to invest time, effort and attention. The barriers to engagement are access (if it captures their interest and provides an amount of and effectiveness of control that meets expectations), and investment of time and effort to learning the rules and environment. Engrossment, the second stage, is defined as when the game features combine in such a way that the gamer's emotions are directly affected and they become emotionally invested in the experience. The barrier to engrossment is game construction, a semi-mystical quality identified by gamers through visuals, interesting tasks, and plot. Total Immersion, the final stage, is described as the feeling of being cut off from reality and detachment to

such an extent that the game is all that matters. The barriers to total immersion are empathy and atmosphere, the growth of attachment with characters and their self-representation and the relevance of atmospheric game construction elements. Interestingly Brown and Cairns use their term Total Immersion and 'Presence' interchangeably. Again, this model relies heavily on significant game interaction to discuss levels of immersion. While there is some interactivity in Eavesdropping, it is not present at the level of engrossment described by Brown and Cairnes. Still, this model provides an interesting relationship between engagement and immersion which offers the possibility of discussing facets of interactivity at a low level as is present in the Eavesdropping audio performance. This is explored further in section 2.4.1.2.

Of the models discussed here, the Slater and Wilbur definition most suits the needs of this project. The factors described by this model are: *inclusive, extensive, surrounding, vivid, proprioceptive matching,* and *self-representation.* As is described above, *inclusive* and *propriceptive matching* are too particular to virtual worlds and are not applicable in our model. Additionally the factor *self-representation* very narrowly describes the virtualization of the user which, while it does have a relationship to augmented reality, doesn't fully describe the phenomenon. For the purposes of our environment, we propose merging the dimension of *self-representation* with the notion of *inclusive* to more appropriately align these factors to a mediated environment. In this case we are looking to define a phenomenon where the users hear an effective representation of themselves that fits well within the augmented environment (rather than transports them to a virtual environment). It is inclusive and not out of place with the environment but compliments their presence, thus increasing their sense of immersion.

This new definition of *inclusive*, combined with the remaining three factors of *extensive*, *surrounding* and *vivid* make up a clear model which we can use to design methods to affect and measure immersion in the type of low interactivity augmented reality art performance environment particular to Eavesdropping.

2.4.1.2 Engagement

Engagement is a vast, heavily researched subject that is often targeted as the ultimate goal in interactive and video game system rather than a sub-dimension of immersion. Focusing the broader definitions of engagement to those that are particular to our needs in describing the impact of interactivity on immersion requires some narrowing of the field. O'Brien and Toms, in order to develop a generalized instrument for measuring engagement, assembled an extensive definition through a broad, multi-disciplinary literature review on the subject and define engagement as:

> *A cognitive, affective (specifically intrinsic motivation) and behavioral state of interaction with a computer application that 'makes the [user] want to be there'. Engaging interactions were thought to involve attention, intrinsic interest, interactivity, perceived control and choice, functionality, and motivation ... [E]ngagement is impacted by: media richness through the use of animations and video; format (e.g., text, audio, and video); interactivity and exploration; communication or socialization with others; aesthetics and sensory appeal; intellectual challenge; and affective involvement. [Engagement] is both a process and product of interaction; its intensity may change over the course of an interaction depending on the combination of users' needs, goals, emotions, actions, and thoughts, or the format, visual presentation, and organization of the computer interface.*
> *(O'Brien & Toms, 2010)*

In the context of interactive art, Batras defines engagement as "a transformative dialog between the user/ participant and the system ..., during which user's/participant's

initial intentions and expectations about the art object may evolve" (Batras, 2009). This evolution of engagement as a process has been further detailed by Edmonds in an effort to explore design features with an effect on creative engagement. He defines attractors as those things which encourage the audience to take note of the thing in the first place, sustainers as those attributes which keep the audience engaged during the initial encounter, and relaters as aspects which encourage a continuing relationship to grow so the audience remained engaged and returns on future occasions (Zafer Bilda, Edmonds, & Linda Candy, 2008; Edmonds, Muller, & Connell, 2006).

Key areas of engagement which do not overlap our prior definition of immersion and target the dimensions of interactivity relevant for our environment can be defined as: *Intrinsic Interest, Interaction, and Control*. *Intrinsic interest* describes those features which attract a person to the system and sustain their curiosity about the future outcomes which may transpire. *Interaction* describes effective user agency where the user feels they are affecting the outcome of events in a real and meaningful way. *Control* identifies issues of mechanics and free will, describing whether the user is able to understand the interface, whether the interface is easy to operate, and whether they feel they have a choice of when to manipulate the controls or not.

An individual's experience with the performance environment of Eavesdropping can thus be described via a model where *Engagement*, affected and described by *Intrinsic Interest, Interaction*, and *Control*, correlates to *Immersion*, affected and described via *Inclusive* (situated self-representation), *Extensive*, *Surrounding* and *Vivid*. These experiences are individual, and while they may be impacted by other participants, it is not necessary for other participants to be present for these effects to occur. The remainder of

this section extends this model to social effects in a mediated environment. Immersion is a term that is often interchangeable with Presence and thus correlates directly with Social Presence.

2.4.2 Social Presence

Primary research in social presence, performed for telecommunications technologies was conducted by Short, Williams and Christie in 1976 (Short, 1976). They define social presence as (1) "the degree of salience of the other person in the interaction" and (2) "the consequent salience of the interpersonal relationships". Recent work in social presence has focused more on evaluating and increasing the affordances to social presence offered by a mediated environment such as telecommunications, virtual worlds, games or augmented, mixed realities (Kreijns, Kirschner, Jochems, & van Buuren, 2004). Some of the most compelling recent scholarship in social presence can be found in the work of Biocca and Harms, who propose the following definition, "Social presence in a mutual interaction with a perceived entity refers to the degree of initial awareness, allocated attention, the capacity for both content and affective comprehension, and the capacity for both affective and behavioral interdependence with said entity" (F. Biocca, C. Harms, & Gregg, 2001; Frank Biocca, Chad Harms, & Judee K. Burgoon, 2003; C. Harms & F. Biocca, 2004).

Biocca and Harms are looking for generalizable dimensions of interaction between individuals in a social medium which factor into a measurable value of social presence. Their work has produced several iterations of a general instrument to measure social presence, the Networked Minds Measure of Social Presence (NMMoSP), which has been adopted for evaluation of the Eavesdropping project as is discussed further in

Chapter 4: Methodology. Biocca and Harms have factored social presence into six key factors. *Co-presence* is the degree to which an observer is aware of the presence of another or others, the sense that one is not alone and the sense that the other is aware of them. *Attentional Allocation* identifies the amount of attention the user allocates and receives from another. *Perceived Message Understanding* is the ability of the user to understand messages being received from another and their perception that the other is understanding their messages. *Perceived Affective Understanding* is the user's ability to understand another's emotional state and their perception of the other's ability to understand their emotional state. *Perceived Emotional Interdependence* describes the extent to which a user's emotional state affects other's emotional states and is affected by the emotional states of others. *Perceived Behavioral Interdependence* is the extent to which a users behavior affects and is affected by others.

The Biocca and Harms definition is pervasive in social presence studies and effectively satisfies the needs of this project to provide a comprehensive framework for understanding social presence in an interactive audio art installation. The complete model of Engagement contributing to Immersion contributing to Social Presence, utilized as a framework for the development and analysis of social presence for the low interactivity, co-located, audio art performance Eavesdropping, can be seen in Figure 2-1 below.

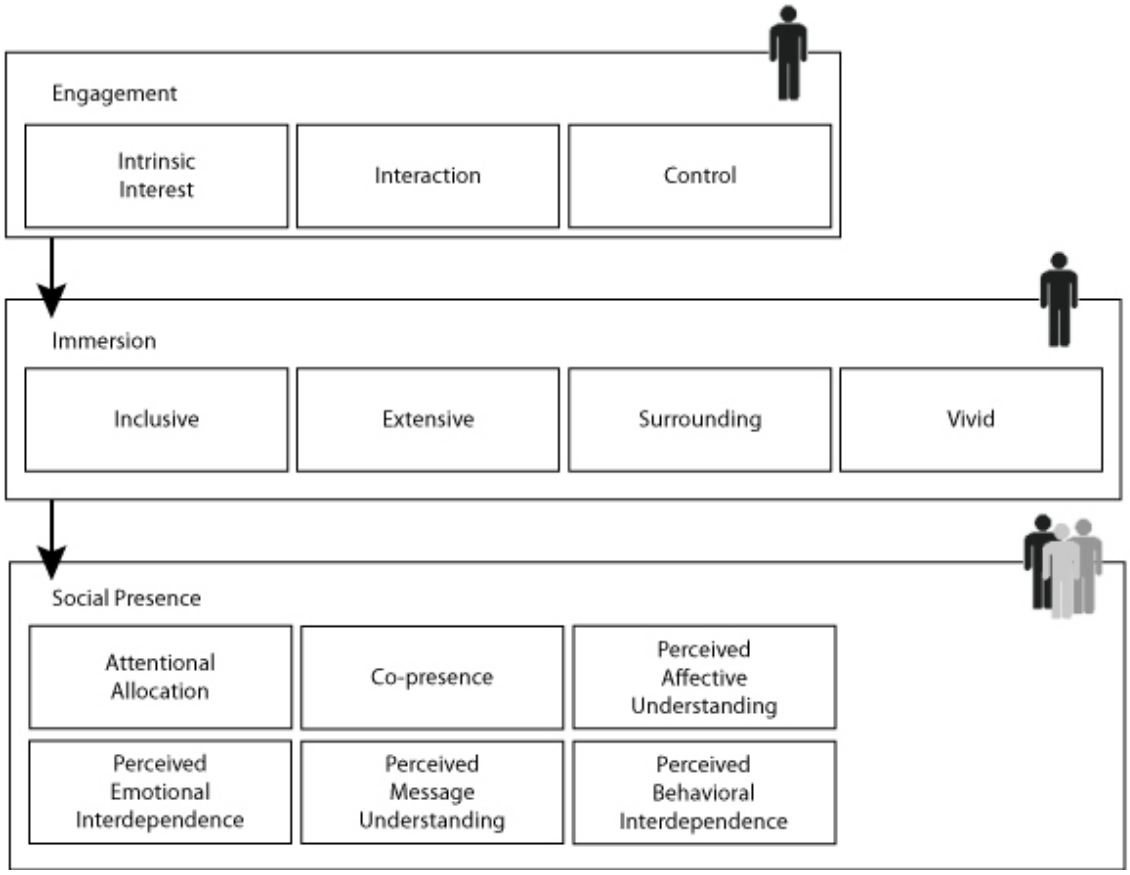


Figure 2-1: Model of Engagement / Immersion / Social Presence

3: EAVESDROPPING: SYSTEM DESCRIPTION

Eavesdropping is an internet-based, client-server architecture comprised of three components: (1) an audio preparation interface, (2) an interactive performance interface, and (3) a server-based, conductor. The following chapter details these three components and the differences between them in the Version 1 and Version 2 systems. The first section, 3.1, gives a high-level walkthrough of a performance. The second section, 3.2, details the audio preparation interface where the audio sets are created and the composition environment particular to Version 1. Section 3.3 presents the interactive interface which is available to participants in the performance. Section 3.4 details the decision making core of the system, the Conductor modules, and how they affect the performance. Finally, 3.5 details the machine learning system involved in the mood matching process for Version 2.

3.1 Performance Walkthrough

Eavesdropping is a web-based system designed for public spaces where several computer users are gathered, such as a café. Anyone can initiate a performance by visiting the website, selecting a composition and entering the name of their location into the system. The initiator would then announce to anyone in the room who is willing to participate that the performance is about to begin and they can join by visiting the website and selecting the performance listed at their location. A simple password is then presented which the initiator will need to announce in order for participants to join the performance. The system will adjust audio based on how many participants are in the

room, the password system prevents people who are not present at a location from joining a performance and skewing the results. Once everyone has joined, they will be taken to one of two performance screens depending on which version they have joined.

3.1.1 Version 1

Version 1 is a mood composition performance system. A musician creates a composition based on layered moods in advance using the audio preparation system described in section 3.2. Individual audio files based on those moods are played at each participant's laptop computer. The audio files are also loaded by the musician in advance but the system decides which files to play at any given time based on the number of participants and where the performance is along the mood composition timeline.

When Version 1 participants have joined the performance they are taken to a simple screen (described in detail in Section 3.3) which shows the playing audio file and some basic information about the performance, for instance, the performance location, the number of players, and the time remaining.

Audio files are between 5 and 60 seconds long depending on the composition. Once an audio file has completed playing at a participant's computer the browser will request another file from the server. The server will continue playing different audio files at the browser until the performance has completed. During the performance participants will hear the audio from their own computer as well as the audio from other participants' computers around the room.

A performance runs for an amount of time identified by the musician when creating the composition. When the performance is finished a simple thank you screen is shown to participants.

3.1.2 Version 2

Version 2 is an interactive mood performance system designed to provide an auditory display of each participant's mood. Again a musician uploads audio files to the system in advance but there is no underlying composition in Version 2. When Version 2 participants join the performance they are shown a mood selection screen where they identify their current mood and click a button to continue. The system then begins sending audio files to the browser to match the mood the participant selected.

The performance screen for Version 2 also shows the currently playing audio file and basic information regarding the performance, as well as the mood the participant selected. The mood cannot be modified during the performance. The Version 2 performance screen also presents an optional question to participants regarding whether the current audio matches the mood they set. Participants can answer Yes or No by clicking the appropriate button. These buttons refresh each time a new audio file is sent to the browser.

A performance runs for an amount of time identified by the initiator when they create the performance. When the performance is finished a simple thank you screen is shown to participants.

3.2 Audio Preparation

The audio preparation interface is accessible to musicians via a password protected web page and in both versions has two main functions: definition of audio sets, and working with audio files contained within a set. The first function, defining audio sets allows a musician to create a new audio set or to edit the properties of an existing set. A performance can utilize any one audio set, thereby imparting a specific style or genre as defined by the files the musician has included in the set. The properties available to audio sets are: the name of the set, the creator of the set, and a description identifying the style of the audio files contained in a set. The second function available in this interface, working with the files contained in the set, allows the musician to upload, delete, and tag audio files within a set.

The Version 1 system additionally includes a composition environment for the creation of layered moods along a timeline described in section 3.4.

3.2.1 Audio Tagging

A musician with secure access uploads mp3 files to a set via the audio preparation interface (see Figure 3-1). First, the file is given a name for identification by the musician. Second, the length in seconds of the audio file is required. This allows the system to inform a participant's web browser when a file has completed so it can request a new audio file. Finally, the musician tags the file with a variety of representational characteristics which denote formal and abstract properties of the audio. These characteristics were developed to give the system basic information to associate audio files to moods (energy, emotion) as well as to apply specific formal musical properties in

layering multiple audio files (tempo, timbre, harmony/melody, and sound density). The system requires each audio file to be defined by all six characteristics.

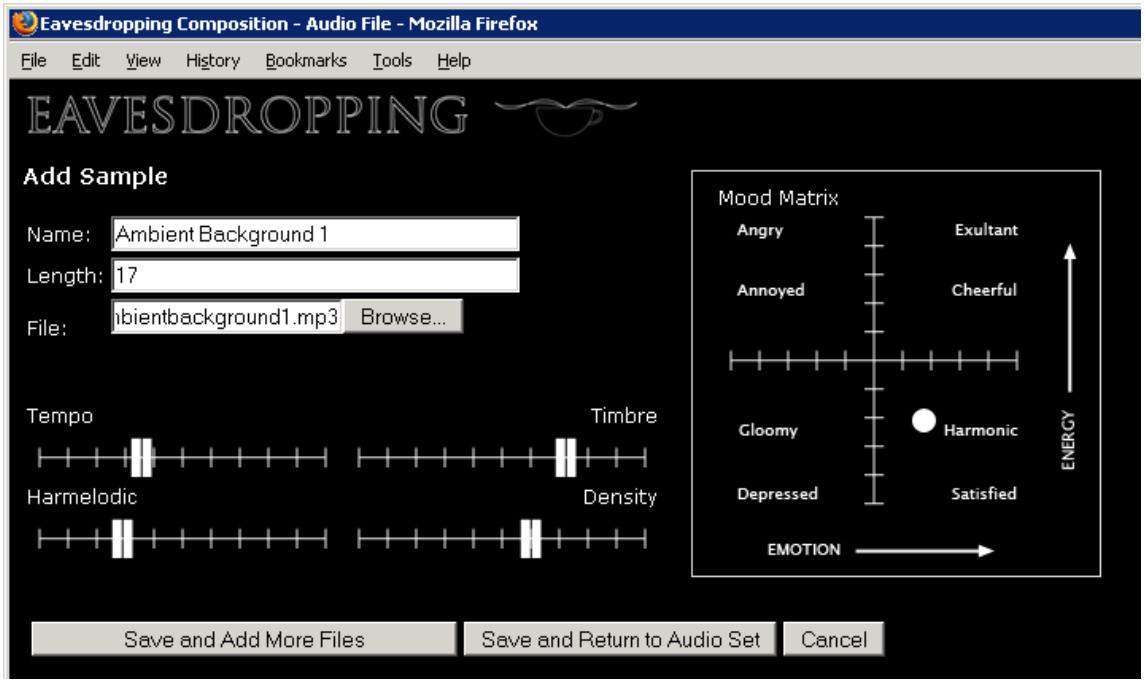


Figure 3-1: Audio File Preparation Interface

Four of the characteristics relate to formal musical parameters which allow the system to provide variety in the mix of audio that is being performed. All are abstracted from formal musical definitions and are normalized on a linear, numeric scale from 1 to 100. *Tempo* translates to the speed of the audio represented. *Timbre* lies on a scale from noisy or atonal to smooth or sonorous. *Harmelodic* relates to the clarity of the voicing and ranges from rich harmonies to solo melodies. *Density* identifies whether the file is dense or sparse, indicating whether it would be better played alone or with other samples. *Density* allows the conductor to adapt to the number of participants by selecting samples of minimal density when the participants are many or selecting rich, high density samples

when the participants are few. These characteristics are set by the musician via sliders on a web page as is exhibited on the lower left side of Figure 3-1.

3.2.2 Mood Classification

Moods in this system are represented by a circumplex ordering of affect around two dimensions (J. A. Russell, 1980; Thayer, 1989). On one dimension, *valence* relates to the pleasantness of an affective experience, on the other, *arousal* relates to the perception of arousal associated with such an experience. Therefore audio with a depressed or gloomy mood will have a low valence and low arousal, and audio with an angry mood will have a low valence and high arousal. Likewise, audio with a contented or satisfied mood exhibits a high valence and low arousal while audio which is ecstatic or exultant exhibits a high valence and high arousal. Note that while the term mood is often used interchangeably with emotion, in this context mood is defined as having a longer duration than emotion which is more episodic (Feldman, 1995).

Mood selection is presented to musician and participants via a simple mood matrix with a moveable dot which can be dragged to a location to indicate mood. Figure 3-1 shows the Eavesdropping audio preparation interface with a mood selected at (2; -2) on the grid. Position is evaluated where the center of the dot falls between the hash marks on the graphical grid. Note that simplified terms have been used to identify the axes in order to be understandable to a wider audience, with *emotion* used for valence and *energy* used for arousal. This graphical map offers several advantages over common approaches to mood classification in audio using discrete adjective descriptors. First, the model is quite simple to understand for participants as it only applies two scales and therefore requires minimal time investment versus reading through and selecting from a

dictionary of mood adjectives. Second, adjective descriptors have been found to have a variety of meanings across a range of participants (Eladhari, Nieuwdorp, & Fridenfalk, 2006). The use of a mood map with labels minimizes confusion between terms.

Due to the variety of possible moods and the potential variety in the number of participants, audio sets require a reasonably large number of files (approximately 100 or more) which cover a varied mix of characteristic combinations.[1] This will ensure that the system can select files which match participants' moods while providing varied characteristics within the mix of audio in the room. Livingstone and Brown have done extensive work mapping musical characteristics to the comprehensive list of musical emotion descriptors defined by Schubert from Henver's foundational checklist (Hevner, 1936; S. R. Livingstone & Thompson, 2009; Steven R Livingstone, A. R. Brown, & Muhlberger, 2005; Steven R. Livingstone, Mühlberger, A. R. Brown, & Loch, 2007; Schubert, 2003). This mapping provided a framework for developing the audio sets used in Eavesdropping to ensure a broad array of possible audio moods based on a variety of audio characteristics.

Some additional constraints have guided the creation of musical material for this system. The musician has no control over the timing of playback and therefore there is no way to specifically align the beats of several simultaneously playing audio files. Additionally, the lack of pitch information in the representational data creates an environment where files in various keys could be combined. Given these constraints, audio sets which resulted in the most aesthetically enjoyable acoustic ecologies were

[1] The six characteristics define a space resulting in $100^4 * 10^2$ possible combinations for the formal and mood values.

those without strong beats that aligned to a narrow complimentary set of keys, or those that explored a 12-tone range and appealed to the generative possibilities of the system.

3.2.3 Mood Composition

Version 1 of the system requires a musician to create a mood composition which guides the selection of audio when participants join a performance. The composition environment provides a web-based piano roll-style interface for arrangement (see Figure 3-2). The piano roll segments are not associated with the audio files themselves but represent compositional elements encoded with characteristics similar to the audio files, *Plurality* (Density), *Harmelodic*, *Tempo*, *Timbre*, *OuterMood* (Arousal), and *InnerMood* (Valence) (note that Version 1 used an earlier set of terms to describe the audio characteristics). The web-based interface presents a timeline in which composers can layer several compositional segments to shape the composition. For instance, the composer can indicate a long harmonic section to be played with several shorter melodic segments of alternating moods.

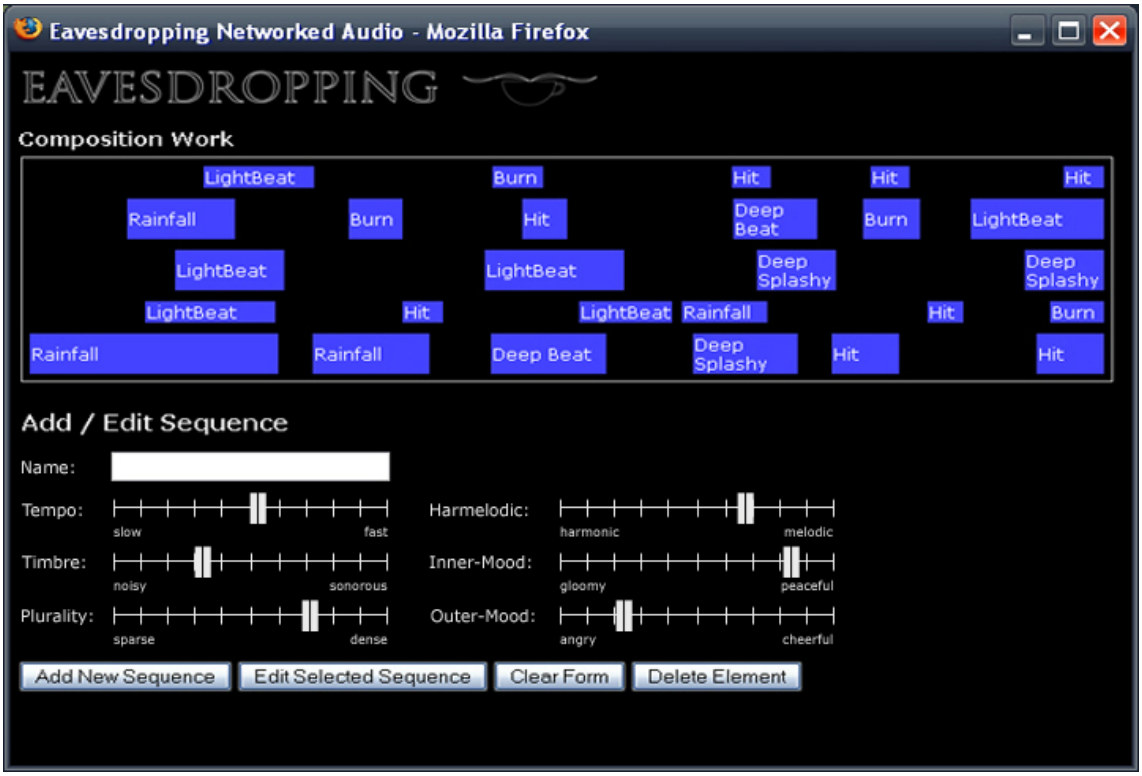


Figure 3-2: Mood Composition Interface

3.3 Performance Interface

The performance interface is designed to provide a minimalist interface with only the most basic information to participants. The information provided is meant to reassure participants that they are involved in the correct performance as well as to minimize any anxiety from uncertainty by always showing the time remaining in the performance. This interface also shows an audio control meant to confirm that the audio track is playing successfully in the event that the audio is periodically too quiet to hear. The Version 1 performance interface is pictured in Figure 3-3.

