ENHANCING NAVIGATION USING AUDITORY FEEDBACK: A CASE STUDY OF A HIERARCHICAL INFORMATION VISUALIZATION SYSTEM

Rafa Absar

School of Information Studies
McGill University, Montréal

April 2012

A Thesis submitted to McGill University
in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

ABSTRACT

Interaction with information systems today mostly consists of a user's unimodal interaction with a text-based or visual information system. However, human-computer interaction studies have illustrated that information can be successfully conveyed through different sensory modalities. This research focuses on the enhancement of the user experience using auditory feedback for the specific case of a 3D-visualized hierarchical information system, by representing some of the structural and navigational cues using nonspeech sounds. It is hypothesized that engaging the auditory modality may aid in navigation tasks, improve users' affective reactions and consequently enhance the overall user experience.

The research involves two studies. In the first study, a user-centred semiotic sound design methodology is used, based on a methodology originally used on visually-impaired users. Three panels of end-users are employed to design the required nonspeech sounds. Based on the results of this study, recommendations are made for extending the sound design method to novel interfaces and sighted users.

The second study is a controlled experiment that compares user experience with the visualization system and with the auditory-feedback enhanced system. The goal is to evaluate the effect of the auditory feedback on user experience. This effect is measured using a measurement model which draws on concepts derived from three conceptual frameworks, based in Human-Computer Interaction (HCI), Information Science (IS) and auditory interface studies. A combination of measures is examined, including utilitarian variables such as time taken on hierarchy navigation and information retrieval tasks, and accuracy of the answers. Hedonic variables, which influence the affective reactions, were also examined. These include preference, perceived ease of use, usefulness and ease of learning, and user engagement and satisfaction.

We observed that 79% of the participants preferred the audio-visual system to the visual-only system. The audio-visual system was also perceived as easier to use and received higher ratings in terms of aesthetic appeal and perceived usability, which are attributes of user engagement. Furthermore, the audio-visual system was often perceived as being faster and more engaging even though no significant differences were observed in terms of utilitarian variables of task times and accuracy. Findings suggest hedonic variables play an important role in enhancing user experience when interacting with information systems. This research contributes to the field of information science by showcasing that designing multimodal information systems with a focus on the user has the potential to improve user experience. Our findings also provides evidence that utilitarian variables need not be the principle focus of user experience enhancement in information navigation and retrieval tasks, as preference appears to be linked to hedonic variables.

RÉSUMÉ

L'interaction avec les systèmes d'information consiste essentiellement, à l'heure actuelle, en une interaction unimodale, où l'utilisateur se sert d'un système d'information textuel ou graphique. Cependant, les études en interaction humain-machine démontrent que le recours à différentes modalités sensorielles permet de communiquer de l'information de façon enrichissante. La présente recherche porte sur le rehaussement de l'expérience utilisateur au moyen de rétroaction auditive dans un système de visualisation de l'information hiérarchique en trois dimensions, en représentant certains éléments structurels et navigationnels à l'aide de sons non oraux. L'hypothèse de base stipule que la modalité auditive pourrait alléger les tâches de navigation, améliorer les réactions affectives de l'utilisateur et ainsi rehausser l'expérience globale d'utilisation.

La recherche comprend deux études. La première étude comporte une méthodologie centrée sur l'utilisateur, basée sur une méthodologie utilisée à l'origine avec des utilisateurs non voyants, pour concevoir des sons ayant une valeur sémiotique. Trois panels d'utilisateurs finaux ont participé à la conception des sons non oraux à utiliser. Les résultats de cette étude permettent de formuler des recommandations afin d'étendre la méthode de conception des sons aux interfaces novatrices et aux utilisateurs voyants.