Figure 3-3: Version 1 Performance Interface

The Version 2 performance interface includes the same information that is present in the Version 1 interface but additionally includes support for the mood selection and reinforcement system explained in more detail in Sections 3.3.2 and 3.3.3.

3.3.1 Performance Initiation

When a participant first arrives at the Eavesdropping website they first are shown options to initiate a new performance or join an existing performance. If the participant opts to initiate a new performance, the system requires selection of an audio set (determining the audio files which will be used for the performance), selection of a

composition (Version 1) or indication of a duration intended for the performance in minutes (Version 2), and the entry of some basic information about the location for the performance (venue, city, province, country). This basic information will show up in a list for subsequent participants who wish to join an existing performance and will allow them to select the performance that is happening in their location. The initiator is then shown a simple, one-word password for their specific performance. This allows the initiator to announce the password to the local room to prevent people from other locations joining a performance and skewing the results. Subsequent participants wishing to join a performance merely have to select their location from the list and enter the password that was announced.

Once the performance is selected, Version 1 participants are presented directly with the performance screen pictured in Figure 3-3. Audio immediately begins playing from their computer. After each audio file is completed the browser requests a subsequent file from the server until the timeline of the composition has been completed.

When Version 2 participants select their location and join the performance they are presented with the mood identification interface described in the following Section 3.3.2.

3.3.2 Mood Identification

Version 2 performances require each participant to initially set their mood on the two-axis mood matrix (pictured in Figure 3-4) by moving a dot to the appropriate location on the grid. Once a participant has set their initial mood, the system will begin sending audio files to their browser based on that mood. While each audio file is playing

the system will continue to show the mood that was initially set but will prevent the participant from editing their mood. Preventing editing of moods is required for two reasons. First, the intent is that a participant will not become absorbed in the interaction at the expense of attention to the audio performance. Second, aesthetically, a participant will not get a sense of the general state of other participants in the room if everyone is adjusting their mood continuously.

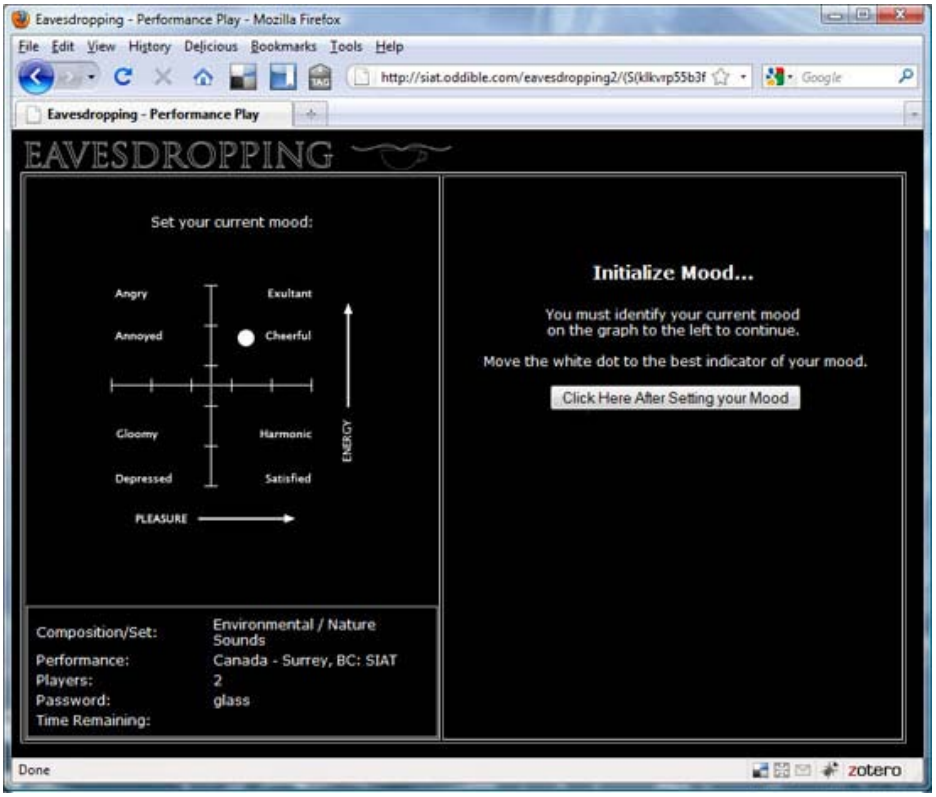


Figure 3-4: Version 2 Mood Identification Screen

3.3.3 Mood Reinforcement Interaction

When the Version 2 system begins playing audio to match a participant's indicated mood, the performance interface showing the mood matrix is additionally

accompanied by an reinforcement question asking the participant whether the audio matches the mood they indicated or not. A response can be given by clicking either of simple Yes / No buttons below the question (seen on the right side of Figure 3-5) This response is recorded by the reinforcement learning system to improve mapping of moods to audio files for subsequent selections (described in section 3.5). This interaction is optional. If a participant wishes to merely listen to the performance and observe the room, the performance will continue to play and request new files without any response.

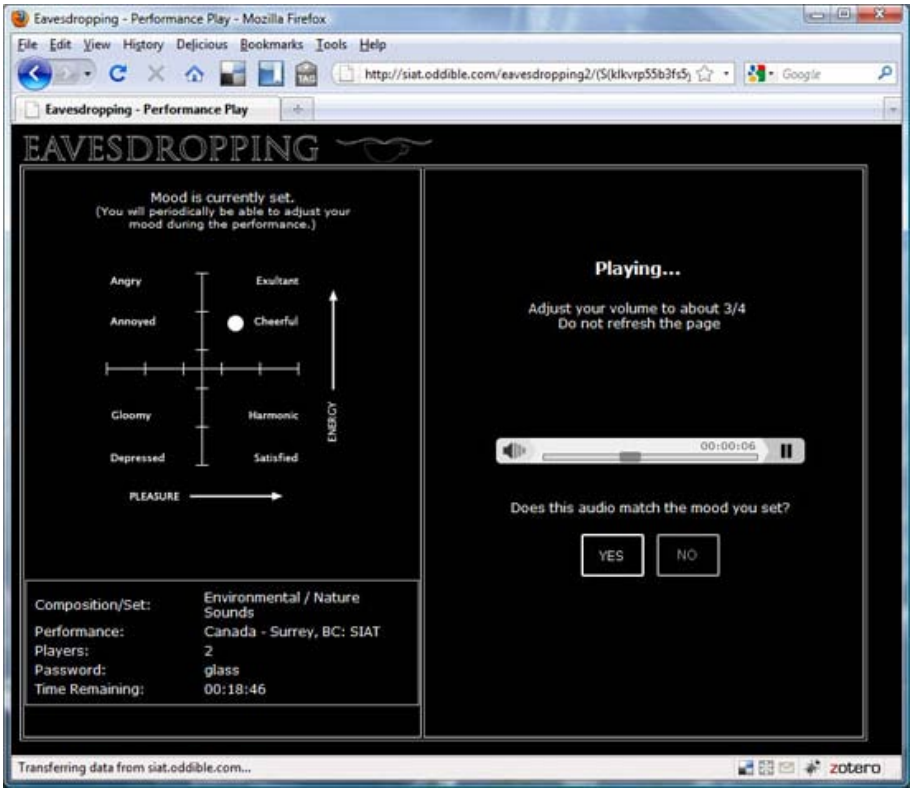


Figure 3-5: Version 2 Performance Interface

The initial mood input by participants serves to increase the engagement and immersion of the audience by giving them the sense that they are affecting the

performance and by providing a self-representation in the audio ecology. This directly addresses the *interaction* dimension of engagement and the *inclusive* dimension of immersion. Reinforcing the mood tagging of audio files further engages participants by encouraging evaluation of the mood representation thus offering them the ability to improve the mood mapping to create a more accurate self-representation and to imprint their legacy on future performances. This serves to address the *intrinsic interest* factor of engagement and both the *inclusive* and *extensive* dimensions of immersion.

During the performance a participant hears a combination of the audio being played from her own computer and from all other players' computers in the room. Each is aware that the audio being performed relates to the moods others have selected. When interesting audio catches the ear of a participant, they become aware of the person whose computer made the sound. It was intended that a shared recognition of the moods of others as augmented publicly through audio would increase *perceived affective understanding* and *perceived emotional interdependence* dimensions of social presence.

3.4 Version 1 Mood Composition Performance

The audio file selection systems in Version 1 and Version 2 are significantly different and are complex enough that they each require their own section. The following section details the audio file selection systems for
Version 1.

The Version 1 system is directed by a mood composition and attempts to fulfill the requirements of that composition by ensuring that all of the moods indicated at the current moment on the timeline are represented in the room. There is no direct

relationship between the moods in the composition and the moods of audio files in the set so it must determine the ideal file to play at any given moment. Version 1 relies on the accuracy of the initial mood values for each file as input by the musician that created the audio set. The system additionally attempts to provide a varied mix of audio in the room, some melodic, some harmonic, dense if there are few participants, sparse if there are many.

3.4.1 Version 1 Conductor

At the center of the audio selection engine for Version 1 is a decision module called a 'Conductor.' Conductors are plug-ins to the system and each has its own variations on how it analyzes the participant information and the composition to determine which samples to play. The use of Conductors allows experimentation with the decision-making logic without re-coding of the server system. When a participant's browser requests a file, the conductor uses its specific algorithm to compare the files which are currently playing with the set of compositional elements assigned by the composer for that specific moment. Only one conductor is active during a performance and the composer can select which conductor works best with their composition and audio.

In the following description, one simple and one complex conductor are discussed. Each conductor has two basic functions: 1) selecting a compositional element to play and, 2) selecting an audio file which matches that compositional element. First the conductor must evaluate the audio files that are currently being played by participant's browsers to the set of elements that has been defined in the composition at that location on the timeline. When a conductor selects a compositional element from the

timeline it increments the play count for that element so it knows which have been selected in prior operations. The simple conductor merely selects the oldest element with the lowest play count. The complex conductor evaluates the individual characteristics to select an element. For instance, it evaluates for the current density of the composition by adding up the *Plurality* values of all the compositional elements at the current location in the piano roll as a target density. It then selects an element which when added to the *Plurality* of all the audio files currently playing in the room is closest to this target density. In choosing a compositional element, the complex conductor weighs certain characteristics over others by sequentially evaluating characteristics and narrowing a candidate set of potential elements in each operation.

Once the conductor has determined a set of characteristics which will keep the composition on track, it then searches the audio file database for a file which most closely matches these characteristics. This is the second primary function of the conductor module. The simple conductor performs a nearest neighbor search by assembling a 'select value' from the sum of the differences between characteristics of the target element and those of each of the audio files. The sample with the lowest 'select value', which represents the lowest deviation from the target element, is sent to the participant. The complex conductor selects a best match file for each characteristic then evaluates the percentage difference of all the characteristics for the selected files. The file with the lowest sum of percentage difference is chosen and sent to the participant.

3.5 Version 2 Machine Learning for Mood Representation

The Version 2 system is attempting to match the moods of individual participants to audio files in the audio set based on a reinforcement learning algorithm and user input

to continuously improve the reliability of its mood matching each time it delivers a file. It is also attempting to provide a varied mix of audio in the room from files which match the target mood profile. The following sections detail the audio file selection systems for Version 2.

The mood classification system for Version 2 functions similarly to a supervised learning model in which the machine is making choices to match an expected response and the participant indicates whether the response is accurate, thus guiding the system to converge on the correct answers. However, in traditional supervised learning the participant is considered an expert with the correct answer. In the case of matching audio to moods there are factors which undermine the participant's correctness. First, the system is designed to function for a variety of participants and different participants classify the mood of audio differently based on personal preferences. Second, mood classification is often relative to the current mood of the participant; individual participants may classify the same audio differently depending on their mood (Knobloch & Zillmann, 2002).

3.5.1 Q-Learning and Mood Values

The algorithm managing the mood representation data based on participant responses is a variation of the popular Q-learning reinforcement learning (S. Russell & Norvig, 1995; Sutton, 1998). In this case, a Q-value table is utilized to record responses as well as for choosing an audio file to represent a specific mood.

Each audio file is associated with a 5x5 data table to store the utility values that have been learned for that file for each possible mood. The two mood characteristics are

each rounded to discrete integers on a scale of 1 to 5 offering 25 possible moods in the table. When an audio file is first added to the system, this table is pre-seeded with very small random values (meaning that this file is nearly equally appropriate for any mood). The higher the Q-value at a specific mood-coordinate for a file, the more likely is that file to get selected to represent that mood. Values in the mood table can range from 0.0 to 1.0. Each time a mood is reinforced for a given file (j), the system also stores the frequency count (nj) to track how many reinforcements a particular file has received.

3.5.2 Mood Reinforcement

During the performance a participant's Yes or No response to the question of whether the audio matches their mood determines if a file will be positively or negatively reinforced to represent the mood the participant has indicated. In this case we simply update the Q-value (Q(j)(x, y)) for the current mood (x, y) and the current file (j) by adding the value of the learning rate (α) multiplied by the reward value (R) to its existing value as seen in Figure 3-6.

$$Q_{(j)}(x,y) = Q_{(j)}(x,y) + \alpha R$$

Figure 3-6: Mood Reinforcement Equation

At present the learning rate is set to a constant, 0.1, and the reward value has been set to 1 for positive reinforcements and -1 for negative reinforcements. Given that the range of values for each mood falls between 0.0 and 1.0, each file can reach its maximum value with ten consecutive positive reinforcements. The reinforcement maximum is

capped at 1.0 rather than rescaling values. Once the maximum has been reached, subsequent positive reinforcements merely increment the reinforcement count.

3.5.3 Exploitation and Exploration

Selection of files to represent a participant's mood utilizes a system that takes into account the fact that in a learning-based model, the best fit for a mood may not be the file that has the highest Q-value for that specific mood. There may be suitable files that have been less reinforced (and thus probably less used) that are worth exploring. In general, reinforcement learning faces the problem of balancing exploitation and exploration. This trade-off, classic in artificial intelligence and machine learning, is about choosing at any given point whether to exploit the file that has the highest degree of confidence to represent a specific mood (in this cases the highest Q-value) or exploring files for which the real Q-value is less known.

A pure exploitation strategy formula as presented in Figure 3-7 would select the file (j) which has the highest Q-value (Q) at the mood location specified (x, y), pondered by its "confidence" (the ratio between the number of reinforcements received for j(nj) and the total number of feedbacks available for the given audio set so far (N).

$$j \leftarrow \underset{j}{\operatorname{argmax}} \frac{n_j}{N} Q_j(x, y)$$

Figure 3-7: Pure Exploitation Equation

In order to balance exploitation and exploration, a Softmax selection policy is utilized to vary the selection probabilities as a graded function of estimated value. In our case, we utilized the Gibbs measure, or Boltzmann distribution, which is commonly used

in machine learning exploration. It chooses file (j) with the following probability (Figure 3-8):

$$\frac{e^{Q_j(x,y)/T}}{\sum_{x,y} e^{Q_j(x,y)/T}}$$

Figure 3-8: Softmax Selection Probability

The greedy action is still given the highest value but others are weighted according to their Q-values. T represents a positive parameter called temperature, with high values for T causing all actions to be nearly equally probable and low values for T causing a greater difference in selection probability for files whose values differ. For our purposes we set T to 0.4. With this value, exploration is significant to allow for fluctuations in participant response while still exploiting known values enough to ensure that users would hear audio appropriate to their mood selection. Because we wanted the system to remain adaptive to variations in participant response there is no mechanism to decrease T over time as is common.

Initial pilots have indicated that this Softmax action selection method will cause convergence of the learning system toward the true Q-values (under the assumption that these exist and are static) and thus an optimal mapping.

The system utilizes this selection strategy to assemble a candidate set of six audio files which are subsequently sent to the conductor.

3.5.4 Version 2 Conductor

Once a candidate set has been identified to match the participant's mood, a single file needs to be chosen based on its ability to mix aesthetically with the other audio files

that are playing in the room. This is determined by setting a limit to the combined sum of any individual characteristic of all the audio files playing in the room at the time. Each formal audio characteristic (Tempo, Timbre, Harmelodic, Density) is given a constant multiplier (M), determined heuristically through pilot evaluation, to limit how it scales with the number of people (P) involved in a performance (listed in the second column of Figure 5).

If there is only a single participant, the limit value equals the maximum for any characteristic, 100. As new participants join the performance, the limit value is adjusted based on the number of participants and the population multiplier via the following equation:

Limit = 100 * M * P

Figure 3-9: Characteristic Limit Equation

	M	Limit *P* = 2	Limit *P* = 3
Tempo	1	200	300
Timbre	.9	180	270
Harmelodic	.8	160	240
Density	.5	100	150

Table 3-1: Example of Calculated Limits

The Conductor sums the values of the each characteristic from all the audio files currently playing and subtracts this amount from the limit values for each characteristic. It then narrows the candidate set provided by the mood selection operation to those files which do not exceed that difference in any characteristic.

As an example, Table 3-1 shows the limits for each characteristic for 2 and 3 participants. So if there are two participants playing audio files with densities of 50 and 35, and a third participant joins the performance, the program would narrow the candidate set to only those audio files with densities less than 65.

Once the candidate set has been narrowed to those files which ensure that no limit has been exceeded, the Conductor randomly selects a single file from the remaining set. If no candidates exist that do not exceed the limit in any one characteristic or if the limit value has already been surpassed by the audio files currently playing, the Conductor then looks for a candidate that exceeds all three characteristics the least. The selected file is sent to the participant's browser for performance and evaluation.

4: METHODOLOGY

This research uses a mixed methods approach encompassing qualitative arts research-based strategies in the creative process and initial contextual evaluation, and a quantitative statistical comparison of engagement, immersion and social presence via self-report questionnaire to evaluate the two versions of the art installation and assess their ability to meet the original design criteria. Section 4.1 describes a study of cafés as a context for the performance of this project and the audio and social phenomenon investigates. Section 4.2 presents an investigation of user experience in performances of both versions of the project in several different environments. Section 4.3 details the construction and adoption of self-report quantitative questionnaires for the evaluation of engagement, immersion and social presence during interactive computer performance art and explains the design of user studies for a comparative analysis between the two versions of the system.

4.1 Contextual Investigation

During the conceptualization and design of this project it was necessary to perform an initial contextual investigation into the social acoustic ecology of the type of public environment for which this installation was designed. The investigation was conducted as a passive, participant observation, without a structured list of expected effects. This exploration was designed to observe audio phenomena in the environment, the relationship between participants and their audio space, the interactions that occurred

between people as a result of audio phenomena, and attention given to laptops in the environment.

A café environment was chosen for this investigation due to the high variability of audio phenomenon, the voluntary nature of peoples' presence in the space, and the very public nature of the space; that there would be a variety of dissociated people with the potential to interact. This study involved a single investigator, the artist, visiting several cafes over several months. There were two types of sessions, focused sessions and participant sessions. In focused sessions the investigator primarily attended the café to specifically monitor the environment. In these sessions the investigator ordered a drink and sat quietly at a table writing in a journal, observing and documenting phenomena. In participant sessions the investigator engaged in normal café activity (such as working on a laptop and reading a book) and only notated effects which explicitly drew attention or otherwise broke concentration.

All of the sounds of the café were notated with particular attention to those sounds which were made by people. Interactions between people at different tables were notated as well as whether these were initiated by sounds. The activities or assumed purpose for individuals' visit to the café were noted and how attentive they were to this purpose versus being distracted by external effects. Significant events were noted and how much attention they attracted.

The results of this investigation contributed to a general field of knowledge concerning the social acoustic ecology in public environments prior to the construction of the art installation and are discussed in 5.1 Contextual Observations.

4.2 Pilot Studies

Upon completion of each version of the art performance system a series of pilot studies were conducted to allow the artist to examine how the work performed in public contexts and to evaluate the outcomes and effect on audience members. Pilot studies were conducted in classrooms, in lecture halls, in a gallery, and in a café (specific locations and results are documented in 5.2 Art Performance Pilot Studies). Documentation in these studies was performed by a known participant observer, the artist. This was primarily a convenience issue due to the fact that the performance was introduced and initiated by the artist and was therefore known prior to the installation.

The pilot studies were observed to identify the effectiveness of the system to produce an audio ecology in the environment, participant behavior in response to audio phenomenon, and interactions between participants. Notes were taken on a laptop computer situated in the installation environment after the performance had completed.

Informal question and answer periods followed all of the pilot studies as two way group interviews between the artist and the participants. Since this was a group phenomenon being explored, group interviews offered some benefit to the interview process. First, participants could hear each other's questions and answers and therefore acknowledge and contribute to the questions and answers given. Second, discussion could take place between participants which resulted in a richer interview which could wander into areas beyond the interviewer's conceptualizations. This did however require some direction from the interviewer to prevent discussions from getting too far afield. Questions and answers as well as any follow-up comments were documented on a laptop computer.

4.3 Quantitative Analysis

It is becoming more common that digital artists and designers evaluate their technological systems employing techniques from the domains of Human Computer Interaction (HCI) (Morrison, Mitchell, & Brereton, 2007). The research question in this study explores the changes made between the two versions of this art system and a comparison of impact these changes had on social presence. This study employs a cross-sectional research design which involves quantifiable data collection from groups of subjects who have participated in the art performance to identify variance between the two versions of the art system. The differences between the two versions of the system were classified into two areas, self-representation as a factor of immersion and interaction as a factor of engagement. The model defined in Section 2.4 additionally identified that Engagement contributes to Immersion and Immersion contributes to Presence and thus Social Presence. These factors were evaluated individually as well as in relation to each other and for their effect on Social Presence.

Construction of questionnaires to assess the level of Immersion and Engagement of participants and the Social Presence between participants began with an extensive review of existing instruments in the fields of telecommunications, virtual environments, augmented reality, video games, computer supported collaborative work (CSCW), HCI, and interactive arts. This review looked at the context in which these instruments were delivered, factors used by each instrument, and the terms used to construct questions regarding those factors. A selection of this review can be found in Appendix B: Questionnaire Evaluation.