La seconde étude est un essai contrôlé comparant l'expérience d'utilisation du système de visualisation avec celle du système enrichi avec rétroaction auditive. L'objectif est d'évaluer l'effet de la rétroaction auditive sur l'expérience utilisateur. Cet effet est mesuré à l'aide d'un modéle de mesure fondé sur des concepts provenant de trois cadres théoriques: l'interaction humain-machine, les sciences de l'information et l'étude d'interfaces auditives. Une combinaison de mesures est effectuée pour les variables utilitaires telles que le temps requis pour la navigation hiérarchique et pour les tâches de repérage de l'information, et l'exactitude de la réponse. Les variables hédoniques, qui influent sur les réactions affectives, ont également été examinées. Celles-ci comprennent la préférence, la facilité d'utilisation perçue, l'utilité perçue et la facilité d'apprentissage perçue ainsi que l'engagement et la satisfaction de l'utilisateur.

Les résultats démontrent que plus de 79% des participants préfèrent le système audiovisuel au système uniquement visuel. Le système audiovisuel est également perçu comme étant plus facile à utiliser et a obtenu des scores plus élevés relativement à l'apparence (l'esthétisme) et à la convivialité perçue, deux attributs de l'engagement de l'utilisateur. Par ailleurs, le système audiovisuel est souvent perçu comme plus rapide et plus engageant, même si aucune différence significative n'a été notée pour les variables utilitaires comme le temps requis pour effectuer les tâches et l'exactitude. Les résultats suggèrent que les variables hédoniques jouent un rôle important dans le rehaussement de l'expérience utilisateur lors d'interactions avec des systèmes d'information. Cette recherche contribue au domaine des sciences de l'information en démontrant que la conception de systèmes d'information multimodaux centrés sur l'utilisateur a le potentiel de rehausser l'expérience utilisateur. Les résultats démontrent également que les variables utilitaires n'ont pas á être le

principal indicateur du rehaussement de l'expérience utilisateur pour les tâches de navigation et de repérage de l'information, puisque la préférence semble être liée aux variables hédoniques.

ACKNOWLEDGEMENTS

I would like to thank all my professors, colleagues, friends and family without whose support and contributions this research would not have been possible. First and foremost, I would like to express my deep gratitude to my supervisor, Catherine Guastavino. Her knowledge, guidance and support every step of the way has been invaluable and has kept me motivated throughout my studies. I would like to thank members of my doctoral committee, Professors Jamshid Beheshti, Andrew Large and Marcelo Wanderley for providing me with valuable feedback and encouragement during the research proposal and dissertation writing stages, and France Bouthillier for the support she provided over the past five years. I would also like to thank Professor Sue Whitesides for encouraging me to apply to the School of Information Studies (SIS) in the first place and providing unquestionable support at every step proceeding that.

I would also like to convey my gratitude to the administrative assistants, Susann Allnutt, Kathryn Hubbard and Ancy Joseph. They have always been more than gracious with their time and assistance, and enabled my entire time spent at the School to proceed as smoothly as I could have ever hoped.

I am very grateful to Charles-Antoine Julien for going above and beyond what is required to help me familiarize myself with the code of the visualization system, and providing support and helpful tips whenever I needed it, no matter how busy his own schedule was. I would like to express my thanks to our Multimodal Interaction Lab (MIL) members, who have become like an extended family to me. Interaction with them in our many lab meetings and feedback from them during research presentations have been a most valuable and enjoyable form of exchange. A special thanks goes out to Jonathan Dorey for helping me with French translations whenever I needed it, including the French Abstract for this dissertation, even if he was swamped with his own deadlines. In addition, I would like to thank all my participants and panellists, students of McGill University, for their time and enthusiasm during the experiments.

This research has been funded by Fonds de Recherche Nature et Technologies Quebec (FQRNT) and the McGill Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT). SIS and CIRMMT have also generously provided travel grants whenever I needed to travel to conferences. I am also grateful for being able to use the MIL lab facilities for the experiments, which were supported by Catherine Guastavino's Canada Foundation for Innovation (CFI) grant.

Last but not least, I would like to thank my parents and parents-in-law for their unwavering moral support during these years. One of my greatest motivations was to fulfill my father's dream of having his child receive the high honor of a doctoral degree. And I can never begin to express my gratitude to my husband, Junaed, whose support, love and unhesitating help in taking care of everything when I did not have the time, helped me keep my focus and motivation. This dissertation would not have been possible without him, so this is for him and my little Nadyne. She reminds me that the past five-year journey has been rewarding not only for the diversity of knowledge and experiences I have received, but also because she was another precious gift that came during this period.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

"..to you they are birds, to me they are voices in the forest."