4.3.1 Social Presence Questionnaire

A significant number of recent studies on social presence have derived their instruments from Biocca and Harms Networked Minds Measure of Social Presence (NMMoSP) (F. Biocca et al., 2001; Frank Biocca et al., 2003; C. Harms & F. Biocca, 2004). This instrument was primarily devised to measure social presence between two individuals involved in completion of a task together and showed a successful ability to differentiate social presence between pairs of subjects who were completing the task in face-to-face versus mediated environments. Of the six factors representing social presence in this instrument, four identify dimensions specifically relevant to the social presence environment created with the Eavesdropping performance. Those factors are *Co-presence, Perceived Affective Understanding, Perceived Emotional Interdependence, and Perceived Message Understanding*. The two remaining factors *Attentional Allocation* and *Perceived Behavioral Interdependence* were more relevant to one-on-one and task oriented interactions but still offer some insight into the social relationship between subjects in the study, where subjects fixate their attention on individuals and where behavior is a reaction to others or seen as initiating the action of others.

The complete 6-factor NMMoSP questionnaire was adopted for use in measuring social presence between the different versions of the Eavesdropping system. Minimal wording changes were required to adapt the questionnaire to this context, amounting only to changes in number in the pronouns used (for instance, the NMMoSP question “I could tell how [my partner] felt” was changed to “I could tell how [others] felt”). This NMMoSP questionnaire involved 36 paired questions in 6 factors answered on a 7-point

Likert scale. The full NMMoSP used in the Eavesdropping study can be found in Appendix A: Questionnaire.

4.3.2 Immersion / Engagement Questionnaire

From the review of engagement and immersion questionnaires (a sample of which can be found in Appendix B: Questionnaire Evaluation), an initial set of 33 factors of engagement and 20 factors of immersion were assessed to construct a model for the Eavesdropping study (see Figure 2-1 in section 2.4.2). The 33 factors of engagement were reduced to three primary factors: *Intrinsic Interest*, *Interaction*, and *Control*. From the set of terms and questions related to these factors, a set of 14 paired questions was derived from the questionnaire evaluation through assessment of face validity and content validity to determine how well they captured the three dimensions of engagement as explained in section 2.4.1.2.

An adaptation of Slater and Wilbur's factors were used to define four dimensions of immersion: *inclusive, extensive, surrounding, and vivid* (Slater & Wilbur, 1997), see Section 2.4.1.1 for details. From the set of terms and questions in the 20 initial factors of immersion, 12 paired questions were derived through assessment of face validity and content validity to capture these four dimensions.

The complete questionnaire can be found in Appendix A: Questionnaire.

4.3.3 Pre-testing, Assessment, and Refinement

The full questionnaire, including the adapted NMMoSP and the 26 question Immersion / Engagement questionnaire, was presented for review to two additional

academics in related fields. Some clarifications of wording were made based on feedback obtained in these sessions.

This survey was then constructed into a single-page web form and pre-tested by a group of five participants who were observed responding to the survey. Subjects were first presented with a performance of the Version 1 system. They were then presented with the questionnaire and asked to verbally raise any issues while they read through and answered the survey, their issues were noted. Comments pertaining to the visual design, functionality, navigability, and verbal clarity of the survey were used to improve the wording in the survey and its delivery system. A second pretest was conducted after the first round of modifications by an additional group of five different participants who were presented with the Version 2 Eavesdropping performance prior to questionnaire evaluation. After the second session the questionnaire was reviewed question-by-question and participants as a group were asked to explain what they thought each question meant. Further clarification was made to the wording of the questions as well as the survey design. The participants in these pretests were a convenience sample of graduate students from two different labs at the School of Interactive Arts and Technology of Simon Fraser University, and were not associated with this research. Changes made to the survey included minor formatting, spelling and wording issues, as well as screen size format issues. The final version of this survey design was used in the Round 1 study.

4.3.4 System Data

In addition to the self-report questionnaires, the Eavesdropping system also collected a variety of data from participant users. It stored each participant by

anonymous ID and recorded the choices that they made as they interacted with the performance.

Version 1 of the system offered no interaction to participants and therefore only logged the date and time of each performance, how long it lasted, the audio set chosen, and how many participants joined the performance throughout its duration as well as the participant ID and the file ID of each audio file performed at each laptop.

Version 2 of the system also logged the date and time of each performance, how long it lasted, the audio set chosen and how many participants joined the performance. Additionally, it logged the participant ID, each participant's initial mood setting, the file ID of each audio file each participant received, and any reinforcement answer or non-answer that a participant may have given regarding whether each audio file matched the chosen mood.

4.3.5 User Study

Two rounds of controlled user studies were conducted for the purpose of evaluating social presence and the effectiveness of changes made to the system between Version 1 and Version 2. These studies were conducted with isolated groups, primarily of five participants, and were followed up with a questionnaire. In each round both versions of the software were tested. Groups were randomly assigned to Version 1 or Version 2.

Subjects were solicited via announcements at the end of undergraduate classes at Simon Fraser University during the summer of 2010. The announcement script can be found in Appendix C: Announcement. Groups were self-assembled and a demographic

question inquired the extent to which they knew the other participants. Incentive for participation in the first round consisted of a lottery for two FutureShop gift certificates (value $50 each) to be drawn from amongst the pool of everyone who participated in the first round. Second round participants each received a $10 Blenz Coffee gift card for their participation. The system was set up in classrooms outside the class that was solicited.

The physical arrangement for both rounds was identical, consisting of 5 laptops around a set of tables, two on one side, two on the other, and one at the end (see Figure 4-1: Participant arrangement). All laptops were identical, with 15 inch screens, a touchpad mouse, and had their audio volume turned up to the maximum prior to the performance.

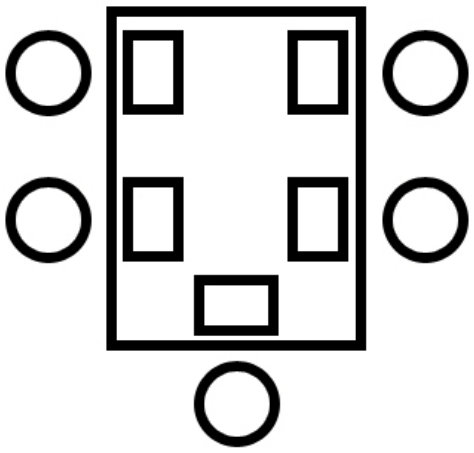


Figure 4-1: Participant arrangement

At the beginning of each session the laptops were all set to show the title screen of the performance system in a full screen browser window. Users were read the introduction script which differed slightly based on whether they were being presented

the Version 1 (v1) system or the Version 2 (v2) system. Introduction scripts can be found in Appendix D: Introduction Script.

At the end of the introduction script participants were instructed to click the link to begin, and to proceed through the tutorial and informed consent authorization until they arrived at the page that told them to 'Wait Here'. The tutorial showed the screens they would encounter to initiate the performance as well as screens that they would see during the performance, with text and arrows highlighting any interaction that would be required or optionally available. The tutorials can be viewed in Appendix E: Tutorial Pages and the informed consent authorization page can be viewed in Appendix F: Informed Consent.

When all participants reached the 'Wait Here' page, the moderator initiated the performance via a sixth laptop but did not join the performance. The Environmental / Nature audio set was chosen for all Version 1 and Version 2 performances and consisted of sound effects between 12 and 17 seconds long primarily of bird and animal noises as well as some weather effects that had been coded into the system by the artist.

Participants were then instructed to click the link to proceed and were taken to the *Join Performance* screen. On this screen the moderator indicated to participants to click the performance listed and to type in the password (given verbally), then to click the Join Performance button.

For v1 participants, the performance started immediately at this point and each computer began playing audio. For v2 participants, they first arrived at the mood selection screen. The moderator prompted them to click and drag the mood indicator to the location on the mood map that represented their current mood and to click the

Continue button. Once they clicked continue, the v2 people began hearing audio from their computers and were offered the reinforcement question on their screens.

During the performance the moderator either waited out of the room until the performance completed, or at a distant table and avoided visually monitoring the participants. When the performance had completed the system presented participants with a link to the questionnaire and all began taking the questionnaire without prompting. Subjects took between 7 and 15 minutes to complete the questionnaire.

The questionnaire was administered via browser-based web pages presented on the same computers as the performance and from the same server as the performance system. It was developed as a custom application in ASP.NET with data stored in SQL Server database. The formatting of the questions on the web page was modified between rounds. Discussion of the reasoning for these changes can be found in the following section, 4.3.6.

Round 1 participants received a questionnaire where all questions were on a single web page and could be scrolled through from top to bottom with the page's vertical scroll bar operating as a progress bar. Questions were divided into groups of 10 to give subjects a sense of progress and were separated into two main sections. The first and most substantial section of the survey consisted of questions from the NMMoSP and the Immersion / Engagement questionnaire (described in section 4.3). Context effects in this section were negated by randomizing items; everyone who completed the survey viewed the items in a different order. The main section was followed with a section of demographic questions. The concluding page thanked respondents for their time and offered a web address where analysis and results would eventually be posted.

Round 2 participants received a questionnaire which was broken down into multiple pages of 10 questions each with a progress bar visible horizontally across the bottom of the screen. The questionnaire was broken down into four main sections. First some initial demographic questions were presented, followed by randomized questions from the NMMoSP, then randomized questions from the Immersion / Engagement questionnaire, and finally some additional demographic questions regarding the participant's experience with similar art works. The concluding page thanked respondents for their time and offered a web address where analysis and results would eventually be posted. The formatting for this questionnaire can be seen in Figure 4-2.

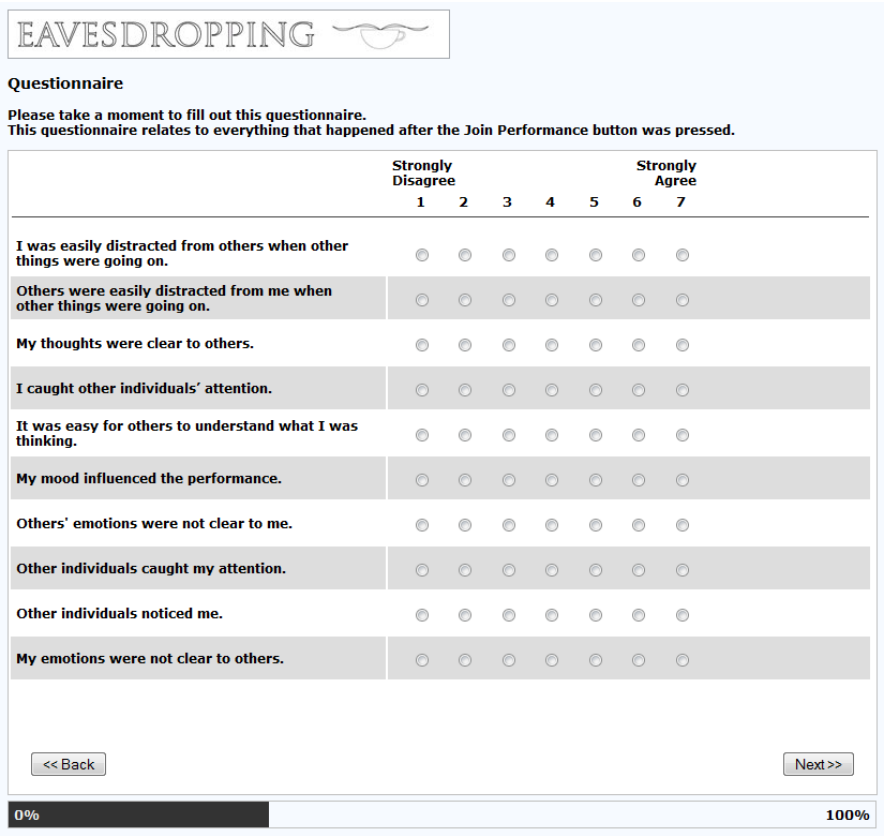


Figure 4-2: Questionnaire Formatting

Following the questionnaire participants were thanked verbally for their participation, given their compensation or entered their name into the drawing, and the session was adjourned.

4.3.6 Interim Results and Questionnaire Modification

During the first round of studies the data was examined to assess if any issues arose during testing. Some irregularities were present. First, it was noted that an unusually high number of university undergraduates had indicated that they had completed a 4-year college degree in the education demographic question. The demographic questions in Round 1 were randomized at the end of the survey which might suggest that subjects were fatigued at this point and were not answering questions accurately.

Similarly, it was noted that a high number of subjects were indicating an above average familiarity with fine arts, also a late questionnaire demographic question. Upon noticing this I began asking some informal exit questions from subjects. In one 5-subject test I asked each of the 5 what they answered for that question. All of them had answered higher than 3 on a 5-point Likert scale. I then asked how many of them had ever been to the Vancouver Art Gallery. All answered that they had never been. I then asked if they had ever been to any art gallery. Two responded that they had walked through the university gallery once. The others had not ever been to any art gallery.

I then examined paired questions in the questionnaire, both by comparing means and on a subject-by-subject basis. Many errors existed where subjects gave contradicting answers to the same paired question. This was not specific to a small number of subjects,

many subjects produced contradicting answers to several questions. Errors seemed to be widespread across the questionnaire, however particular attention was paid to issues in the Engagement / Immersion questionnaire since that was a new instrument constructed for this project. Additionally, many subjects who had participated in the Version 1 performance without interaction answered in a way to suggest they had a strong interaction with the performance system.

The broad incidence of issues indicated that there may have been problems with the study design. This was addressed by modifications to three areas of the study. First, the introduction script was refined to ensure that participants were aware that the questionnaire was specific to the performance itself and was not asking questions related to their interaction with the system while they were reading the tutorial or joining the performance. Second, the questionnaire was visually and functionally redesigned to reduce fatigue. The color scheme was changed from white text on a black background to black text on a white background, the format was changed to ten questions per page, and a clear progress bar was developed along the bottom of each page. Lastly, the paired questions in the Engagement / Immersion questionnaire were rewritten as single questions, each with their own Likert range related specifically to the question. To reduce ambiguity for subjects who were not given an interactive performance, questions which mentioned interaction were rewritten to a more general wording that did not imply that there was an interactive element.

A second round of the study was run with the new design. In order to evaluate the impact of the changes made between rounds, the new factor, Rounds, was added to

statistical evaluation as a blocking variable. Paired questions from Round 1 were combined as a mean for comparison with the new questions in Round 2.

4.3.7 Statistical Methods

The total number of people to participate in both rounds of the study was 101, however since each session involved 5 people, the 5-person group will be considered the unit for this study. Therefore, the N for this study was 20 units, 11 for Round 1, 5 in v1 and 6 in v2, and 9 for Round 2, 4 in v1 and 5 in v2. Social presence is determined by the interaction between the 5 specific people in attendance at a performance rather than by any one individual, the treatment was applied to a 5-person unit. Variability of individuals certainly contribute to social presence effects but this also affects the group and is not a particularly relevant unit to measure for the purposes of this study.

The study design involved one primary treatment factor, Version. Rounds were included as a blocking effect to assess variability between session dates and changes in protocol between the rounds.

Since in our model Engagement contributes to Immersion and Immersion contributes to Social Presence, effects were analyzed in that order. First a 2-factor ANOVAs was run against Engagement, with Rounds and Versions as factors. Subsequently a 2-factor ANCOVA was run, this time against Immersion, with Version and Rounds as factors and Engagement as a co-variate. A final 2-factor ANCOVA was run against Social Presence with Version and Rounds as factors and both Engagement and Immersion as co-variates. Using ANCOVA removes the effects of the co-variates

and lets us evaluate whether any effect we are experiencing comes from the main factor or other factors in the model.

Similar tests were then run with the individual dimensions of Engagement, Immersion and Social Presence to look for more precise effects.

5: ANALYSIS AND FINDINGS

This section describes the details and outcomes of the three exploratory evaluation stages run in this project. First, in Section 5.1 we evaluate the contextual studies which established the framework and basis for the art work as was explored through participant observation in cafes and discuss the findings that arose from these observations. Second we explore the audience experience, comments and questions that arose during the pilot and field studies, the findings from Version 1 pilots which directed design decisions for Version 2, and findings from Version 2 pilots which evaluated the success of those designs. Finally we analyze the quantitative data captured during the social presence user studies and compare the results from the two systems.

5.1 Contextual Observations

During the contextual participant observations the study focused on three general areas, the sounds in the environment, the interactions between people and with the sounds, and the purpose for people's presence in the space.

In general the sounds in the space filled the soundscape spectrum including continuous sounds, a general din which established the keynote sounds of the space (including spatial acoustic resonance and reflection), and a variety of sound signals which broke through the background sound and attracted conscious attention to them (Truax, 2001). The range of what could be called keynote sounds, sounds particular to an environment which establish the culture of a space, varied according to the place. All

cafes had some sort of music playing which primarily remained as background sound but periodically attracted its own attention. Some cafes were on busier streets and had the sound of traffic going by. All of the cafes had the sounds of the baristas performing their coffee and tea orchestra behind the counter. The most prominent of these sounds were the steam from the espresso machine and the coffee grinder, which, while quite loud did not often attract any attention. The same can be said for the general counter work of washing dishes, clinking silverware, etc. The verbal conversation at the counter offered significant variety which attracted the most attention of any of the counter activities. Additionally, if there was any line-up at the counter, the people waiting in line were in almost continuous distraction by events and sounds taking place around them. Conversely, these people standing, often in the middle of the room, were a frequent source of distraction to folks sitting down at the café as well.

The sounds in the customer area of the café can be broken down into two types: on one hand people created sounds because they were doing things that made noise such as talking to someone, stirring coffee, moving a chair, typing on their laptop keyboard, setting down a ceramic coffee cup on a wooden table, etc. I will identify these as action sounds. On the other hand, people made what I will call fidget sounds that had no apparent functional purpose, these sounds were primarily the result of fidgeting, clicking a pen, scratching at a book edge, tapping a foot. Human-made sounds seemed to cover a broad range of how consciously aware people were that they were making the sounds and how aware they were that other people could hear the noises they were making. Fidget sounds seemed to almost always be unconscious and typically happened when people were reading, staring at their computer monitors or staring out the window.

Action sounds were much more complex. There were common actions which were regularly performed as part of the range of activities one might be involved in at a café. Different people performed these actions with a broad range of volume levels. Some people set their cups down loudly, others quietly. Some people laughed loudly, others chuckled in more subdued ways. Some action sounds were clearly behavioral responses to the actions of others. People used sounds to attract another's attention, they used loud sharp sound signals to interrupt or quiet the loud continuous sounds of others. Still other sounds seemed to contribute to a serendipitous acoustic ecology of the space. Some sounds seemed to arise to break the silence, to accompany others' sounds, or to contribute to the general atmosphere once the silence was broken. There was some serendipity to the sound making in the space where several sounds seemed to arise as if someone had opened the door to allow others to make sounds of their own.

People seemed to attend the café for a wide variety of reasons and activities. Many came to meet and chat with others, some came alone and met others at the café, still others arrived as groups. Many people attended the café to work either on paper or on laptops. Some of the laptop folks seemed engaged in less work-like activities including web browsing and even video games. A significant number of café goers attended the café to read. Lastly, others attended the café merely to sip a drink and observe the scene or to think to themselves.

The purpose for attending the café seemed to have some impact on the level of attentiveness to events and sounds happening in the café. Those working or reading or involved in conversations were less attentive than those just sitting at tables enjoying their coffee. Likewise, there were café goers who isolated themselves acoustically via

headphones with their own personal soundtrack. However, everyone in the café participated in some people watching. Some took breaks from what they were doing to look around. Others would glance up periodically then go directly back to their work. Sometimes these breaks were instigated by some action or sound event which happened in the space, other times they seemed to coincide with breaks in conversation or breaks in the activities in which they were involved.

Discussion of the implications of these observations and how they set the framework for this project are further detailed in Chapter 6: Discussion.

5.2 Art Performance Pilot Studies

Pilot studies of the Eavesdropping Interactive Audio system were performed in three different contexts: demonstration of the system in the classroom, presentation and performance of the system in a conference setting, and performance and exhibition of the system in cafes.

5.2.1 Version One

The version one mood composition and performance system was exhibited in a classroom, in a conference and in a cafe.

The classroom session was attended by ten participants, each with their own laptops arranged in a half-square at tables in a small room with no participant more than 12 feet from another. All participants were facing the center of the room and could easily see all other participants without turning more than 90 degrees to the side. The system was described in detail during a 20 minute presentation prior to beginning the performance. The performance was started and lasted for 10 minutes. After the

performance participants were encouraged to comment on the performance. During the comment session, there was some question and answer between the moderator and the participants to clarify the comments and some of the actions that happened during the performance.

The conference session was attended by approximately 50 audience members, of which approximately 30 participated in the performance with their own laptops. The room was a medium-sized lecture hall with seats arranged in several rows in a semi-circle around the front screen. The system was again described in detail during a 20 minute presentation prior to the performance. The performance lasted 10 minutes and afterward a similar discussion and Q&A session ensued.

Figure 5-1: Cafe performance

The cafe performance was attended by approximately 10 audience members, 5 of which had their own laptops and participated in the performance. The room was approximately 20 feet square and all participants sat at cafe tables on one side of the room (see Figure 5-1). The performance lasted 10 minutes and was video taped but did not have a post-session discussion.

While each performance was underway, participants primarily listened to the audio, looked around the room, and laughed when unusual sounds occurred while looking toward the direction of the audio and the person sitting at that laptop. Several participants made funny faces to correspond to funny noises that their computers made, in a performative reaction intended seemingly to acknowledge yet distance themselves from the sound – in many cases their heads did even move away from the computer in making these faces. Several participants leaned their heads into their laptops to more clearly hear the sounds their computer was making over the audio from other computers. Several participants also turned their heads to better direct their ears at other participants' computers that were making interesting sounds.

Participants in the cafe session were the most conservative in their participation and allowed the performance to complete without disruption. This was likely due to the very public nature of the setting. Participants in the conference setting were somewhat difficult to monitor individually as there were so many of them, however, most seemed to stay focused on the performance and other participants, and didn't engage in significant disruptive behaviour. This was likely due to the fact that with so many participants there was a significant number of different sounds in the room creating a diverse and interesting acoustic ecology to pay attention to. The classroom participants were a bit

more explorative in their interaction with the performance system and each other. Observations on some of the more unusual interactions that happened during the performances follow.

Two of the participants in the classroom session turned their laptops around to face the group. One participant in the conference session turned his laptop around and raised it over his head. When asked during the Q&A why they did this, they responded that they had trouble hearing some of the sounds from computers that were far away from them so they wanted to make sure their sounds were heard by others. One participant in the classroom session opened Windows Media Player and began playing her own music during the performance. She was quietly asked to turn it off shortly thereafter. When asked during the Q&A session why she had done it, she responded that she thought we were all contributing audio to the room and that she felt the audio from the system wasn't that enjoyable and wasn't doing very much. Lastly, one participant in the classroom session opened multiple browser windows of the same performance so that his laptop was playing several audio samples. He mentioned this aloud which caused several other participants to also open several browser windows. They found that the system would play several samples at the same time but that only the top most browser window would refresh and play subsequent samples. They began refreshing all of their windows manually so that their systems continued to play several audio samples. When asked in the Q&A why everyone did this, they mentioned that they were just exploring – looking for interactive possibilities in the interface.

During the post-performance discussion session there were several proposals as to why the participants behaved in the ways that they did. One issue was that the

performance co-opted individuals' laptops for the performance. Participants sensed a loss of control of their machines and engaged in behaviour to manipulate their machines in the context of the performance as a means to regain a sense of control. A second issue was that participants did not feel connected to the audio coming from their machines yet other participants were associating them to the audio by looking at them when an unusual noise occurred at their machine. This made people feel uncomfortable and they began behavior that distanced themselves from this discomfort or to create a scene which they felt better represented them.

Resolutions were proposed and discussed to minimize the issues that arose. One resolution to the control and disconnectedness issue was to allow participants to input their own mood into the system. They would feel more invested in the audio coming from their own computer and would be less inclined to create their own audio and would be more interested in the audio that was presented. Another proposal was to give participants some sort of interaction to engage in during the performance that would make them feel like they had more control. With some added meaningful interaction, participants would be more engaged and less likely to explore external interactions or a gaming of the system.

Overall the first version satisfied the goals of this art system to create a balanced mix of audio in the room based on the number of participants and to achieve some social interaction and awareness between individuals through the proximity of participants to sounds. The café, the originally intended location for this system, was a loud environment with people coming and going and talking and ordering coffee, and cash registers and espresso machines beeping and hissing. Laptop speakers in this

environment turned out to be rather quiet, which made the performance a subtle and intimate audio experience where localized interactions only occurred between participants who were closest in proximity.

5.2.2 Version Two

The version two interactive mood ecology system was exhibited in a classroom, in a conference and in a gallery.

The classroom session was attended by six participants, each with their own laptops arranged in two rows at tables in a small room with no participant more than 8 feet from another. All participants were facing the front of the room. The system was described in detail during a 20 minute presentation prior to beginning the performance. The performance was started and lasted for 10 minutes. After the performance participants were encouraged to comment on the performance with similar Q&A as in the Version 1 sessions.

The conference session was in a large black box space with approximately 120 attendees, of which approximately 20 participated on their own laptops. This environment was incredibly loud due to the large number of audience members shuffling in their seats and having conversations on the periphery, the buzz and hiss of fans and speakers from the other audio and video equipment in the room, and questions and discussion that began while the performance was under way. Some discussion and Q&A happened during and after the performance but this was clearly not an appropriate environment for this system and results were left out of the analysis.

The gallery session was a unique situation where 5 laptops were arranged on a single table with chairs approximately 2 feet apart from each other, all facing the center of the table. 12 groups of 5 people participated in the performance with several spectators looking and listening over their shoulders. As each new group sat down at the table, the art installation was briefly explained and then started for a 5 minute performance. After each performance a short discussion ensued with some Q&A for clarification of comments. This environment can be seen pictured in Figure 5-2.

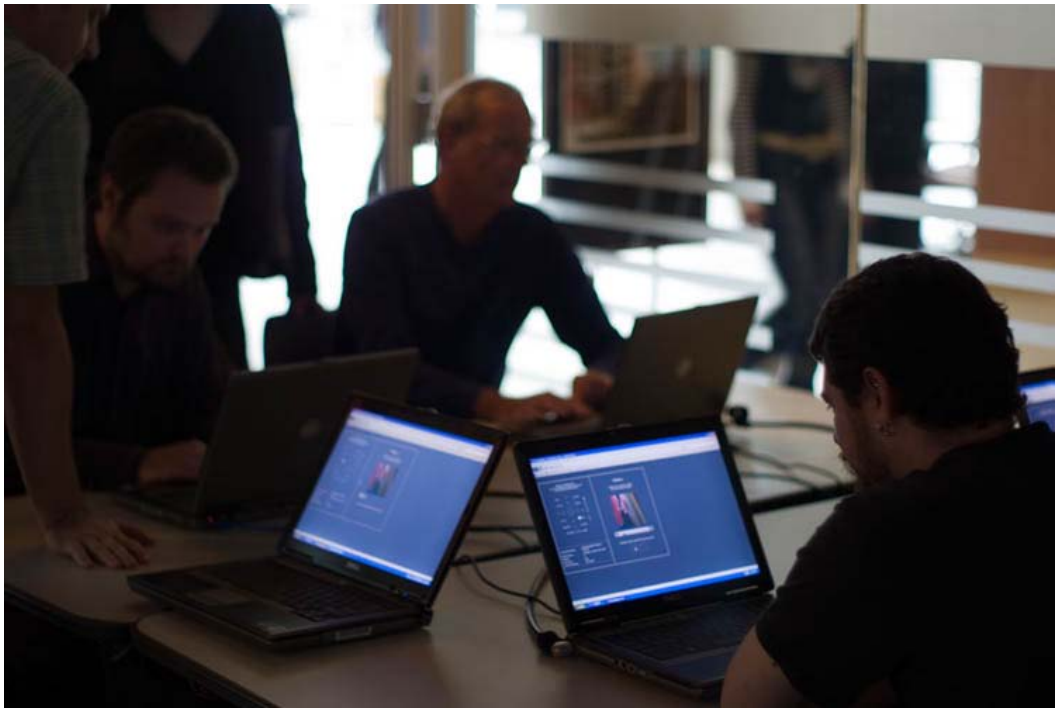


Figure 5-2: Gallery performance

Participants to these Version 2 performances were significantly more engaged by the interface than in Version 1. There was minimal behaviour outside of the expected set of actions – no one re-oriented their computers, no one played other audio, no one opened multiple browser windows. All interaction was strictly with the interface in front of

them. There was much more attention paid to the computer screen than to other participants as well. During observation of participants' interaction with the system, it was clear that many participants were merely waiting for the next audio file to start so that they could input their vote regarding the audio-to-mood matching. Several didn't even look up from their computer monitor but watched the audio player for the entire performance so that they could input their vote every 15 seconds when the next audio sample ended.

After the performances there were two recurring suggestions which arose. The first related to the setting of the mood on the mood matrix. The Version 2 system only allowed participants to set their mood at the beginning of the performance, not during the performance. Several participants indicated wanting to be able to change their mood while the performance was under way. The second significant criticism related to the task-oriented nature of the reinforcement learning question. Several participants indicated that they felt obligated to vote on every audio file. This correlated to observations made during the performances of a reduction in awareness between individuals caused by more interaction at the interface.

5.3 Social Presence Data

This section presents an analysis of questionnaire data to assess the impact of changes made between versions of the system on engagement, immersion and social presence.

The first round of user studies (R1) involved 56 subjects participating in 11 performances consisting of 10 groups of 5, and 1 group of 6 participants. 5 groups were

given the Version 1 (v1) performance, 6 groups were given the Version 2 (v2) performance. The second round of studies (R2) involved 45 subjects participating in 9 performances consisting of 9 groups of 5 participants. 4 groups were given the v1 performance and 5 groups were given the v2 performance. The following table (Table 5-1) shows the arrangements of groups.

Subject Grouping			N
Round	1	Version 1	5
		Version 2	6
	2	Version 1	4
		Version 2	5

Table 5-1: Subject Grouping

First a test of skewness and kurtosis was run to examine the normal distribution of the data. Skewness results for v1 and R1 showed that the overall data was within an acceptable range for engagement, immersion and social presence. However, significant negative skewness was present for immersion in v2 and for social presence in both v2 and R2. This indicates that scores were bunched up at the high end of the scale. Kurtosis appeared normal for engagement but was leptokurtic for immersion for R1 and v2 and for social presence in R2 and v2. This indicates a sharp peak in values. These results are visible in Table 5-2 below.

Engagement * Round

Engagement

Round	Mean	Median	N	Std. Deviation	Kurtosis	Skewness
1	4.0534	4.0750	11	.20295	.622	-.815
2	4.2435	4.4333	9	.42772	-.033	-.996
Total	4.1390	4.1125	20	.32882	-.182	-.303

Engagement * Version

Engagement

Version	Mean	Median	N	Std. Deviation	Kurtosis	Skewness
1	4.3139	4.3000	9	.26448	-.393	-.057
2	3.9958	4.0250	11	.31529	-.471	-.235
Total	4.1390	4.1125	20	.32882	-.182	-.303

Immersion * Round

Immersion

Round	Mean	Median	N	Std. Deviation	Kurtosis	Skewness
1	4.0230	4.0750	11	.29627	2.566	-.885
2	4.3167	4.4500	9	.83703	.652	-.736
Total	4.1552	4.1125	20	.60304	1.485	-.132

Immersion * Version

Immersion

Version	Mean	Median	N	Std. Deviation	Kurtosis	Skewness
1	4.3174	4.1500	9	.64414	.388	.621
2	4.0224	4.0750	11	.56211	2.198	-1.305
Total	4.1552	4.1125	20	.60304	1.485	-.132

Social Presence * Round

Social Presence

Round	Mean	Median	N	Std. Deviation	Kurtosis	Skewness
1	4.0732	4.0667	11	.14423	-.217	-.076
2	4.1381	4.2417	9	.24348	4.823	-2.070
Total	4.1024	4.1260	20	.19237	2.314	-1.274

Social Presence * Version

Social Presence

Version	Mean	Median	N	Std. Deviation	Kurtosis	Skewness
1	4.1153	4.1021	9	.18046	-.778	-.256
2	4.0918	4.1500	11	.20973	4.351	-1.879
Total	4.1024	4.1260	20	.19237	2.314	-1.274

Table 5-2: Normal Distribution Assessment

These issues may in part be a result of the low N in this study and may not be indicative of a wider problem with the instrument or method. In a more detailed analysis with a larger subject pool it may be advantageous to apply a transformation to this data to bring it closer to a normal curve. Since a t-test assumes a normal distribution we relied on the more robust ANOVA for analysis. ANOVA controls the Type I error rate such that skewness and kurtosis have little effect (Field, 2009).

5.3.1 Correlation

Given our model, the first 2-factor ANOVA was performed with Engagement as the dependent variable, and Round as a blocking variable. Engagement in this test was derived as the mean of all questions contributing to this factor. The treatment variable in

these studies of correlation is always Version since we are looking for variance between the two versions. The results of this test showed there was a significant effect of Version ($p < .05$) on Engagement after controlling for any effect from Round (see Table 5-3: Engagement ANOVA). Round was not found to have a significant impact on the Engagement measure.

Engagement Tests of Between-Subjects Effects

Dependent Variable: Engagement

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.686[a]	2	.343	4.259	.032
Intercept	342.620	1	342.620	4255.744	.000
Round_first	.179	1	.179	2.222	.154
Version_first	.507	1	.507	6.296	.023
Error	1.369	17	.081		
Total	344.674	20			
Corrected Total	2.054	19			

a. R Squared = .334 (Adjusted R Squared = .255)

Table 5-3: Engagement ANOVA

Descriptive Statistics

Dependent Variable:Engagement

Round	Version	Mean	Std. Deviation	N
1	1	4.1250	.16105	5
	2	3.9937	.22855	6
	Total	4.0534	.20295	11
2	1	4.5500	.13472	4
	2	3.9983	.42803	5
	Total	4.2435	.42772	9
Total	1	4.3139	.26448	9
	2	3.9958	.31529	11
	Total	4.1390	.32882	20

Table 5-4: Descriptive Statistics for Engagement

The descriptives show that the direction of this effect was negative, that the mean of Engagement was higher in v1 and lower in v2. This can be seen visually represented in the boxplot in Figure 5-3 and numerically represented in Table 5-4.

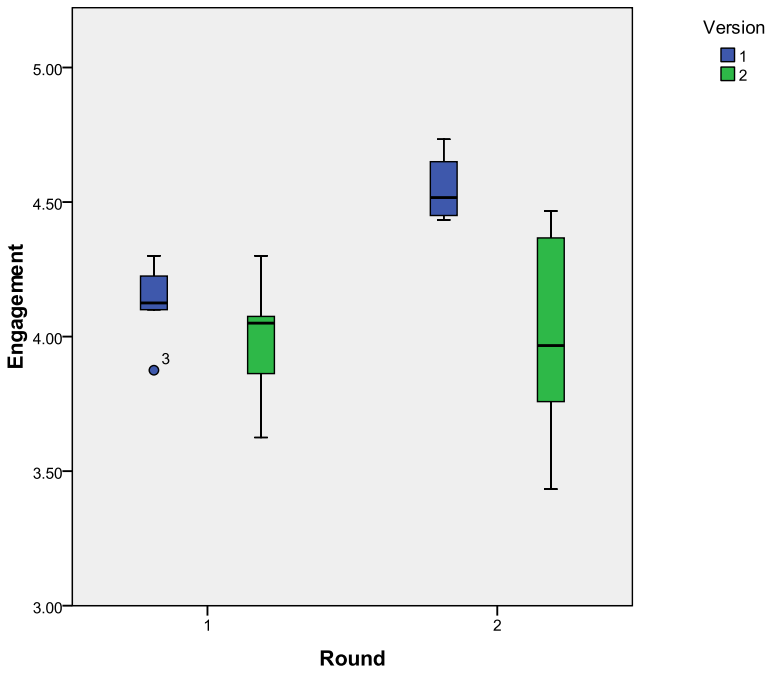


Figure 5-3: Boxplot of Engagement Means

The impact of Version on Immersion was then tested via an ANCOVA with Immersion as the dependent variable, Rounds again as a blocking variable, Version as the treatment factor, and the Engagement variable this time as a co-variate. The Immersion variable was derived as the mean of all questions contributing to the Immersion factor. Effect was removed in the following order: Rounds, Engagement, and then Version. The results can be seen in Table 5-5 below. Adding Engagement as the co-variate in this equation allows us to test whether Version had an impact on Immersion after the significant contribution discovered from Engagement has been removed. The results indicated that there was no significant variance in Immersion and that Version did not have a significant impact on Immersion. The impact of Rounds on Immersion was also insignificant. See Table 5-5 below.

In this test, Engagement did not show statistical significance as a factor in Immersion at the P < .05 level. This undermines the validity of the model which states that Engagement contributes to Immersion. However it should be noted the p-value for the impact of Engagement on Immersion was relatively quite low and close to the generally accepted measure of statistical significance, in this study p = .053. For an exploratory study with a very low N, this gives us some confidence that the model still has relevance.

Immersion Tests of Between-Subjects Effects

Dependent Variable: Immersion

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1.815[a]	3	.605	1.900	.170
Intercept	345.306	1	345.306	1084.421	.000
Round_first	.427	1	.427	1.341	.264
Engagement	1.384	1	1.384	4.348	.053
Version_first	.004	1	.004	.011	.917
Error	5.095	16	.318		
Total	352.216	20			
Corrected Total	6.910	19			

a. R Squared = .263 (Adjusted R Squared = .124)

Table 5-5: Immersion ANCOVA

The effect of the treatment factor, Version, was then tested via an ANCOVA with Social Presence as the dependent variable, derived as the means of all Social Presence questions, Rounds as a blocking variable, and both Immersion and Engagement as co-variates. Effect was removed in the following order: Rounds, Immersion, Engagement, and then Version. This test showed that there was a significant (p < .05) variance in

Social Presence. It was not however the result of the treatment variable; Version had no significant impact on Social Presence. Immersion, however, did show up as strongly correlated to Social Presence ($p < .01$), therefore, confirming our model that Immersion contributes to Social Presence. Results are shown in Table 5-6.

Social Presence Tests of Between-Subjects Effects

Dependent Variable:Social Presence

Source	Type I Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.390[a]	4	.097	4.668	.012
Intercept	336.589	1	336.589	16118.594	.000
Round_first	.021	1	.021	.998	.334
Immersion	.357	1	.357	17.105	.001
Engagement	.005	1	.005	.254	.622
Version_first	.007	1	.007	.314	.584
Error	.313	15	.021		
Total	337.292	20			
Corrected Total	.703	19			

a. R Squared = .555 (Adjusted R Squared = .436)

Table 5-6: Social Presence ANCOVA

A closer look at the individual dimensions which made up these factors, three for Engagement, four for Immersion and six for Social Presence, revealed that only one showed statistically significant results at the $p < .05$ level. It is not uncommon that individual dimensions do not show since the statistical power of an instrument comes from the combination of dimensions that contribute to each factor. However, looking at the relative strengths of the p-values for these individual dimensions may help clarify some of the results.

In Engagement, of the three factors which contributed to this measure, both Interaction and Control both showed low p-values, while Intrinsic Interest did not (Table 5-7). As can be seen in Table 5-8, Version 1 resulted in higher means for both Control and Interaction.

Dependent Variable	Type I Sum of Squares	df	Mean Square	F	Sig.
Intrinsic Interest	.345[a]	2	.173	1.023	.381
Control	.708[a]	2	.354	3.181	.067
Interaction	2.040	2	1.020	2.745	.093

Table 5-7: Significance of Engagement Dimensions

Descriptive Statistics

Dependent Variable: Control

Round	Version	Mean	Std. Deviation	N
Total	1	4.3500	.30362	9
	2	3.9773	.34250	11
	Total	4.1450	.36980	20

Dependent Variable: Interaction

Round	Version	Mean	Std. Deviation	N
Total	1	4.5333	.63713	9
	2	3.9432	.58182	11
	Total	4.2088	.66319	20

Table 5-8: Descriptives of Engagement Dimensions

In Immersion, of the four factors which contributed to this measure, Surrounding showed a low p-value (Table 5-9). In this case, the primary variable which contributed to this low p-value was Round, resulting in a higher mean for Round 2. This indicates that the rewording of the question resulted in variance.

Dependent Variable	Type I Sum of Squares	df	Mean Square	F	Sig.
Inclusive	1.250[a]	3	.417	.797	.513
Extensive	2.995[a]	3	.998	1.686	.210
Surrounding	2.465[a]	3	.822	2.727	.079
Vivid	3.075[a]	3	1.025	.879	.473

Descriptive Statistics

Round	Version	Mean	Std. Deviation	N
1	Total	4.0125	.53895	11
2	Total	4.5333	.61644	9

Table 5-9: Significance and Descriptives for Immersion: Surrounding

In Social Presence, of the six factors which contributed to this measure, Perceived Emotional Interdependence showed significance ($p < .05$) and Perceived Behavioral Interdependence showed a very low relative p-value ($p = .052$) seen in Table 5-10. In these cases, both resulted from a strong correlation with Immersion ($p < .05$) rather than as a result of any effect from Round or Version.

Dependent Variable	Type I Sum of Squares	df	Mean Square	F	Sig.
Co-presence	.725[a]	4	.181	1.446	.267
Attentional Allocation	.366[a]	4	.091	.977	.449
Perceived Affective Understandig	.623[a]	4	.156	1.146	.373
Perceived Emotional Interdependence	1.495[a]	4	.374	3.163	.045
Perceived Message Understanding	.225[a]	4	.056	.448	.772
Perceived Behavioral Interdependence	1.186[a]	4	.296	3.013	.052

Table 5-10: Significance for Social Presence Dimensions

6: DISCUSSION

This section discusses the results from the three stages of this project and assesses implications in each. Implications from the participant observation stage were incorporated into the original design for Version 1, and implications from the first pilot studies were incorporated into Version 2. The discussion of the quantitative analysis reflects on the early hypotheses of the project and evaluates the assumptions made as well as the wider implications of the work and the model used.

6.1 Contextual Observation

There were two primary areas where the contextual observations shaped the production of this project, the general soundscape for the construction of an augmented acoustic ecology in the space, and the sonic events that initiated interaction between people in the space.

The general soundscape in cafés was inhabited by background noise and sound events. Background noise ranged from single long continuous sounds to combinations of many closely-occurring, low to mid-volume sounds. When continuous sounds were extremely loud or there were several continuous sounds occurring simultaneously any sound events were lost in the din. If there were many sound events occurring in succession it was difficult to isolate any one specific sound and they all merely became background noise. Sound events that broke through the background noise were those that were either significantly louder than the rest or were of a significantly different sonic

character. Conversely when there were no continuous or background sounds at all every little sound created an alarming sound event.

These observations reflected in the social interaction that happened in these cafés. When the level of background sound diminished, there seemed to be a conscious lowering of the volumes of action sounds in the space. Since every sound event attracted maximum attention in the absence of background noise, most participants avoided the attention and muted the sounds they were making. Likewise when there were loud continuous and background sounds the volume of every other sound in the café seemed to increase in volume to accompany them.

Rarely did the continuous sounds break participants' attention. This allowed the sound events to peek out from the general background noise to a greater and lesser degree and gave participants a choice of which events would attract their attention. Again it was the loudest or most peculiar sounds that attracted the most attention. Frequently when loud or peculiar sounds occurred, people in the café would look up from what they were doing and look in the direction of the sound. In the case that they made eye contact with the person that created the sound, the response was typically either an embarrassed return glance or a smile. Though frequently the sound creator would hide from eye contact by increasing their focus in what they were doing. In other cases, spectators to a sound event would look up and make eye contact with other spectators. This resulted a wide range of interactions based on the sounds and what caused them and seemed to hint at a growing social presence in the space. Spectators who had made prior eye contact, frequently exchanged eye contact on subsequent occasions as a result of new sonic

events. The interactions that occurred between spectators were varied widely and were too numerous to effectively conjecture a categorization for the purposes of this study.

6.1.1 Implications

There were three key social effects which became the target of this project. First, that participants in this environment would feel comfortable to make sounds when the general background noise in the space was continuous and loud, and would moderate and reduce their volume when the background sounds faded. Second, that participants would look up when there was a peculiar noise and focus their attention on the direction of the sound and relate the person sitting in the direction of the sound to the noise that was made. Third, that there was often a shared moment when participants who were not proximal to a sound source caught each other's eye – these often recurred between two individuals.

In order to effectively take advantage of this social acoustic environment it was necessary to create a system which would generate a varying soundscape which could reproduce and augment these types of sounds in the space. The system would have to maintain an awareness of the level of background sound it was creating to be able to ensure that it varied in both character and volume throughout the performance. Likewise the system would have to have an awareness of the audio files it was choosing so as to create sound events which would emerge from the background sounds in the performance.

It was clear, through the purposeful production of sounds by participants in the space as well as the universal people-watching going on in these environments, that

whatever activities people were engaged in at a café, there was a simultaneous exhibitionism and voyeurism which could be activated and augmented through a properly tailored sonic performance. The performance system would have to be implemented via a distributed sound system which placed the source of sounds proximal to individuals in the environment. There would have to be a voluntary participation where participants would feel comfortable owning the sounds which would arise from the distributed source closest to them. Any interface should be minimal so as not to distract participants from the possibility of visually responding to audio cues. Lastly, due to the transitory nature of this setting the system would have to support a variable number of participants and allow people to come and go during the performance, integrating them when they arrived and adjusting for their absence when they left.

These are the art-design lessons taken from the participant observation stage of this research which served as the basis for creating the system for Eavesdropping Version 1 which is documented in Chapter 3.

6.2 Pilot Studies

This section describes the implications of the two pilot studies that were run and the resulting design changes incorporated into Version 2.

6.2.1 Version 1

The pilot studies of Version 1 confirmed some of the original assumptions regarding the proximity of a participant to a sound event source and the attention it would attract to that participant. This resulted in many successful social interactions between spectators and the participant sitting at a sound. Likewise, unusual or loud sound events

also successfully caused spectators to look around and see how other spectators were reacting to these sounds. This caused interaction between spectators which also seemed to impact social presence during the performance.

However, the Version 1 pilot studies raised issues of immersion and engagement. First, participants indicated not feeling a strong connection to the audio and acted with the intention of imparting some personal affect on the audio coming from the computer in front of them (such as moving their laptops around, opening browser windows, and playing their own audio). While the original intent of the project did not assume that the mood audio being performed at each computer would specifically relate to the participant at that machine, there was an assumption that the participant would merely react by giving some public display of the extent to which the audio did or did not relate to their state and that these behaviors would be exchanged between participants in the room. While this was successful, the dissociation caused subjects to engage in additional activities which were disruptive to the performance.