(Feld, 1990, p. 44)

We use a combination of vision, audition, olfaction and haptic perception in our everyday lives to gather information from our surroundings and interact with the world. Such a combination could be a powerful tool for interaction with human-computer and human-information interfaces. However, common interfaces today are mostly visually-oriented; very little information is presented via the other modalities (Brewster, 2003). Hence, the myriad advantages that can be offered by integrating the visual medium with other presentation media are not yet fully realized.

Although there are numerous areas in the computer interface that can benefit from the interaction with multiple modalities to enhance the information provided, one such important area is the navigation of an information space. *Navigation* is a term used to describe the process users employ to access any of several objects available in a computer interface (Isa, Ogden, Wolfe, & Korenshtein, 1986). Hence, the term navigation refers to the means by which a user can find his way through any interface or system. It involves tasks such as traversing through an interface to find an object, interacting through a multi-level menu, browsing through a file system or searching an information system during the process of information seeking. In the context of information spaces specifically, *navigation* is the "means by which a user can describe movement between pieces of information" (Dourish & Chalmers, 1994).

Such tasks are typically carried out in the visual domain. The aim of this research is to design and evaluate the addition of auditory feedback to enhance the user experience when carrying out such tasks in an information space. The next few subsections describe the problems in navigating information spaces, the advantages of using multiple modalities in information presentation, the definition of the problem, the specific research questions and finally, an outline of the dissertation.

1.1. Navigation and information seeking

Effective navigation is a process that is of great interest in recent times due to the explosion of information available today. *Effectiveness* can be defined as "the accuracy and completeness of users' tasks while using a system" ("Ergonomic Requirements for office with visual display terminals - Guidance on usability", 1998). It is necessary to effectively navigate through a vast amount of information, be it when trying to search or find a document in a file system, browse through the internet, or search for specific data in an information system.

However, this is not always an easy task because of the size, structure and complexity of most information systems (Ahuja & Webster, 2006; Yu & Roh, 2002). The effectiveness of navigation effects are determined by the structure of the information presented, the link mechanisms and the design of the structural cues (Shneiderman, 1998). If the structural layout and cues of the system do not allow users to create an accurate mental model of the system, it can lead to several navigation problems.

Benyon and McCall (1999) have identified three main navigational problems encountered by users in electronic information spaces. They are: not knowing where to go next, knowing where to go but not how to get there and not knowing the current position relative to the overall structure, or in other words, *navigational disorientation*.

Disorientation is one of the most common navigation problems, especially in hypertext systems like the World Wide Web. Disorientation is defined as the tendency to lose the sense of location and direction in a non-linear document system (Conklin, 1987). This may be caused by the user having difficulty understanding the non-linear structure or the relationship between the system and interface (Yu & Roh, 2002). Since there are so many more options to be selected in most electronic systems than in traditional text, this can cause a user to lose the sense of direction and location. Any information system with a poor navigational structure can leave users feeling disoriented and frustrated (Ahuja & Webster, 2006) and lead them to feel that they lack control over the interaction (Rozell & Gardner, 2000).

Another navigation problem that can arise in complex information systems is the phenomenon of cognitive overload. This is defined as the additional effort and concentration necessary to maintain several tasks or trails simultaneously (Conklin, 1987). Navigating the system requires the users to create a mental model of the navigation map or path of the links (Yu & Roh, 2002). This extra mental effort, which is not related to the content of the documents being searched, means that it could lead to user cognitive overload. And together, the disorientation problem and the increased cognitive load can lead users to being lost in hyperspace (Sand, 1996). Hence, this problem of getting "lost" in virtual information spaces has been noted as a severe difficulty in navigation by experts in the field of information science (Dillon, McKnight, & Richardson, 1990).

These navigation problems may be solved by designing structural cues which allow users to retain a mental model of the structure of the information system. Since the addition of only visual cues may lead to higher cognitive and visual overload, one solution could be the design and addition of auditory cues to complement and reinforce the visual structural cues.