Second, participants felt that the performance had taken over their laptops and that they had lost control of their systems. They engaged in several activities to regain control of their systems. These actions were outside the bounds of the intended behavior for this art project where the goal relied on participants feeling a connection to their computers to encourage an association with the audio by proximity.

These two negative implications were the impetus for the redesign of the project for Version 2. In order to solve the issue of a lack of immersion due to the mismatch between the audio and the participant's actual mood, we sought to create a more accurate self-representation. This was addressed in two ways. First, allowing participants to input

their own moods at the beginning of the performance ensured that they would feel they had input in the performance and that it would relate to them. However, since the mood-to-music representations have been shown to often be subjective, the music might still be seen as a poor representation. The second solution, an optional mood reinforcement question, gave participants the ability to correct the system and let it know whether the sounds matched their moods each time an audio file played on their computer to ensure that participants would not feel helpless behind a computer making sounds that did not represent them.

The issue of control that had come up in the Version 1 pilot studies was also intended to be addressed by the optional mood reinforcement. Giving participants a meaningful interaction would offer a sense that they were having an effect and that they had agency over the outcome of the performance.

6.2.2 Version 2

The results of the pilot studies for Version 2 indicated that the changes made to the system had solved the initial issues but had also introduced some problematic effects. The task oriented nature of the interaction system was a significant distraction and participants were drawn away from attention to other participants in the room by an interest in the brief moment of interaction with their computer each time the audio file changed, which in the case of most audio sets was about 15 seconds. This seemed to have a negative effect on the social presence in the space since people were less freely looking around the room and allowing themselves the time to foster any connection with other participants via eye contact and reaction to sound events.

While the subjects all indicated that they felt a close relationship with the audio from inputting their mood at the beginning of the performance, they also indicated that they wanted to input their mood more often. Despite the fact that participants acknowledged that increased interaction created problems with the original design motivations of the art performance, they wanted even more interaction which would no doubt have even further reduced their social awareness of each other. Almost all suggestions from the Version 2 pilots indicated a desire to increase interaction with the system, perhaps this was to consciously or unconsciously to avoid the social interaction.

While computers are often used in various environments in a completely passive way, for instance playing background music or watching a movie, when a user is sitting directly in front of the computer there seems to be a strong desire to be in constant manipulation of that system. In the original café observations people with laptops remained focused on their interfaces until they were distracted by audio events or because what they were doing on their computer no longer held their interest and they looked up to find some distraction. I did not document how long each laptop user remained head down and focused on their laptop in these sessions between looking up and around the room, there may be some valuable relationship between the amount of time interactive interest can be maintained between distractions. Perhaps the performance interface offered too much interest to participants and they remained focused on the novelty of the experience at the interface for the entire length of the 5 minute performance. If the performance were to run for 10 or 15 minutes instead of 5 they may have gotten bored with responding to the 15 second event on the computer screen and might have just let the opportunity to interact pass without their involvement. The interest may have

changed focus to the more variable set of events in the room found in the social interactions and other peoples' responses to the audio.

6.3 Social Presence Analysis

The social presence study looked to validate the model of Engagement, Immersion and Social Presence constructed for this project and to evaluate any variance between the two versions of the Eavesdropping system within this model.

6.3.1 Engagement, Immersion, Social Presence Model

The results showed a positive confirmation of the model in which Engagement contributes to Immersion and Immersion contributes to Social Presence. Validation of this model required showing a correlation between Engagement and Immersion, a correlation between Immersion and Social Presence, and confirmation that there is no direct correlation between Engagement and Social Presence.

In evaluating the correlation between Engagement and Immersion, we did not get statistical significance at the $p < .05$ level, however, we did achieve a very low p-value which suggests that we are on the right track. One explanation for this may have been the changes made to the Engagement questionnaire between R1 and R2 since the significance of Round on the Engagement variable was also quite high. A further round of study using the identical questionnaire that was used in R2 might show a stronger relationship between Engagement and Immersion.

Immersion and Social Presence were strongly correlated in this study, giving a very clear validation of the relationship between these two factors in the model. Additionally, there was no significant correlation found between Engagement and Social

Presence directly, further confirming that Engagement effects are mediated through Immersion.

6.3.2 System Evaluation

The overall expectation was that, due to the changes made to the system, participants would show increased Engagement and Immersion resulting in increased Social Presence in Version 2 of the Eavesdropping system. It was anticipated that there would be a distracting effect from the increased interaction at the user interface in Version 2 and that this would show up in the Engagement variable, but that it would not be great enough to undermine the positive effects of Engagement on Immersion and subsequently on Social Presence.

In fact, the results showed that there was no statistically significant difference in Social Presence between versions. Additionally there was no statistically significant difference in Immersion between versions. The only statistically significant effect as a result of changes between versions showed up in Engagement. The data confirmed that the increases in interactivity and control added to Version 2 did in fact have a negative impact on Engagement, however they were so strong that they undermined any positive effect and the Version 1 system resulted in higher Engagement than in Version 2.

While none of the individual dimensions of Engagement showed statistical significance at the $p < .05$ level, both the Control and Interaction variables resulted in comparatively low p-values suggesting that these were the areas which had the most impact on the result in Engagement. The effect from the Control variable was anticipated. This variable reflected a participant's perception of the difficulty of the

interface and the level of free will they felt they had in the interaction (versus feeling that they had to pay continuous close attention to the system).

The effect from Interaction described the perception that participants had agency over what was happening. The decrease in Interaction from this variable in Version 2 might suggest that although we gave participants the ability to input their moods and to teach the system, which alone may have resulted in positive results, we also severely constrained these options. Participants could set their mood only at the beginning of the performance, and although we showed the same exact mood matrix screen during the performance, we did not let them modify their moods. Additionally, we encouraged participants to teach the system with the mood reinforcement question but we did not give this interaction an immediate response to allow them to feel agency in their action. For instance, if they gave a 'No' answer, the audio merely continued playing and due to the way the exploration of the mood A.I. works, the participant may even be given that same audio file again. This is actually fundamental principle in interaction design, suggesting that results from user's actions should show an immediate and significant response. Asking the participants to do work for the system every 15 seconds and merely changing the color of 'No' button in response likely resulted in some negative feelings toward the interactivity of this system.

Another issue with the Interaction variable is in the wording of the questions. Despite the fact that we got no significant impact from Round as a blocking variable, some questions might still be misleading and not accurately depict the interaction effect with the performance system. Round 1 Interaction questions made reference to 'interaction with the system' when for Version 1 participants there was no interaction

with the system. The wording could have been misinterpreted and might have led to errors in the results. To correct this issue the wording was made more general, but it was taken too far possibly prompting errors in the opposite direction (see Appendix A: Questionnaire for specific wording of each question and question pair). The wording should probably specifically identify that it is looking for impact to the system. Perhaps the questions should be modified to: "The effect I had on the art performance computer system was...", and "The effect I had on which sounds the system chose was…".

One expectation which did not occur was an increase in the Inclusive variable of Immersion, meant to measure the effectiveness of the self-representation. It was expected that the Version 2 system would result in participants feeling much closer to the audio that was coming from their machine but the data does not confirm this. The results from the Immersion variable were not statistically significant, as were the results from the Inclusive variable. A possible explanation might be that participants were willing to role-play whatever sounds their computer made in Version 1and adopted the persona of the sound at their computer as a means to interact with other participants. Conversely, when given the specific option to input their mood in Version 2, it is likely that participants were less willing to accept errors in self-representation. Additionally, this negative reaction to mismatched representations was probably strengthened by the fact that the interface allowed them to specifically answer 'No' in these instances. A thorough analysis of the mood reinforcement system data has not yet been completed for this study, however, a quick look at the mean of the sum of positive (1) and negative (-1) responses shows that there were slightly more negative responses than positive (mean = -1.0318).

6.3.3 Limitations of the Study

This research was designed as an exploration of an art system and as such there were several limitations which arose primarily as a result of the vast number of variables involved in this study. Some of those limitations are discussed in this section.

Some issues arose from the subject sampling done for this study. For the most part this study used a convenience sample from the undergraduate population on the Surrey Campus of Simon Fraser University, with a large portion of subjects coming from one first year class. This convenience sample did not select subjects which were necessarily café-goers, which have already been identified as people with a penchant for the voyeuristic and exhibitionistic exchange that goes on in a public place like a café. The intended audience for this performance would likely have shown a ready willingness to interact socially across the room which may have had an impact on both the interest in the performance and the social presence results.

Instead, the demographics show that the subjects selected were quite young and had very little experience with art, interactive art, and performance art. This lack of familiarity with art phenomenon may have undermined their ability to fully participate in the social experience that the art system was attempting to instigate. The novelty of this experience may also have led subjects to focus on inconsequential effects rather than to focus on more meaningful interactions. Several subjects mentioned in comments that they had never participated in anything like this before. (This is not to say that café-goers are necessarily more apt to be older or art savvy but only to mention that the demographic of the subject pool for this study was a fairly homogenous group, unlike what one might find in a café).

This population also had a high number of students born in foreign countries and who spoke English as a second language. After the first round of studies it was suggested that perhaps there were language barriers to understanding the questionnaire. This contributed to some of the rewording in the Engagement / Immersion questionnaire and the language was reduced from a seventh grade reading level to a fifth grade reading level. I explored the possibility of asking an English proficiency question in the questionnaire but research into this showed that it is difficult to get an effective measure without asking a significant number of questions.

Finally, despite the fact that 101 subjects participated in the study, due to the study design where each unit consisted of a group of 5 subjects, the study resulted in low N of 20 units.

There were other types of limitations which were a result of context effects. For instance, the performance was designed for a public café environment yet the studies were run in a black box. In a larger public environment like a café with non-participants, background noise, and perhaps a more familiar setting to subjects, the results might have been different. Additionally, in the café environment each participant would have their own laptop. In the user studies, laptops were supplied and set up ahead of time. This likely reduced the comfort subject had with manipulating the machine during the performance (though in pilot studies Version 2 still saw a reduction in disruptive actions, even when subjects were using their own machines).

Despite the limitations, the study produced several positive results, including implications for the design of the system, design of quantitative studies in media arts experience, and a model which relates engagement, immersion and social presence.

6.3.4 Implications

This model has been designed to cover a broad set of non-overlapping, interface and social factors which effectively map user experience in a low interactivity, co-located, augmented reality, art performance. In its intent to relate the interactive and environmental factors of engagement and immersion with social presence this model allows us to directly address social elements in HCI designs. As a model for designing social art phenomenon this presents a framework for understanding the effects we can anticipate from interactive and immersive elements.

7: SUMMARY AND CONCLUSIONS

This work explores the creation of a co-located, networked audio art installation through a mixed method approach employing arts research and a quantitative user study. The motivation behind this work is to increase social presence in public spaces like a café where several computer users are gathered via an audio augmented reality. During the process of developing this work a functional model relating the HCI concepts of engagement and immersion to social presence was constructed to guide the interaction design for the system. This chapter reviews the work done through the frameworks used for exploration, development and investigation of this project and proposes future work within these contexts with this project and in this area.

7.1 Art Research

Arts practice research offers a post-disciplinary perspective for knowledge generation through systematic inquiry which may employ several methodologies and post-hoc investigations. The goal of arts practice research is frequently focused on identifying possibilities or transforming awareness. A common framework for research in the arts is evaluation through an outside perspective. In interactive art, the most applicable outside perspective is that of human computer interaction (HCI). The process of arts practice research is often cyclical, operating through processes of practice, theory generation, and evaluation.

The Eavesdropping project was generated through a contextual investigation, development of the first version, pilot studies, development of a second version, further pilot studies, and finally quantitative user studies.

7.2 Acoustic Ecology

This project takes advantage of existing behavior in acoustic ecology to create social awareness in a public space. In these spaces there are background sounds and sound events generated by humans as part of our daily activities. We associate meanings and emotions with different sounds and when we hear a sound that attracts our attention we look in the direction of the sound. We subsequently associate the meanings we hear in the sound with whoever is at the sound source. By utilizing localized sound sources like laptops in a café to generate sounds this project seeks to associate the person at that laptop with the meaning identified from the audio. When someone looks up in response to a sound and makes eye contact with someone else in the room, they have a shared moment in acknowledgement of that sound. These shared moments can develop a connection and social presence between people in the same room.

7.3 Social Presence

One of the contributions of this research work is the development of a model that addresses Social Presence in mediated interactive art performances. Social presence is identified by the degree of salience of another person in an interaction. Biocca and Harms identify six factors of social presence as awareness or co-presence, allocated attention, the capacity for message understanding and affective comprehension, and the capacity for both affective and behavioral interdependence (Frank Biocca et al., 2003).

Their work examines social presence in mediated environments which is constructed on a notion of presence, or the sense of 'being there' in a mediated environment. Presence is often used interchangeably with immersion, which in augmented reality is marked by a self-representation which is inclusive in the mediated environment, the extensiveness of the representations, whether the generated environment surrounds the subject, and the vividness or quality of the media. Where interaction is involved, it impacts immersion through a sense of engagement, which can be represented through intrinsic interest, interactivity, and control.

These theories served as the foundations for construction of a model where engagement contributes to immersion which further contributes to social presence. By affecting any of the factors we can effectively impact social presence.

7.4 Eavesdropping System

Two versions of the Eavesdropping art installation were created to facilitate social presence in public spaces. Version 1 performs a passive mood composition where participants join the performance and audio files are then played at their computers and at all other participants' computers. Version 2 asks participants to enter their moods into the interface and then selects audio files to match those moods to play back at their computer. They are given an optional question allowing them to let the system know whether the audio file matches their mood or not.

Participants in both versions hear a combination of audio from their own computer and from all other participant's computers. When a loud or interesting sound is

made at one computer, it may pique a participant or several participants' interest at which point they may look up and make a social presence connection with others in the room.

7.5 Three Studies

The iterative process to create a co-located, networked audio art installation went through a series of theory, practice and evaluation phases. Three evaluation phases informed the development and assessed the results of this project through: contextual observations, pilot studies, and quantitative user studies.

The contextual observations saw café goers participating in a social acoustic ecology to attract and hide from the attention of others in this highly public space. As background noise rose, so did the comfort level of people in the environment to make noises themselves. As background noise fell, likewise people in the environment lowered their sound profile in the absence of masking sounds. Loud and peculiar sound events attracted attention. The implications of this study suggest that the people feel an association with sounds that are in close proximity themselves and conversely make an association between others and sounds in others' proximity.

Two rounds of pilot studies, after each version of the system was developed, offered an opportunity to examine the art performance in context. In the first pilot studies, run on Version 1, participants were involved in all sorts of antics with their computers during the performance. This was evaluated as a response to not accepting the association between themselves and the sounds from their computer, and feelings of loss of control over the computer at which they were sitting as they yielded their system to the performance for a period. Utilizing the model of engagement, immersion and social

presence, these issues were addressed by adding an effective self-representation to Version 2 of the system to increase a sense of immersion, and by adding a meaningful interaction to increase engagement. The second round of pilot studies, run on Version 2, saw none of the disruptive behavior of the Version 1 pilots but also seemed to lose some of the social presence. This was attributed to an overt focus on the interaction task, an over-attentive engagement which withdrew the participant from the social environment.

The third phase compared the two versions of the system using the model of engagement, immersion and social presence via quantitative questionnaires to both validate the model and assess the effectiveness of the changes made between versions. A questionnaire was constructed based on the model to measure engagement and immersion in this study, social presence was measured using the popular Networked Minds Measure of Social Presence. Results strongly confirmed the relationship between immersion and social presence and suggested that the relationship between engagement and immersion may be correlated since statistical analysis showed an almost significant p-value ($P < .053$). External issues such as variations in the questionnaire between rounds and a low N may have confounded this result. The data went on to show that version had a significant impact on engagement but on none of the other variables. This impact however indicated that the non-interactive Version 1 had a higher engagement than the interactive Version 2. This confirmed assumptions that the interactivity added to Version 2 both added complexity and reduced participants feeling of control over what they had to do in the performance. Adding the mood reinforcement task had too large an impact and drew participants away from attention to the social environment.

7.6 Future Work

This research has been exploratory by nature, as such, it moved rapidly through a rich area and left many areas untouched. Additionally, it uncovered new areas for research as well as collected far more data than was initially explored in this investigation. Some of the more interesting areas for future work are explored in this section.

The initial investigation into social audio interaction in public spaces revealed a vast area for research which likely relates to research in sociology and anthropology as well as studies in non-verbal communications. People in the café environment seemed to be speaking mostly unconsciously in a meta-language of audio cues, fidgets, action sounds and gestures. The meanings of most of these interactions were unclear and the scope of this study did not allow for the extensive study design required to fully examine audio interaction in this environment. If this non-verbal language in public spaces was documented, even in part, it could lead to manipulation through localized automated or performative sonic events generated in the space.

Another facet for exploration in a public space is a deeper study of attention, either as a cognitive-science study of a subject's focus, or as an HCI experience looking at how long a subject will remain on task in the presence of intentionally distracting audio events. The development for Version 2 of the Eavesdropping system made incorrect assumptions about the draw of task-oriented interaction. A more substantial study of attention might show that subjects will only remain focused by the novelty of a task-oriented interaction for a limited time. Such a study might explore how this task-based attention can be minimized for a quicker acclimation to allow attention to other

effects. An even narrower study might be done in the café environment by examining amount of time focused on a laptop (or book, or other focused action) between looking up and around the room. This also hints at running the existing Eavesdropping study but offering a longer performance in the hope that subjects will get bored with answering the mood reinforcement question. Another option would be to add some example mood reinforcement interactions to the tutorial to reduce the novelty when subjects encounter it during the performance.

Context was raised as a limitation in this study. Certainly an obvious direction for future work would be to re-run this study in a café environment or several café environments. This of course introduces many new variables such as existing noise present in the café, or the relationship between people in the café, or whether people are regulars versus having attended the café for the first time, or the varied placement of the participants during the performance, however, the more natural environment might produce more natural effects. However, running the study in several contexts would also allow adding contextual location as a blocking variable in the data to separate out context effects from the data.

Evaluation of the effect of the proximity of audio in garnering attention to a subject could be performed via a control group in a room with no audio cues and ear plugs (evaluating whether subjects naturally look at each other even without audio impetus). Special attention would be required to ensure that there was some minimal information on a laptop screen to mimic the environment in the general study.

In relation to the development of the engagement, immersion, social presence model, the terminology for the factors of immersion should probably be updated. It was

convenient to adopt the terminology used by Slater and Wilbur but it is confusing to change the meaning of one of their terms. Additionally, due to the fact that their model originally had six factors and the new design has four, the entire set of terms should be redesigned. The new meanings are relevant; it is the reuse of a pre-defined term with a different meaning that causes confusion.

Lastly, this research project collected a significant amount of data that was not fully tapped for information. ANOVA and ANCOVAs were run to explore effects within the model but no post hoc analysis was run to seek out further information from these relationships. Also, the data set collected by the server system during the performances contains more information that can help to shed more light about the participants and the art system. For instance, this data includes every file played by every participant as well as all moods set and mood reinforcements given. There may be interesting relationships between number of reinforcements (positive or negative) and interaction, or positive / negative reinforcements and the effectiveness of self-representations for immersion. Subject profiles might be further developed by correlating mood with any of the engagement, immersion, social presence factors, or through exploring how many reinforcements were given, or whether a subject was giving mood reinforcements consistent with those given by other subjects for each audio file.

7.7 Final Words

This thesis created a co-located, networked audio art installation which has received international recognition for its design. Arts phenomena have the power to transform social relationships in public spaces starting at the most basic levels of co-presence and attention.

Inceptive in the arts research, this project produced a model for evaluation of low interactivity, augmented reality, art installations in a social context. This model has established a framework for the development of an instrument to measure engagement and immersion effects and their impact on social presence. This model and instrument can guide and evaluate future forays into locallized audio art for social presence.

APPENDICES

Appendix A: Questionnaire

Key

- Question: the exact wording of the question in the questionnaire
- Type: they type of question (select list, yes/no, Likert scale, Textbox). Note that Varying Scale is merely a Likert scale with question-specific wording.
- R1 / R2: indicates whether the question occurred in the Round 1 or Round 2 questionnaire or both.
    - X: indicates that the question did occur
    - O: indicates that the question was not in this questionnaire

Paired questions from Round 1 are followed with their replacement single questions for Round 2.

Demographic

Demographic 1

Question	Type	R1	R2
What is your gender?	SelectList	X	X
What is your age?	SelectList	X	X
What is the highest level of education you have completed?	SelectList	X	X
Do you own the laptop that you used for this performance?	Yes/No	X	X
Have you participated in a performance of this project before?	Yes/No	X	X

Demographic 2

Question	Type	R1	R2
How many people in this performance are your friends?	Varying Scale	X	X
How well do you know your closest associate in the performance?	Varying Scale	X	O
How well do you know your least close associate in the performance?	Varying Scale	X	O
What is your experience level with fine arts?	Varying Scale	X	X
What is your experience level with interactive art?	Varying Scale	X	X
What is your experience level with audio installation art?	Varying Scale	X	X
Any further comments you would like to share regarding this performance?	Textbox	X	X

Engagement

Intrinsic Interest

Question	Type	Scale Low	Scale High	R1	R2
The performance was interesting to me.	Likert (Agree / Disagree)			X	O
I lost interest in the performance.	Likert (Agree / Disagree)			X	O
The performance was...	Varying Scale	Not Interesting	Very Interesting	O	X
I didn't like the sounds.	Likert (Agree / Disagree)			X	O
The sounds were aesthetically appealing.	Likert (Agree / Disagree)			X	O
I felt the audio was...	Varying Scale	Unenjoyable	Enjoyable	O	X

Interaction

Question	Type	Scale Low	Scale High	R1	R2
My level of interaction with the software was...	Varying Scale	None	Very Much	O	O
My interaction with the computer had no effect on the performance.	Likert (Agree / Disagree)			X	O
My interaction with the computer impacted the performance.	Likert (Agree / Disagree)			X	O

The effect I had on the art performance environment was...	Varying Scale	None	Very Much	O	X
I could hear the results of my input in the sounds my computer produced.	Likert (Agree / Disagree)			X	O
I didn't have any impact on the sounds my computer produced.	Likert (Agree / Disagree)			X	O
The effect I had on the sounds was...	Varying Scale	None	Very Much	O	X

Control

Question	Type	Scale Low	Scale High	R1	R2
I was not able to enjoy the performance due to the computer interface.	Likert (Agree / Disagree)			X	O
I enjoyed interacting with the computer interface.	Likert (Agree / Disagree)			X	O
Rate your level of control over the software.	Varying Scale	Very Low	Very High	O	O
The computer screen required my attention.	Likert (Agree / Disagree)			X	O
I did not feel compelled to pay attention to the computer screen.	Likert (Agree / Disagree)			X	O
The software...	Varying Scale	Required my attention	Did not require my attention	O	X
The interface was simple to understand.	Likert (Agree / Disagree)			X	O
I had difficulty understanding the interface.	Likert (Agree / Disagree)			X	O
The interface was...	Varying Scale	Difficult to understand	Easy to understand	O	X

Immersion

Inclusive

Question	Type	Scale Low	Scale High	R1	R2
The audio playing from my computer effectively represented my mood.	Likert (Agree / Disagree)			X	O
The computer was playing sounds didn't sound like my mood.	Likert (Agree / Disagree)			X	O
I felt like the sounds from my computer...	Varying Scale	Did not represent my mood	Represented my mood well	O	X
I felt like the sounds from my computer were my sounds.	Likert (Agree / Disagree)			X	O
I felt no connection to the sounds from my computer.	Likert (Agree / Disagree)			X	O
I paid more attention to...	Varying	My sounds	Others' sounds	O	O

	Scale				
I could not hear myself represented in the audio.	Likert (Agree / Disagree)			X	O
I felt represented in the audio ecology in the room.	Likert (Agree / Disagree)			X	O
I paid more attention to...	Varying Scale	The software interface	Other people	O	O

Extensive

Question	Type	Scale Low	Scale High	R1	R2
There were a variety of moods represented.	Likert (Agree / Disagree)			X	O
All the moods sounded the same.	Likert (Agree / Disagree)			X	O
The moods the audio tried to create...	Varying Scale	Sounded the same	Were a wide variety	O	X

Surrounding

Question	Type	Scale Low	Scale High	R1	R2
I could not tell whose computer was playing specific audio.	Likert (Agree / Disagree)			X	O
It was easy to hear which audio was coming from specific computers.	Likert (Agree / Disagree)			X	O
I could tell which computers the sounds were coming from.	Varying Scale	Strongly Disagree	Strongly Agree	O	X

Vivid

Question	Type	Scale Low	Scale High	R1	R2
I could not hear the audio very well.	Likert (Agree / Disagree)			X	O
The audio was of good sound quality.	Likert (Agree / Disagree)			X	O
The audio was...	Varying Scale	Poor quality	High quality	O	X

Social Presence

Attentional Allocation

Question	Type	R1	R2
Others were easily distracted from me when other things were going on.	Likert (Agree / Disagree)	O	X
I was easily distracted from others when other things were going on.	Likert (Agree / Disagree)	O	X
Other people were easily distracted when other things were going on.	Likert (Agree / Disagree)	X	O
Others did not receive my full attention.	Likert (Agree / Disagree)	O	X
Other individuals seldom gave me their full attention.	Likert (Agree / Disagree)	X	O
I seldom gave my full attention to other individuals.	Likert (Agree / Disagree)	X	O
I remained focused on others throughout our interaction.	Likert (Agree / Disagree)	O	X
Others remained focused on me throughout our interaction.	Likert (Agree / Disagree)	O	X
Specific people sometimes paid close attention to me.	Likert (Agree / Disagree)	X	O
I was easily distracted when other things were going on.	Likert (Agree / Disagree)	X	O
I did not receive others' full attention.	Likert (Agree / Disagree)	O	X
I sometimes paid close attention to specific people.	Likert (Agree / Disagree)	X	O

Co-Presence

Question	Type	R1	R2
I noticed other individuals.	Likert (Agree / Disagree)	X	X
My presence was obvious to other individuals.	Likert (Agree / Disagree)	X	X
Other individuals noticed me.	Likert (Agree / Disagree)	X	X
I caught other individuals' attention.	Likert (Agree / Disagree)	X	X
Other individuals' presence was obvious to me.	Likert (Agree / Disagree)	X	X
Other individuals caught my attention.	Likert (Agree / Disagree)	X	X

Perceived Affective Understanding

Question	Type	R1	R2
My emotions were not clear to others.	Likert (Agree / Disagree)	O	X
I could tell how others felt.	Likert (Agree / Disagree)	X	X
I could describe others' feelings accurately.	Likert (Agree / Disagree)	O	X
Others could tell how I felt.	Likert (Agree / Disagree)	X	X
Others could describe my feelings accurately.	Likert (Agree / Disagree)	O	X
My moods were not clear to the others.	Likert (Agree / Disagree)	X	O
I could describe others' moods accurately.	Likert (Agree / Disagree)	X	O
The moods of the others were not clear to me.	Likert (Agree / Disagree)	X	O
Others could describe my mood accurately.	Likert (Agree / Disagree)	X	O
Others' emotions were not clear to me.	Likert (Agree / Disagree)	O	X

Perceived Behavioral Interdependence

Question	Type	R1	R2
Others' behavior was closely tied to my behavior.	Likert (Agree / Disagree)	X	X
Others responded to my actions.	Likert (Agree / Disagree)	X	X
My behavior was often a direct response to others' behavior.	Likert (Agree / Disagree)	X	X
My behavior was closely tied to others' behavior.	Likert (Agree / Disagree)	X	X
The behavior of others was often in direct response to my behavior.	Likert (Agree / Disagree)	X	X
I reciprocated other's actions.	Likert (Agree / Disagree)	X	X

Perceived Emotional Interdependence

Question	Type	R1	R2
Others' moods influenced the performance.	Likert (Agree / Disagree)	X	X

My attitudes influenced how others felt.	Likert (Agree / Disagree)	X	X
The other individuals were influenced by my moods.	Likert (Agree / Disagree)	X	X
My mood influenced the performance.	Likert (Agree / Disagree)	X	X
I was sometimes influenced by other individuals' moods.	Likert (Agree / Disagree)	X	X
Others' attitudes influenced how I felt.	Likert (Agree / Disagree)	X	X

Perceived Message Understanding

Question	Type	R1	R2
My thoughts were clear to others.	Likert (Agree / Disagree)	X	X
Others had difficulty understanding me.	Likert (Agree / Disagree)	X	X
Others' thoughts were clear to me.	Likert (Agree / Disagree)	X	X
Understanding others was difficult.	Likert (Agree / Disagree)	X	X
It was easy to understand what others were thinking.	Likert (Agree / Disagree)	X	X
It was easy for others to understand what I was thinking.	Likert (Agree / Disagree)	X	X

Appendix B: Questionnaire Evaluation

Context

- CSCL: Computer Supported Collaborative Learning
- VE: Virtual Environments
- AR: Augmented Reality
- HCI: Human-Computer Interaction

	Concept	Context / Field	Dimensions	Terms
(F. Biocca et al., 2001)	Social Presence	HCI	Co-presence: - Isolation / Inclusion - Mutual Awareness Psychological Involvement: - Mutual Attention -Empathy -Mutual Understanding Behavioral engagement: -Behavioral Interdependence -Mutual Assistance -Dependent Action	Isolation: alone Mutual awareness: Notice, Aware, Felt alone(?) Attention: pretend, paid attention, distracted, ignored Empathy: I / Other happy, influenced, affected, (moods) Mutual Understanding: clear (opinions, thoughts), understood Behavioral Interdependence: dependent, direct response, affected Mutual assistance: helped, worked with, Dependent action: could not act
(Bente, Rüggenberg, & Krämer, 2004)	Social Presence, Embodiment	CSCL, VE	Co-presence, **closeness**, comprehension, contagion and coordination, interpersonal trust (social presence, common ground, group awareness) co-presence, comprehension, connectedness and contingency	
(Bystrom et al., 1999)	Immersion, Social Presence	Tele-presence, VE	Immersion, Sensory Fidelity, Task Requirements, Performance, Presence	

(Kreijns et al., 2004)	Social Presence	Asynchronous, CSCL	Sociability, Social Presence, Social Space	**Sociability**: enables contact, feel lonely, impression of, enabled development, relationships, identify myself, comfortable, friendships **Social Presence**: partner in mind's eye, real persons, abstract anonymous persons, face-to-face **Social Space**: felt free, reached understanding, get / kept in touch, maintained contact, personal information, open and lively conversations, spontaneously conversed, attacked, criticized, suspicious, dislike, obstructed, unreasonable, disagreed, conflicts, gossiped
(Abeele, Roe, & Pandelaere, 2007)	Social Presence, Connectedness	CSCW	Social Presence, Emotional Presence, Connectedness	Social (Perceptual) Presence: feeling someone was Physically in the room, impression …, felt physically alone Emotional presence: emotionally in the room, impression …, emotionally alone Connectedness: connected, supported, presence help, presence feel less lonely
(Dow et al., 2007)	Immersion, Presence, Engagement	Game, VE, HCI	Physical and Self Presence, Social Presence, Dramatic Presence,	
(Vastfjall, 2003)	Presence, AR			
(Kumar & Benbasat, 2002)	Social Presence	HCI	Immediacy, Affect, Similarity, Receptivity, Composure, Formality, Dominance, Equality and Involvement	**Immediacy/Intimacy:** Close / Closeness, Distance / Aloof, Detached, Impersonal **Sense of Understanding:** Understand / Understood, Knows, No Clue (goals, trying to do, desires, wanted) **Positivity:** Likable, pleasant, unfriendly, fun, dislike, positive feelings. **Involvement:** Absorbed, Involved, Holds

				attention, Interested, Excites, Aroused **Dominance:** Tried hard, Persuade, Assertive, Influences, Pushy, Controlled, Aggressive, Over-selling
(Short, 1976)	Presence / Immersion	Telecom		personal–impersonal, sociable–unsociable, sensitive–insensitive, and warm-cold
Rourke and Anderson	Social Presence			personal–impersonal, warm–cold, trusting– untrusting, dis-inhibiting–inhibiting, close–distant, and friendly–unfriendly
Gunawardena and Zittle	Social Presence			Social Presence Scale stimulating–dull, personal–impersonal, sociable–unsociable, sensitive– insensitive, warm–cold, colorful–colorless, interesting–boring, appealing–not appealing, interactive–non-interactive, active–passive, reliable– unreliable, humanizing–dehumanizing, immediate– non-immediate, easy–difficult, efficient– inefficient, unthreatening–threatening, and helpful– hindering.
(Witmer & Singer, 1998)	Presence / Immersion	VE	Control: Degree of control, Immediacy of control, Anticipation of events, Mode of control, Physical environment, Modifiability Sensory: Sensory modality,	Presence Questionnaire & Immersive Tendencies Questionnaire In control, Responsive to actions, natural interactions, all senses engaged, visual involvement, auditory involvement, natural controls, event awareness, control awareness, compelling sense of objects moving through space, inconsistent or disconnected senses, anticipation of

			Environmental richness, Multimodal presentation, Consistency of multimodal information, Degree of movement perception, Active search Distraction: Isolation, Selective attention, Interface awareness Realism: Scene realism, Information consistent with objective world, Meaningfulness of experience, Separation anxiety/ disorientation	response to actions, completeness of visual survey, identify sounds, localize sounds, survey with touch, compelling sense of movement, closely examine objects, move or manipulate objects, confused or disoriented, involved in the experience, distracting controls, delay between actions and outcomes, adjust to virtual space, proficiency with controls at end, visual display quality impact on actions, control interference, concentration, learn techniques, lost track of time Emotionally involved,
(Wijnand A. IJsselsteijn, de Ridder, Freeman, & Avons, 2000)	Presence			
Engagement				
(Chen, Koldo, Cuddihy, & Medina, 2005)	Engagement, Immersion, Flow theory, Fun	Game, Serious Games	Interest, Attention, Immersion (Interface & Fidelity)	**Overall**: Like, Add/Detract Experience with, Problems with interface, Problems with meaning **Flow**: Challenging, motivated to continue, enjoyable **Immersive**: involved, trouble getting your attention, mentally alert, aware of things around you, identify with characters, inside the screen, healthy, block external

				distractions, react physically to things on screen, quality of concentration, daydream, residual emotion after gaming, excited, apprehensive / scared, lose track of time, deeply involved, solving puzzles **Engagement**: anticipate, interaction delay, appropriateness of controls, understand the controls, appropriateness of UI, understand the UI, gain proficiency controlling, enjoyable graphics, enjoyable sound, identify the sounds, consistency between UI elements, lost track of time, inside the game world, Familiarity with genre, enjoyable content, play again, UI interfere, spatially confused or disoriented, explore more, noise distract
(Batras, 2009)	Embodiment, Interaction, Engagement	Autonomous Agents, Interactive Art	Intent	
(Bryan-Kinns & Hamilton, 2009)	Mutual Engagement, Collaboration (social, non-work)	CSCW	Quality, Preference, Content Assessment, Musicality, Communication, Attunement, Interaction, Contribution, Mutual Modification, Proximal Acknowledgement, Mirroring, Transforming, Complementing Interaction: contribution, mutual modification, proximal	Best, Satisfied, Enjoyed, Felt Involved, Understood, Frustrating, Complex,
(O'Brien & Toms, 2010)	Engagement	HCI	Aesthetics, Affect, Focused attention, Challenge, Control, Feedback, Interest,	Forgot about surroundings, involved, ignored, lost myself, lost track of time, blocked out things, time slipped away, absorbed, let myself go, frustrated, confusing, annoyed,

			Motivation, Novelty, Perceived Time	discouraged, mentally taxing, demanding, in control, could not do what I wanted, attractive, aesthetically appealing, liked graphics and images, appealed to visual senses, visually pleasing, worthwhile, success, did not work out as planned, recommend, incited my curiosity, interested, drawn into, fun Sex, Age, Residence (Urban, rural), Education, Occupation,
(Sgouros, 2000)	Engagement	Multimedia performance		
Webster & Ho, 1997 (Webster & Ho, 1997)	Engagement	Presentations, CSCL	attention focus, curiosity, and intrinsic interest, Challenge, Feedback, Presenter Control, Variety, Engagement	Challenge: Challenges, encourages me to think, Feedback: provides direct feedback, clear feedback, Control: maintain control, control of delivery, Variety: incorporates change, uses variety, Engagement – Attention focus: keeps me totally absorbed, holds my attention Curiosity: excites my curiosity, arouses my imagination Intrinsic Interest: fun, intrinsically interesting Overall: engaging
(Zafer Bilda et al., 2008; Edmonds et al., 2006)	Engagement, Interaction	Interactive Art		Don't know what to do, understand interaction, experienced before, didn't get the idea, uncomfortable, critically examine
(Argyle & Dean, 1965)	Social Presence, Engagement	Eye-contact, social interaction, communication		

(McMahan, 2003)	Immersion, Engagement, Presence	Video games		
Immersion				
(Gilroy, Cavazza, & Benayoun, 2009)	Flow	Interactive Art, Augmented Reality		
(Mania & Chalmers, 2001)	Immersion	VE, Memory / Task	Pleasure – arousal - dominance	
(W. IJsselsteijn et al., 2007)	Immersion	Video Games	GEQ – Game Experience Questionnaire	
(Grimshaw & Schott, 2007)	Immersion	Video games		
(Slater & Wilbur, 1997)	Immersion, Presence	VE		
(Lombard & Ditton, 1997)	Presence			
(Ermi & Mäyrä, 2005)				
Connectedness				
(J. van Baren, W. A. IJsselsteijn, N. Romero, P. Markopoulos, & B. de Ruyter, 2003)	Social Presence	Telecom		

(Gonzales, 2009)				
(Wijnand IJsselsteijn, Joy Baren, Panos Markopoulos, Natalia Romero, & Boris Ruyter, 2009)				
(Biemans et al., 2008)	Connectedness			
(Edward S. De Guzman et al., 2004)				
(Panos Markopoulos et al., 2004)	Connectedness			
(Wijnand IJsselsteijn et al., 2009)	Connectedness	HCI	Benefits: Personal Effort, Thinking about each other, Sharing experiences, Staying in touch, Recognition, Group Attraction Costs: Obligations, Expectations, Invasion of Privacy, Process Effort	Obliged to contact, keeps thinking about me, infer how they are doing, contacts take a lot of time, informed, part of a group, stay in touch, feel special, invasion of privacy, knows what I feel, avoid contact, keep to myself, share, regular contact, invest energy, think back, involved, feels valuable, disappointed in lack of contact, share experiences, unity, learns about me, put efforts into making a nice contact, more effort, identify, should respond, hardly thinks about, expects, knows what I am doing, identify with, should respond, infer how I'm doing, feel valuable,
(J. K. Burgoon et al., 2000; Judee K. Burgoon & Hale, 1987)	Connectedness	Communications	dominance, emotional arousal, composure, similarity, formality, task vs. social orientation, intimacy, familiarity, affection, inclusion, trust and intensity	

			of involvement	
(Seo & Gromala, 2007)	Immersion	Art		
(Lessiter, Freeman, Keogh, & Davidoff, 2001)	Presence	IMAX	physical space, engagement, naturalness, and negative effects	*ITC Sense of Presence* *Inventory* (ICT-SOPI)

Appendix C: Announcement

A user study on social presence in an audio art installation is being held today in room … . The entire study takes 15 minutes of your time and consists of a 5 minute audio performance followed by a 5-10 minute questionnaire. This study supports work being done toward my Master's thesis and all data gathered is anonymous. For participation in this study …

Round One

… you will be given a chance to win one of two $50 FutureShop gift certificates.

Round Two

… you will each be given a Blenz coffee gift card worth $10.

Please join me in room … to participate.

Appendix D: Introduction Scripts

Round 1

Version 1

Welcome to the Eavesdropping User Study. You will be presented with a five minute interactive art performance followed by a short questionnaire. Please follow the instructions on your screen to continue.

Eavesdropping is a networked audio performance in which a musician has designed a composition based on moods. The generative system will select audio files from its library to represent the moods the composer has chosen and play them from each participant's laptop. The performance will last for five minutes. Click 'Start' to join the performance.

Version 2

Welcome to the Eavesdropping User Study. You will be presented with a five minute interactive art performance followed by a short questionnaire. Please follow the instructions on your screen to continue.

Eavesdropping is a networked audio performance that plays a mix of audio based on the moods of participants in the room. Each participant enters his or her mood into the interface and the generative system will select audio files from its library to represent that mood to perform from the participant's laptop. You will be offered optional questions while the audio is playing to improve the system's ability to represent moods. The performance will last for five minutes. Click 'Start' to join the performance.

Round 2

Version 1

Welcome to the Eavesdropping User Study. This study is designed to evaluate social presence in an audio art installation. This exhibit was created for an environment such as an internet café where several people are gathered with their laptops – that is the type of environment simulated here today. During the performance you will hear audio playing from your own computer as well as the computers of other participants. Treat this like any art exhibition – take in the sounds from your computer, the sounds around you, be aware of what is happening in the room, and how others in the room are responding to the project and the sounds.

The exhibition will require a bit of setup prior to joining the performance including a short tutorial and an informed consent document. You will then arrive at a 'join performance' button after which you will be presented with the 5 minute art performance followed by a questionnaire.

Note that questionnaire evaluates what happens after you click the 'join performance' button and not the setup portion.

This version of the system is a networked audio performance where a musician has designed a composition based on a series of moods. The server selects audio files to match those moods and tries to achieve a balanced mix of audio in the room based on the number of participants and the time you joined the performance.

Please click the ‘Start or Join a User Study button’ to begin the tutorial. When you get to the screen that says ‘Wait Here’, please wait for everyone to catch up and we will join the performance together.

Version 2

Welcome to the Eavesdropping User Study. This study is designed to evaluate social presence in an audio art installation. This exhibit was created for an environment such as an internet café where several people are gathered with their laptops – that is the type of environment simulated here today. During the performance you will hear audio playing from your own computer as well as the computers of other participants. Treat this like any art exhibition – take in the sounds from your computer, the sounds around you, be aware of what is happening in the room, and how others in the room are responding to the project and the sounds.

The exhibition will require a bit of setup prior to joining the performance including a short tutorial and an informed consent document. You will then arrive at a ‘join performance’ button after which you will be presented with the 5 minute art performance followed by a questionnaire.

Note that questionnaire evaluates what happens after you click the ‘join performance’ button and not the setup portion.

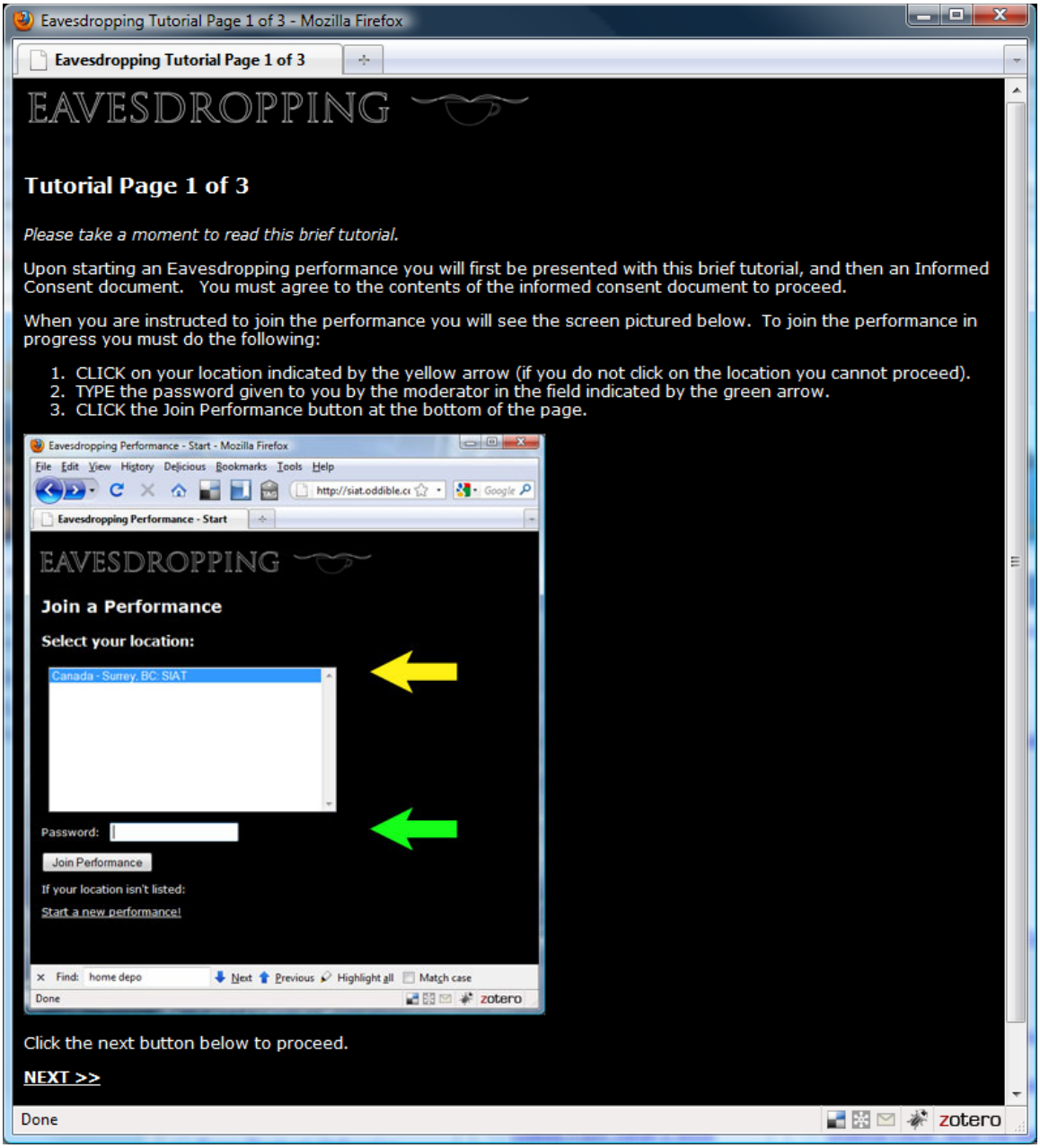
Once the performance has begun, you will indicate your mood in the software interface. The system’s artificial intelligence will attempt to match audio files to your mood to play at your computer. During the performance you will be presented with an