1.2. Multimodal interfaces

Past studies have shown that information spread over multiple modalities rather than a single one helps to minimize users' cognitive load in various types of environments (Oviatt, Coulston, & Lunsford, 2004). Oviatt et al. showed that multimodal interface users spontaneously respond to higher cognitive loads by shifting to multiple modalities as load increases with task difficulty and communicative complexity, such as in mobile interfaces and map systems. It has also been found that users prefer to interact multimodally when given a choice, especially in spatial tasks such as map navigation (Oviatt, 1997).

In educational literature, it was found that a dual-mode presentation format using visual diagrams and audiotapes supported the working memory more than a unimodal visual format (Mousavi, Low, & Sweller, 1995). This has been shown for other types of tasks and presentation materials as well, including multimedia animations and presentations (Andres & Petersen, 2002; Mayer & Moreno, 1998; Tindall-Ford, Chandler, & Sweller, 1997), confirming that multimodal presentation formats help support the working memory in a way that is advantageous for classroom instruction. It has also been shown to have advantages when presenting more difficult instructional materials compared to simpler ones (Tindall-Ford et al., 1997). Specifically in the area of audio-visual processing, empirical results compared with unimodal perception have shown that combined audio-visual stimuli provide advantages in perceptual discrimination, i.e. bimodal target stimuli are responded to much faster and are identified more accurately than the unimodal target stimuli (Calvert, Spence, & Stein, 2004; Teder-Slejrvi, McDonald, Di Russo, & Hillyard, 2002).

In addition to literature supporting multimodality due to the positive effects on working memory and cognitive load, sensory integration is also believed to influence usability and enjoyment derived by users of a technology (Quesenbery, 2003). Sensory appeal through the use of multiple channels such as text, graphics and sound is believed to increase stimulation and lead to pleasurable experiences (Jennings, 2000; Laarni, Ravaja, Kallinen, & Saari, 2004; Laurel, 1993).

Hence, there is converging evidence that multimodal or dual-modal interfaces are effective in minimizing users' cognitive load, especially when performing complex tasks (Oviatt et al., 2004). The integration of sensory modalities such as audio and video may also contribute to the enhancement of affective reactions. As such, there is a need for further research into the effects of dual-sensory integration in complex tasks such as information retrieval.

Chapter 2 gives a more detailed review of the literature relevant to information visualization and sonification; it presents a comprehensive review of how auditory feedback has been used to enhance the effectiveness of various types of interfaces and how the effects are evaluated.

1.3. Problem definition

The focus of this project is on auditory enhancement of navigation tasks in structured information interfaces, and how sound can be used to help users understand and better respond to such a system. Auditory feedback is just one component in a wide range of multimedia-based interactions, and can be used to reinforce or complement visual stimuli presented to the user. This research includes exploring how the two modalities, visual and audio information, can be integrated in an effective and synchronized way, specifically for navigation tasks in an information system, and studying usability issues to understand the actual effects of complex interactions with such data sources on user experience and performance.

The scope of the study will be limited by using the specific case study of a hierarchical information system. Such a system presents a special and challenging case of navigation, as it explicitly requires the formation of an internal model of the hierarchical structure, as discussed in Section 2.1.

For this purpose, one general research question is posed and to address this, it has been broken down into three sub-questions.

1.3.1. Research question.

What is the effect of adding auditory feedback to a visual information system on user experience for information navigation and seeking tasks?

Before addressing the above research question, a few terms need to be properly defined in the current context.

- *User experience*:
  - Conceptual definition: The entire set of affects that is elicited by the interaction between a user and a product, including the degree to which all our senses are gratified (aesthetic experience), the meanings we attach to the product (experience of meaning), and the feelings and emotions that are elicited (emotional experience) (Desmet & Hekkert, 2007).
  - Operational definition: The user experience in the current context will be evaluated by measuring the effects on performance, as well as affective reactions and subjective preferences of users.
- *Enhancement*:

– Conceptual definition: an improvement that makes something more agreeable (as defined by WordNet (r) 2.0).
– Operational definition: In the current context, an enhancement will be seen as any improvement in any aspect of the performance or affective reactions of users, or in other words, in the user experience as defined above.