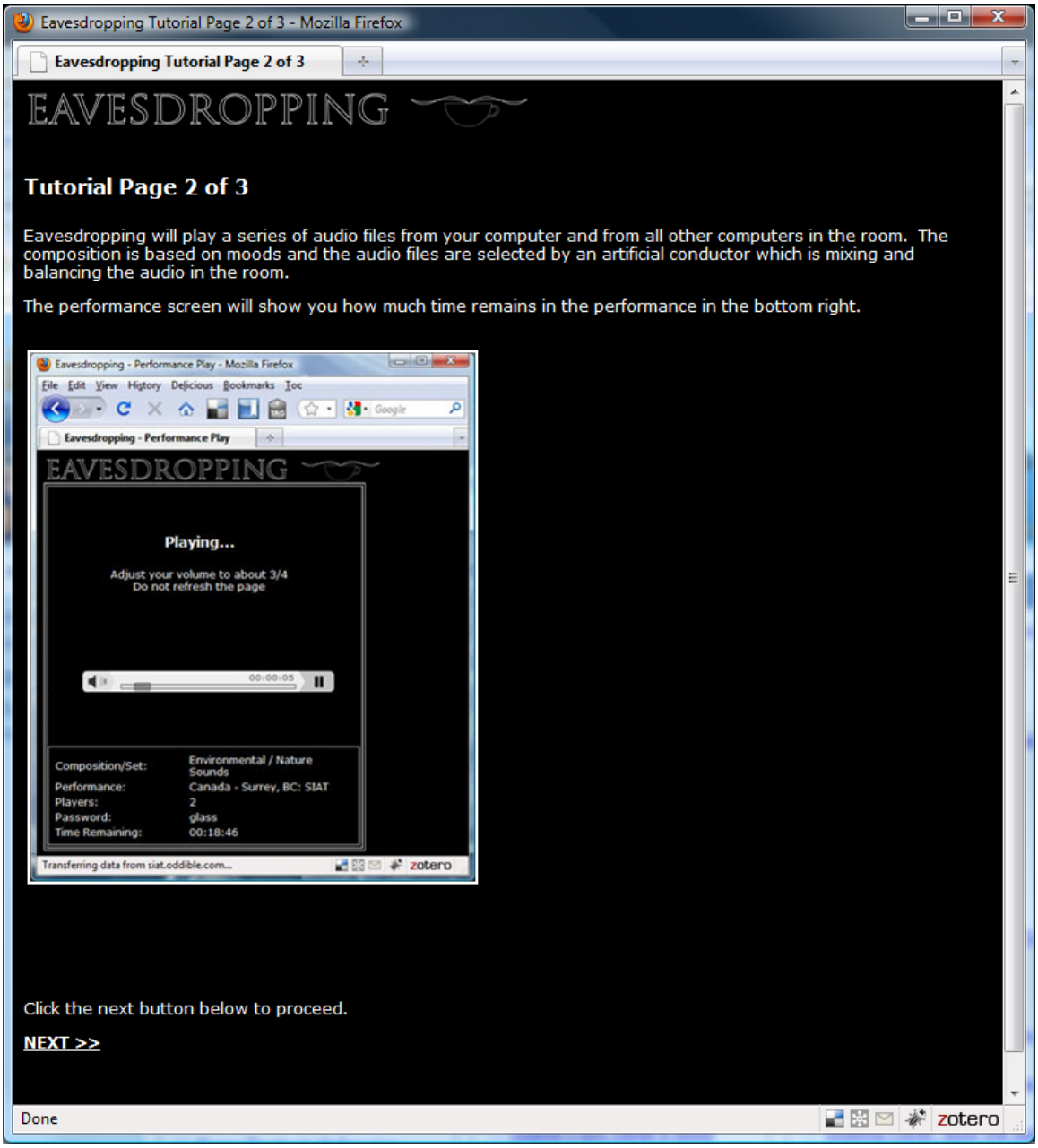
optional question asking whether the audio matches your mood or not, to help train the system to match audio to moods. Note that answering this question will not have an immediate effect and will not change the current audio but will help improve the reinforcement learning artificial intelligence to make better choices of which audio to use for your mood. You do not have to answer this question, the performance will continue either way.

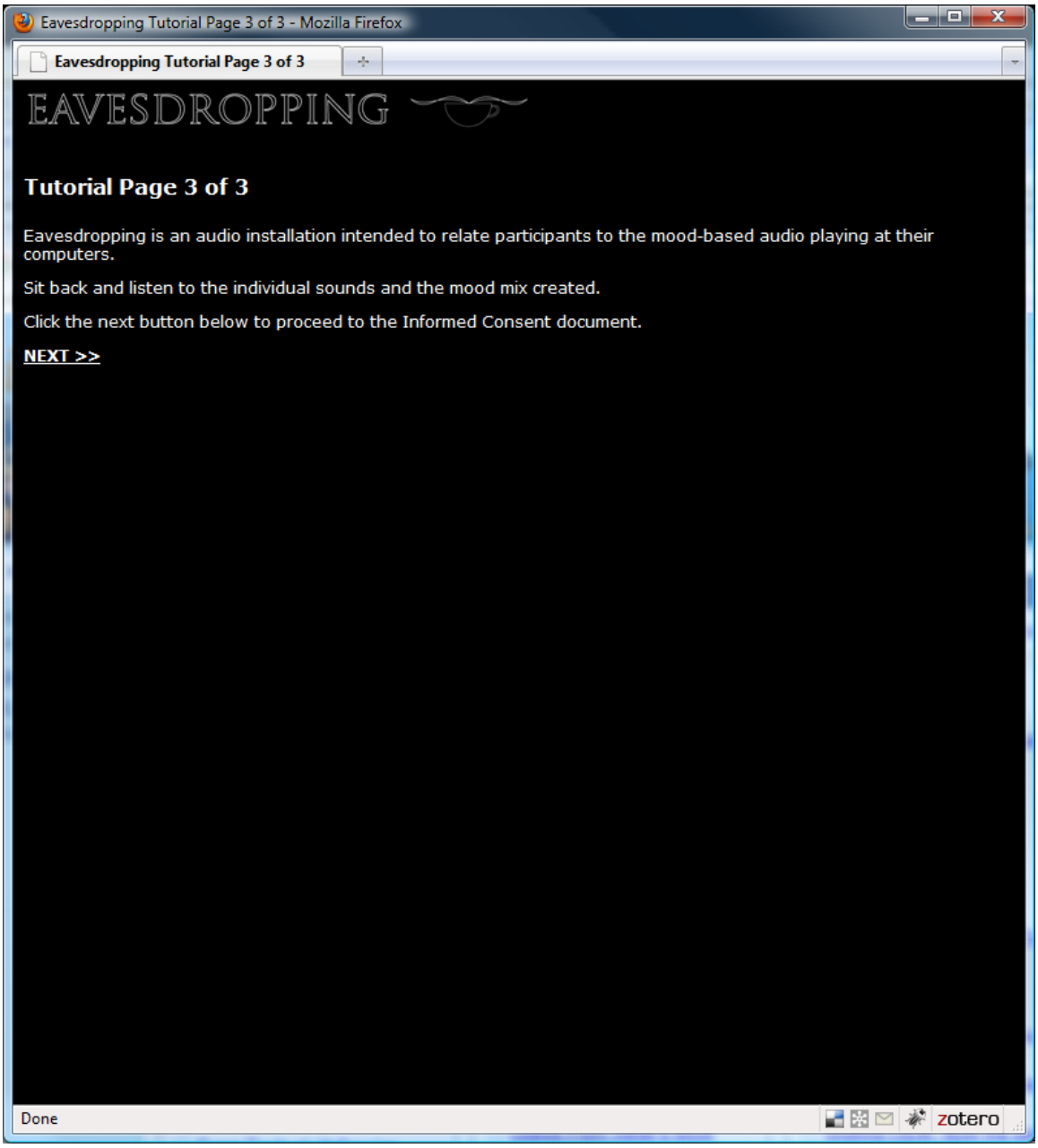
Please click the ‘Start or Join a User Study button’ to begin the tutorial. When you get to the screen that says ‘Wait Here’, please wait for everyone to catch up and we will join the performance together.

Appendix E: Tutorial Pages

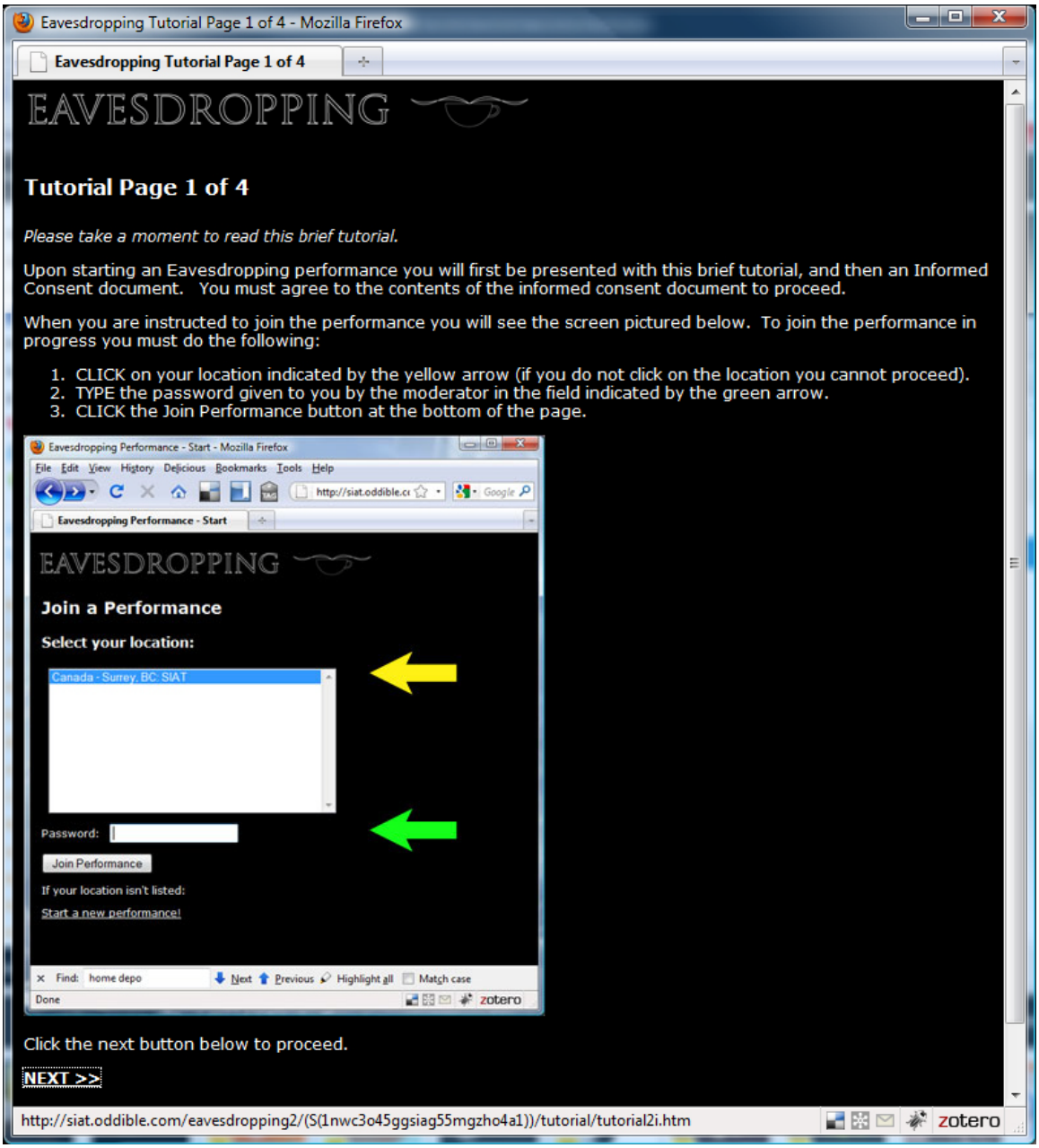
Version 1

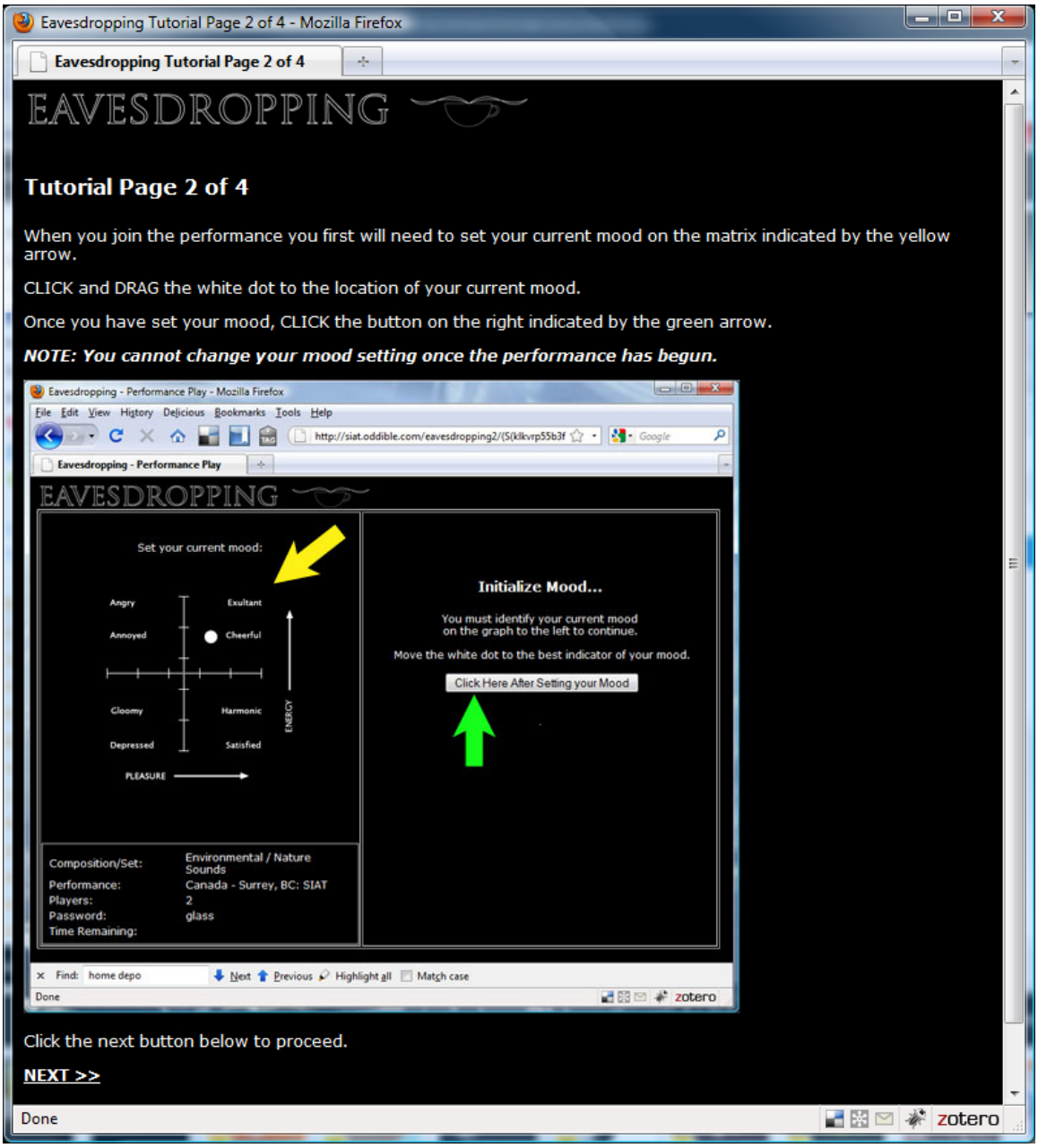


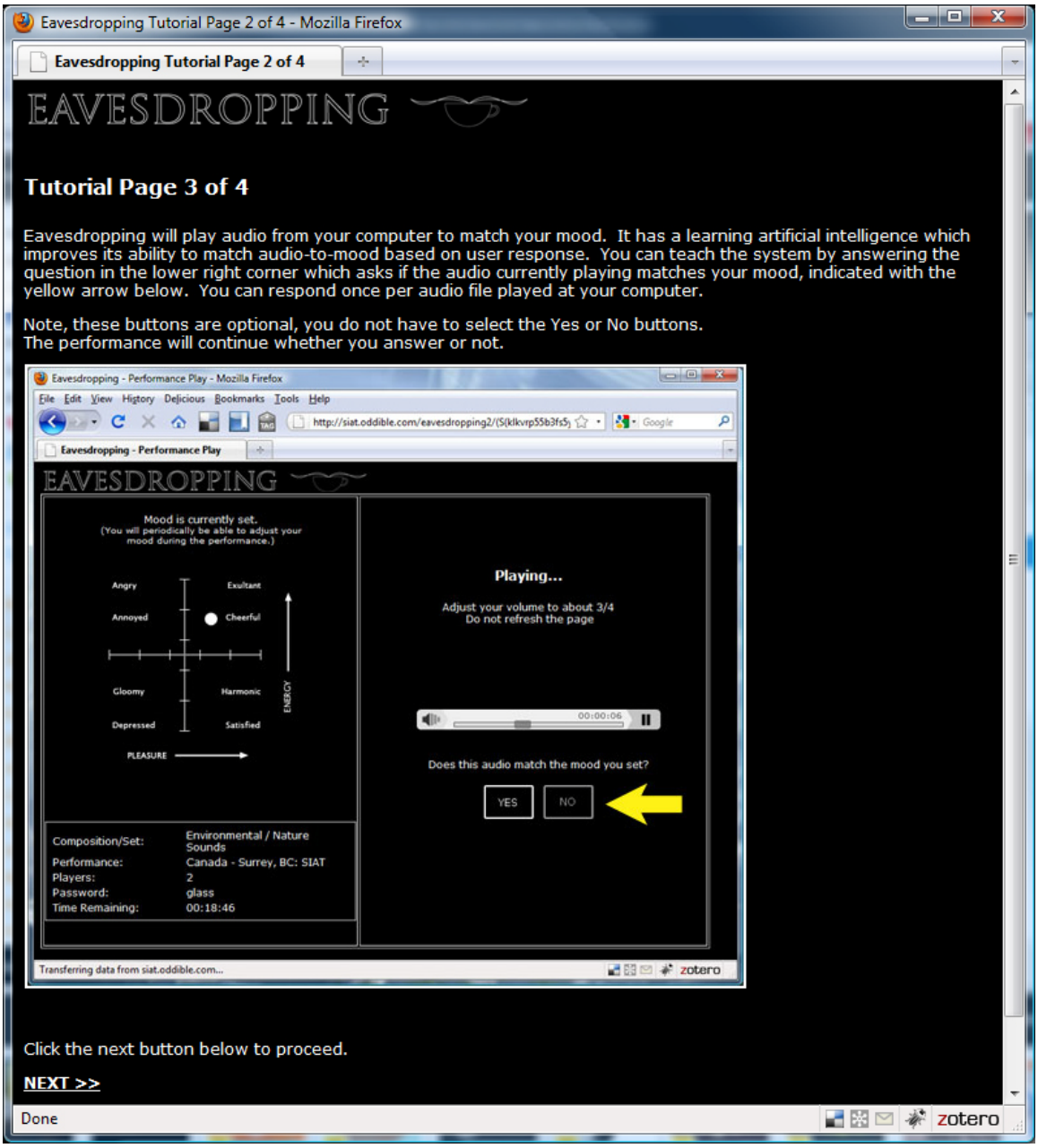

Eavesdropping Tutorial Page 2 of 3 - Mozilla Firefox
Eavesdropping Tutorial Page 2 of 3
EAVESDROPPING
Tutorial Page 2 of 3
Eavesdropping will play a series of audio files from your computer and from all other computers in the room. The composition is based on moods and the audio files are selected by an artificial conductor which is mixing and balancing the audio in the room.
The performance screen will show you how much time remains in the performance in the bottom right.
Eavesdropping - Performance Play - Mozilla Firefox
File Edit View History Delicious Bookmarks Toc
Eavesdropping - Performance Play
EAVESDROPPING
Playing...
Adjust your volume to about 3/4
Do not refresh the page
00:00:05
Composition/Set: Environmental / Nature Sounds
Performance: Canada - Surrey, BC: SIAT
Players: 2
Password: glass
Time Remaining: 00:18:46
Transferring data from siat.oddible.com...
zotero
Click the next button below to proceed.
NEXT >>
Done
zotero

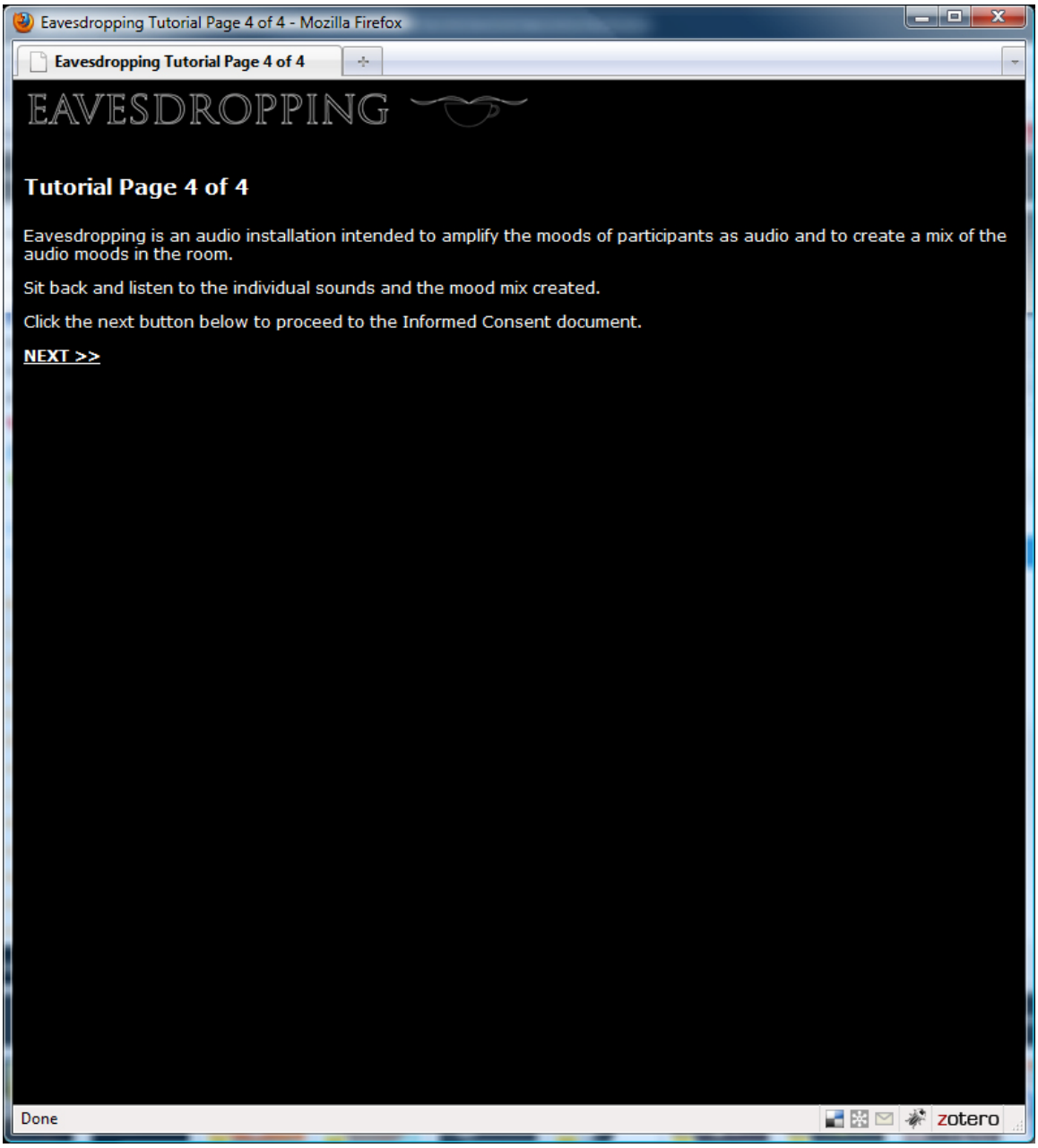

Eavesdropping Tutorial Page 3 of 3 - Mozilla Firefox
Eavesdropping Tutorial Page 3 of 3
EAVESDROPPING
Tutorial Page 3 of 3
Eavesdropping is an audio installation intended to relate participants to the mood-based audio playing at their computers.
Sit back and listen to the individual sounds and the mood mix created.
Click the next button below to proceed to the Informed Consent document.
NEXT >>
Done
zotero