To address the stated research question, the following more specific sub-questions are posed.

(i) What kind of non-speech auditory cues can be designed to convey information in a hierarchical information system using a participatory design method?
(ii) How can the effect of auditory feedback be evaluated for information navigation tasks in such a system, so as to examine the effect on the entire user experience, which combines user performance, preference and affective reactions?
(iii) What are the differences between a visual-only hierarchical information system and an audio-visual one in terms of the user experience?

1.4. Outline of the dissertation

This thesis describes the previous literature and the experiments geared to study the above aims. Specifically, the thesis describes the investigation of the design, addition, and effect of auditory enhancement in visual information-navigation tasks on user experience. In Chapter 2, previous work in the fields related to information visualization, auditory feedback and sonification is discussed. A part of the review on auditory feedback has been published in (Absar & Guastavino, 2008). Chapter 3 describes the theoretical framework of navigation and user experience on which the research is based. Chapter 4 describes the sound design methodology that has been used here and the resulting sounds designed. The results of this methodology has been published in (Absar & Guastavino, 2011). Chapter 5 describes the design and methodology for the controlled comparative experiment of the audio-visual and visual-only systems. Chapter 6 presents the analysis of the data and the results. Chapter 7 concludes the thesis and discusses future research directions.

CHAPTER 2

VISUALIZATIONS AND AUDITORY DISPLAYS

2.1. Information Visualization and Sonification

Information visualization is the communication of abstract data through the use of interactive visual interfaces (Keim, Mansmann, Schneidewind, & Ziegler, 2006). Accessing and managing large information spaces is becoming more and more difficult due to the current exponential growth of information. Information visualization addresses this problem by providing compact graphical and visual presentations for interactively manipulating large numbers of items (Plaisant, 2005). It enables the viewer to gain knowledge about the internal structure of the data and causal relationships in it so that they can create a mental model of the structure of the information.

One of the techniques for visualizing hierarchical information structures is by using *Cone Trees* (Robertson, Mackinlay, & Card, 1991). The hierarchical structure is displayed in 3D to maximize the use of screen space available effectively and give the viewer a view of the whole structure. Interactive animation is also used to help in information access and managing tasks. An example of a cone tree is shown in Figure 2.1.

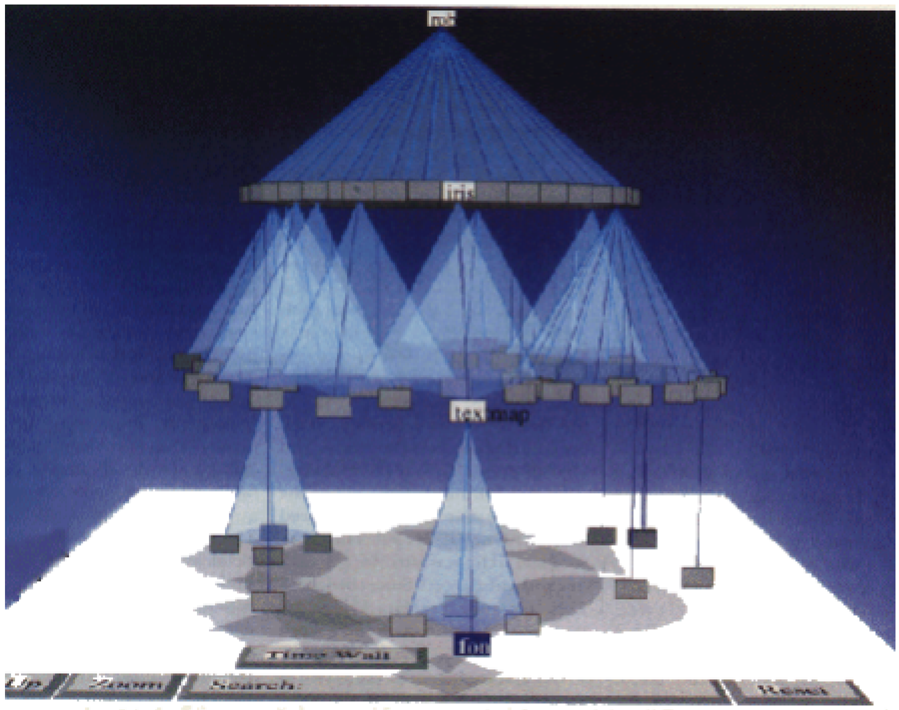


FIGURE 2.1. A cone tree (Robertson, Mackinlay & Card, 1991)