Version 2


Eavesdropping Tutorial Page 1 of 4 - Mozilla Firefox
Eavesdropping Tutorial Page 1 of 4
EAVESDROPPING
Tutorial Page 1 of 4
Please take a moment to read this brief tutorial.
Upon starting an Eavesdropping performance you will first be presented with this brief tutorial, and then an Informed Consent document. You must agree to the contents of the informed consent document to proceed.
When you are instructed to join the performance you will see the screen pictured below. To join the performance in progress you must do the following:
1. CLICK on your location indicated by the yellow arrow (if you do not click on the location you cannot proceed).
2. TYPE the password given to you by the moderator in the field indicated by the green arrow.
3. CLICK the Join Performance button at the bottom of the page.
Eavesdropping Performance - Start - Mozilla Firefox
File Edit View History Delicious Bookmarks Tools Help
Eavesdropping Performance - Start
EAVESDROPPING
Join a Performance
Select your location:
Canada - Surrey, BC: SIAT
Password:
Join Performance
If your location isn't listed:
Start a new performance!
Find: home depo
Next Previous Highlight all Match case
Done
zotero
Click the next button below to proceed.
NEXT >>
http://siat.oddible.com/eavesdropping2/(S(1nwc3o45ggsiag55mgzho4a1))/tutorial/tutorial2i.htm
zotero


Eavesdropping Tutorial Page 2 of 4 - Mozilla Firefox
Eavesdropping Tutorial Page 2 of 4
EAVESDROPPING
Tutorial Page 2 of 4
When you join the performance you first will need to set your current mood on the matrix indicated by the yellow arrow.
CLICK and DRAG the white dot to the location of your current mood.
Once you have set your mood, CLICK the button on the right indicated by the green arrow.
NOTE: You cannot change your mood setting once the performance has begun.
Eavesdropping - Performance Play - Mozilla Firefox
File Edit View History Delicious Bookmarks Tools Help
Eavesdropping - Performance Play
EAVESDROPPING
Set your current mood:
Angry
Exultant
Annoyed
Cheerful
Gloomy
Harmonic
Depressed
Satisfied
ENERGY
PLEASURE
Initialize Mood...
You must identify your current mood on the graph to the left to continue.
Move the white dot to the best indicator of your mood.
Click Here After Setting your Mood
Composition/Set: Environmental / Nature Sounds
Performance: Canada - Surrey, BC: SIAT
Players: 2
Password: glass
Time Remaining:
Find: home depo
Next Previous Highlight all Match case
Done
zotero
Click the next button below to proceed.
NEXT >>
Done
zotero


Eavesdropping Tutorial Page 2 of 4 - Mozilla Firefox
Eavesdropping Tutorial Page 2 of 4
EAVESDROPPING
Tutorial Page 3 of 4
Eavesdropping will play audio from your computer to match your mood. It has a learning artificial intelligence which improves its ability to match audio-to-mood based on user response. You can teach the system by answering the question in the lower right corner which asks if the audio currently playing matches your mood, indicated with the yellow arrow below. You can respond once per audio file played at your computer.
Note, these buttons are optional, you do not have to select the Yes or No buttons.
The performance will continue whether you answer or not.
Eavesdropping - Performance Play - Mozilla Firefox
Eavesdropping - Performance Play
EAVESDROPPING
Mood is currently set.
(You will periodically be able to adjust your mood during the performance.)
Angry
Exultant
Annoyed
Cheerful
Gloomy
Harmonic
Depressed
Satisfied
ENERGY
PLEASURE
Playing...
Adjust your volume to about 3/4
Do not refresh the page
Does this audio match the mood you set?
YES
NO
Composition/Set: Environmental / Nature Sounds
Performance: Canada - Surrey, BC: SIAT
Players: 2
Password: glass
Time Remaining: 00:18:46
Transferring data from siat.oddible.com...
zotero
Click the next button below to proceed.
NEXT >>
Done
zotero


Eavesdropping Tutorial Page 4 of 4 - Mozilla Firefox
Eavesdropping Tutorial Page 4 of 4
EAVESDROPPING
Tutorial Page 4 of 4
Eavesdropping is an audio installation intended to amplify the moods of participants as audio and to create a mix of the audio moods in the room.
Sit back and listen to the individual sounds and the mood mix created.
Click the next button below to proceed to the Informed Consent document.
NEXT >>
Done
zotero

Appendix F: Informed Consent

Informed Consent by Participants

The University and those conducting this research study subscribe to the ethical conduct of research and to the protection at all times of the interests, comfort, and safety of participants. This research is being conducted under permission of the Simon Fraser Research Ethics Board. The chief concern of the Board is for the health, safety and psychological well-being of research participants.

Should you wish to obtain information about your rights as a participant in research, or about the responsibilities of researchers, or if you have any questions, concerns or complaints about the manner in which you were treated in this study, please contact the Director, Office of Research Ethics by email at hal_weinberg@sfu.ca or phone at 778-782-6593.

By choosing to continue with this performance through the completion and submission of the online survey, it will signify that you have read the description of the procedures, whether there are possible risks, and benefits of this research study, that you have received an adequate opportunity to consider the information in the documents describing the study, and that you voluntarily agree to participate in the study. Completing and submitting this online survey signifies that you are either a student of Simon Fraser University, or are 19 years of age or older.

Statement of Confidentiality

Any information that is obtained during this study will be kept confidential to the full extent permitted by the law of British Columbia and Canada. Responses gathered in the online survey will remain confidential through the use of a secure website. All data collected by the system during this performance will be anonymous, including interactions during the performance as well as the questionnaire after the performance. Anonymous data and analysis of data will be used for assessment in a Master's Thesis as well as related publications in the field. Video recordings may be used for documentation purposes and will only be presented in support of this thesis research or for promotion of this art installation. All raw data will be stored on a hard drive and all raw video will be stored on DAT tapes. Both will be stored in a secure area of the SFU SIAT research labs for a period of two years after which the data will be securely deleted and DAT tapes demagnetized.

Purpose and Goals of this Study

This performance and study is designed to assess levels of engagement in a co-located interactive performance achieved through the use of interaction and input of user mood.

What the Participants are Required to Do

You have been invited to participate in a research project which will evaluate audience engagement in a co-located interactive performance. The performance will last five minutes during which audio will be played from the computer in front of you. You will be presented with a web interface which will offer some information or minimal interaction with the performance. At the end of the performance the web interface will present a short questionnaire. You may also participate in an optional interview after the performance.

This performance may be filmed. The video will be used for documentation purposes and will only be presented in support of this thesis research or for promotion of this art installation. Consent to participate in this research also implies consent to be filmed for the above mentioned purposes.

All participation in this research is voluntary and you may withdraw from participation at any time.

The Risks to Participants

There are no risks to participants in this study.

The Benefits to Participants

There are no benefits to participants in this study.

You may register any concern or complaint with the Director of the Office of Research Ethics:

Dr. Hal Weinberg
Director, Office of Research Ethics
8888 University Drive
Simon Fraser University
Burnaby, British Columbia
Canada V5A 1S6
+1 778-782-6593
email: hal_weinberg@sfu.ca

You may obtain results of this study by contacting Jack Stockholm, School of Interactive Arts & Technology, Simon Fraser University, jack_stockholm@sfu.ca.

By checking 'I Consent' below, you are agreeing that you have been informed that the research will be confidential, you understand the risks and contributions of your participation in this study, and you agree to participate. **By consenting to participate, you are confirming that you are either a student of Simon Fraser University or are 19 years of age or older.**

I Consent	I Do Not Consent

REFERENCE LIST

Abeele, M. V., Roe, K., & Pandelaere, M. (2007). Construct validation of the concepts social presence, emotional presence and connectedness and an application of Zajonc's social facilitation theory to social presence research. In *Proceedings of 10th international workshop on presence* (pp. 215-224). Presented at the Presence 2007, Barcelona, Spain.

Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry*, *28*(3), 289-304.

Balkema, A. W., & Slager, H. (Eds.). (2004). *Artistic Research*. Lier en Boog Series of Philosophy of Art and Art Theory (Vol. 18). Amsterdam: Rodopi.

Ballass, J. (1994). Delivery of information through sound. In G. Kramer (Ed.), *Auditory Display : Sonification, Audification, and Auditory Interfaces* (pp. 79-94). Reading, MA: Addison-Wesley.

Barbosa, A. (2003). Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation. *Leonardo Music Journal*, -, 53-59.

Baren, J. V., IJsselsteijn, W. A., Romero, N., Markopoulos, P., & Ruyter, B. D. (2003). Affective Benefits in Communication: The development and field-testing of a new questionnaire measure. In *PRESENCE 2003*.

Batras, D. (2009). Creative Engagement in Interactive Art. Retrieved from ftp://ftp.irisa.fr/local/caps/DEPOTS/BIBLIO2009/Batras_Dimitrios.pdf

Bente, G., Rüggenberg, S., & Krämer, C. (2004). Social Presence and Interpersonal Trust

in Avatar-Based, Collaborative Net- Communications. *Presence 2004*.

Biemans, M. C., van Halteren, A. T., van Dijk, E. M., Rijckenberg, G. J., van Pelt, X., & Poortinga, R. (2008). We'll stay in touch: Intuitive communication means for social connectedness. Retrieved from http://eprints.eemcs.utwente.nl/14270/

Bilda, Z., Edmonds, E., & Candy, L. (2008). Designing for creative engagement. *Design Studies*, *29*(6), 525-540.

Biocca, F., Harms, C., & Gregg, J. (2001). The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity. Presented at the PRESENCE 2001, 4th Annual International Workshop on Presence, Philadelphia, PA.

Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: review and suggested criteria. *Presence: Teleoper. Virtual Environ.*, *12*(5), 456-480.

Brown, E., & Cairns, P. (2004). A grounded investigation of game immersion. In *CHI '04 extended abstracts on Human factors in computing systems* (pp. 1297-1300). Vienna, Austria: ACM.

Bryan-Kinns, N., & Hamilton, F. (2009). Identifying mutual engagement. *Behaviour & Information Technology*. doi:10.1080/01449290903377103

Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., & Allspach, L. (2000). Interactivity in human-computer interaction: a study of credibility, understanding, and influence. *Computers in Human Behavior*, *16*(6), 553-574.

Burgoon, J. K., & Hale, J. L. (1987). Validation and measurement of the fundamental themes of relational communication. *Communication Monographs*, *54*(1), 19.

Bystrom, K., Barfield, W., & Hendrix, C. (1999). A Conceptual Model of the Sense of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, *8*(2), 241-244.

Candy, L., Amitani, S., & Bilda, Z. (2006). Practice-led strategies for interactive art research. *CoDesign: International Journal of CoCreation in Design and the Arts*, *2*(4), 209.

Chen, M., Koldo, B., Cuddihy, E., & Medina, E. (2005). Modelling and measuring engagement in computer games. Presented at the Annual Conference for the Digital Games Research Association (DiGRA), Vancouver, BC, Canada.

Cohen, A. J. (1998). The functions of music in multimedia: A cognitive approach. In S. W. Yi (Ed.), *Proceedings of the Fifth International Conference on Music Perception and Cognition* (pp. 13-20). Seoul, Korea: Seoul University Press. Retrieved from http://www.upei.ca/ musicog/research/docs/funcmusicmultimedia.pdf

Csikszentmihalyi, M. (1991). *Flow: The Psychology of Optimal Experience* (First Edition.). Harper Perennial.

Cumming, N. (2000). *The sonic self : musical subjectivity and signification / Naomi Cumming.*

Davis, B. (1977). *Five Village Soundscapes*. Vancouver: A.R.C. Publications.

Dey, A. K., & Guzman, E. D. (2006). From awareness to connectedness: the design and deployment of presence displays. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 899-908). Montréal, Québec, Canada: ACM.

Dow, S., Mehta, M., Harmon, E., MacIntyre, B., & Mateas, M. (2007). Presence and engagement in an interactive drama. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 1475-1484). San Jose, California, USA: ACM.

Edmonds, E., Muller, L., & Connell, M. (2006). On creative engagement. *Visual Communication*, *5*(3), 307-322.

Eladhari, M., Nieuwdorp, R., & Fridenfalk, M. (2006). The soundtrack of your mind: mind music - adaptive audio for game characters. In *ACE '06: Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology* (pp. 54-61). New York, NY, USA: ACM Press.

Elkins, J. (2004). Theoretical Remarks on Combined Creative and Scholarly PhD Degrees in the Visual Arts. *Journal of Aesthetic Education*, *38*(4), 22-31.

Ermi, L., & Mäyrä, F. (2005). Fundamental Components of the Gameplay Experience: Analysing Immersion. In S. de Castell & J. Jenson (Eds.), *Changing views: Worlds in play.*, Selected papers of the 2005 DiGRA's Second International Conference (pp. 15-27). Retrieved from http://www.digra.org/dl/display_html?chid=http://www.digra.org/dl/db/06276.41516.pdf

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality & Social Psychology*, *69*, 153.

404 Festival. (2008). International Festival of Electronic Art 404. *International Festival of Electronic Art 404*. Retrieved January 29, 2010, from

http://www.404festival.com/eng/autores2008.htm

Field, A. P. (2009). *Discovering Statistics Using SPSS: (and Sex and Drugs and Rock 'n' Roll)* (3rd ed.). Los Angeles: SAGE.

Gaver, W. W. (1991). Sound support for collaboration. In *Proceedings of the second conference on European Conference on Computer-Supported Cooperative Work* (pp. 293-308). Amsterdam, The Netherlands: Kluwer Academic Publishers.

Gilroy, S. W., Cavazza, M., & Benayoun, M. (2009). Using affective trajectories to describe states of flow in interactive art. In *Proceedings of the International Conference on Advances in Computer Enterntainment Technology* (pp. 165-172). Athens, Greece: ACM.

Gonzales, A. L. (2009). Validation of an inventory of social connectedness. In *Proceedings of the ACM 2009 international conference on Supporting group work* (pp. 375-376). Sanibel Island, Florida, USA: ACM.

Grimshaw, M., & Schott, G. (2007). *Situating Gaming as a Sonic Experience: The acoustic ecology of First-Person Shooters*.

Guzman, E. S. D., Yau, M., Gagliano, A., Park, A., & Dey, A. K. (2004). Exploring the design and use of peripheral displays of awareness information. In *CHI '04 extended abstracts on Human factors in computing systems* (pp. 1247-1250). ACM Press.

Harms, C., & Biocca, F. (2004). Internal Consistency and Reliability of the Networked Minds Social Presence Measure. In *Seventh Annual International Workshop*. Presented at the Presence 2004, Valencia: Universidad Politecnica de Valencia.

Hermann, T., & Ritter, H. (2004). Sound and meaning in auditory data display (Vol. 92,

pp. 730-741).

Hevner, K. (1936). Experimental Studies of the Elements of Expression in Music. *The American Journal of Psychology*, *48*(2), 246-268.

Holly, M. A., & Smith, M. M. (2009). *What Is Research in the Visual Arts?: Obsession, Archive, Encounter*. Clark Art Institute.

Hoorn, J., Konijn, E., & Van der Veer, G. (2003). Virtual reality: Do not augment realism, augment relevance. *Upgrade -- Human-Computer Interaction: Overcoming Barriers, IV*(1), 18-26.

IJsselsteijn, W., de Kort, Y., Poels, K., Jugelionis, A., & Bellotti, F. (2007). Characterising and Measuring User Experiences in Digital Games. In *Proceedings of the ACE Conference 2007*. Presented at the ACE, Salzburg: ACM Publishers.

IJsselsteijn, W., Baren, J., Markopoulos, P., Romero, N., & Ruyter, B. (2009). Measuring Affective Benefits and Costs of Mediated Awareness: Development and Validation of the ABC-Questionnaire. In *Awareness Systems* (pp. 473-488).

IJsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: concept, determinants, and measurement. In B. E. Rogowitz & T. N. Pappas (Eds.), *Human Vision and Electronic Imaging V* (Vol. 3959, pp. 520-529). Presented at the Human Vision and Electronic Imaging V, San Jose, CA, USA: SPIE.

Kivy, P. (1989). *Sound Sentiment: An Essay on the Musical Emotions*. Philidelphia, PA: Temple University Press.

Knobloch, S., & Zillmann, D. (2002). Mood Management via the Digital Jukebox. *The Journal of communication*, *52*(2), 351-366.

Knowles, J. G. (2008). *Handbook of the Arts in Qualitative Research: Perspectives,*

Methodologies, Examples, and Issues. Los Angeles: Sage Publications.

Kramer, G. (1994). *Auditory Display : Sonification, Audification, and Auditory Interfaces*. (G. Kramer, Ed.). Reading, MA: Addison-Wesley.

Kreijns, K., Kirschner, P. A., Jochems, W., & van Buuren, H. (2004). Determining Sociability, Social Space, and Social Presence in (A)synchronous Collaborative Groups. *CyberPsychology & Behavior, 7*(2), 155-172.

Kumar, N., & Benbasat, I. (2002). Para-Social Presence and Communication Capabilities of a Web Site: A Theoretical Perspective. *e-Service Journal, 1*(3), 5-24.

Leavy, P. (2009). *Method Meets Art: Arts-Based Research Practice*. New York: Guilford Press.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoper. Virtual Environ., 10*(3), 282-297.

Livingstone, S. R., & Thompson, W. F. (2009). The Emergence of Music from the Theory of Mind. *Musicae Scientiae, 10*(Special Issue), 83–115.

Livingstone, S. R., Brown, A. R., & Muhlberger, R. (2005). Influencing the Perceived Emotions of Music with Intent. In T. Innocent (Ed.), . CEMA Press. Retrieved from http://eprints.qut.edu.au/9526/

Livingstone, S. R., Mühlberger, R., Brown, A. R., & Loch, A. (2007). Controlling musical emotionality: an affective computational architecture for influencing musical emotions. *Digital Creativity, 18*(1), 43.

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence.

Journal of Computer-Mediated Communication, *3*(2).

Mäkelä, M., & Routarinne, S. (2006). *The art of research: research practices in art and design*. University of Art and Design.

Mania, K., & Chalmers, A. (2001). The Effects of Levels of Immersion on Memory and Presence in Virtual Environments: A Reality Centered Approach. *CyberPsychology Behavior*, *4*(2), 247-264.

Manovich, L. (2002). The Anti-Sublime in New Media. *Chair et Metal / Metal and Flesh*.

Markopoulos, P., Romero, N., Baren, J. V., IJsselsteijn, W., Ruyter, B. D., & Farshchian, B. (2004). Keeping in touch with the family: home and away with the ASTRA awareness system. In *CHI '04 extended abstracts on Human factors in computing systems* (pp. 1351-1354). ACM Press.

McMahan, A. (2003). Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Games. In M. J. P. Wolf & B. Perron (Eds.), *The Video Game Theory Reader* (1st ed., pp. 67-86). Routledge.

Morrison, A. J., Mitchell, P., & Brereton, M. (2007). The lens of ludic engagement: evaluating participation in interactive art installations. In *Proceedings of the 15th international conference on Multimedia* (pp. 509-512). Augsburg, Germany: ACM.

O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, *61*(1), 50-69.

Potts, A. (2009). The Artwork, the Archive. and the Long Moment. In M. A. Holly & M.

M. Smith (Eds.), *What Is Research in the Visual Arts?: Obsession, Archive, Encounter* (pp. 119-137). Clark Art Institute.

Rhizome.org. (2008). Rhizome 2008 Commissions. *Rhizome 2008 Commissions*. Retrieved January 29, 2010, from http://rhizome.org/commissions/2008/

Rice, R. E. (1992). Task Analyzability, Use of New Media, and Effectiveness: A Multi-Site Exploration of Media Richness. *Organization Science*, *3*(4), 475-500.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality & Social Psychology*, *39*(6), 1161.

Russell, S., & Norvig, P. (1995). *Artificial intelligence: a modern approach*. New Jersey: Prentice-Hall.

Schubert, E. (2003). Update of the Hevner adjective checklist. *Perceptual and Motor Skills*, *96*(3 Pt 2), 1117-1122.

Seo, J., & Gromala, D. (2007). Touching light: A new framework for immersion in artistic environments. *Technoetic Arts: a Journal of Speculative Research*, *5*(1), 3-14.

Sgouros, N. M. (2000). Detection, analysis and rendering of audience reactions in distributed multimedia performance. In *Proceedings of the eighth ACM international conference on Multimedia* (pp. 195-200). Marina del Rey, California, United States: ACM.

Short, J. (1976). *The Social Psychology of Telecommunications*. London: Wiley.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE) - Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *6*(6), 603.

Stockholm, J. (2008). Eavesdropping: Network Mediated Performance in Social Space. *Leonardo Music Journal*, -, 55-58.

Stockholm, J., & Pasquier, P. (2008). Eavesdropping: audience interaction in networked audio performance. In *Proceeding of the 16th ACM international conference on Multimedia* (pp. 559-568). Vancouver, British Columbia, Canada: ACM.

Stockholm, J., & Pasquier, P. (2009). Reinforcement Learning of Listener Response for Mood Classification of Audio. In *Computational Science and Engineering, IEEE International Conference on* (Vol. 4, pp. 849-853). Los Alamitos, CA, USA: IEEE Computer Society.

Sullivan, G. (2010). *Art Practice as Research: Inquiry in Visual Arts* (2nd ed.). Thousand Oaks, Calif: Sage Publications.

Sutton, R. S. (1998). *Reinforcement Learning: An Introduction* (Vol. 9).

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *Comput. Entertain.*, *3*(3), 3-3.

Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford University Press.

Truax, B. (1999). *Handbook of acoustic ecology*. London: Cambridge Street Publishing.

Truax, B. (2001). *Acoustic communication*. London: Ablex Publishing.

Vastfjall, D. (2003). The Subjective Sense of Presence, Emotion Recognition, and Experienced Emotions in Auditory Virtual Environments. *CyberPsychology & Behavior*, *6*(2), 181-188.

Webster, J., & Ho, H. (1997). Audience engagement in multimedia presentations. *SIGMIS Database*, *28*(2), 63-77.

Weinberg, G. (2002). The aesthetics, history and future challenges of interconnected

music networks. In *Proceedings of the International Computer Music Association Conference* (pp. 349-356).

Weinberg, G. (2005). Interconnected Musical Networks: Toward a Theoretical Framework. *Computer Music Journal*, *29*(2), 23-39.

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225-240.