Hearst and Karadi (Hearst & Karadi, 1997) extended this idea of a cone tree to what they referred to as a *Cat-a-cone*. This is an interface where the searching and browsing of large category hierarchies is integrated with their associated text collections. There are two key components of this interface: one is the display of the representations of the categories and the retrieved documents while the other is the simultaneous display of multiple selected categories together with their hierarchical context. The information visualization system that is used in the current research for auditory enhancement and evaluation (described in Section 5.2.1) is based on this concept of cat-a-cones (Hearst & Karadi, 1997).

The idea of using visualization is that some of the user's cognitive load from gathering information from large and diverse sources (such as from text-based interfaces) can be reduced by shifting some of the load to the human perceptual system by visual representations (Robertson et al., 1991). An extension to this idea is to not only reduce cognitive load by shifting to the visual system, but by also making use of other parts of the perceptual system such as the auditory modality.

Sonification is the use of nonspeech audio to convey information. More specifically, sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation (Kramer et al., 1999). The effectiveness of information visualization can be reduced by several adverse elements such as visual overload, lack of screen space, and visual obscurity. One of the ways these issues may be effectively addressed is by integrating the visual medium with auditory cues, which can display both complementary and supplementary information to the visuals.

Most of the research at the intersection of visualization and sonification have looked into effective ways of sonifying scientific data visualizations, visualizations of numerical information, such as graphs and tables, and other such auditory displays of complex data (Barrass & Kramer, 1999; Minghim & Forrest, 1995; Salvador, Minghim, & Pacheco, 1998). For example, Rabenhorst, Farrell, Jameson, Linton, and Mandelman (1990) looked into the combination of visualization with sonification for multi-dimensional data which computed electron density, hole density, and potential throughout the volume of a three-dimensional semiconductor. L. Brown, Brewster, Burton, Riedel, and Ramloll (2003) described design guidelines for audio presentation of graphs and tables for visually-impaired users. Ramloll, Brewster, Yu, and Riedel (2001) designed such a tool for improving access to numerical information in two-dimensional tables by using audio produced by a MIDI synthesizer. Flowers, Buhman, and Turnage (2005) argue that though most of the time these sonifications are developed primarily for the visually-impaired, it is possible to effectively display basic properties of simple functions and data samples using simple auditory graphs involving patterns of pitch variation over time. As such, they have found that these displays can be easily comprehended by sighted users with minimal practice. Hence, they encourage further exploration of such data representation by sound, which may lead to a variety of useful creative developments in data display technology.

Another area in which sonification has been worked on for augmenting visualizations is that of sonifying maps. H. Zhao, Smith, Norman, Plaisant, and Shneiderman (2005) developed two sonifications for an enhanced table and a spatial choropleth map, which is a map that shows differences by using shading or colors. Their investigations into using interactive sonification to present the geographical distribution pattern of statistical data to vision impaired users showed that users were able to perceive aural distinctions on both familiar and unknown maps, and learn new map geography (H. Zhao, Plaisant, & Shneiderman, 2005).

One interesting map sonification research project is the BATS (Blind Audio Tactile mapping System) project at the University of North Carolina (Parente & Bishop, 2003), in which they use simple spatial audio to sonify maps. Real world sounds are played as the user moves the cursor over the map, such as car noises over cities and birds or crashing waves over forests and beaches. The sound becomes louder as the user gets close to the source of the sound, allowing blind users to explore the map and locate objects with interactivity and immediate feedback.

The most relevant research project found for auditory displays designed for information visualizations though is the "loudSpire" system, which is a auditory display schema designed for "Spire", a system for visualizing large amounts of document-based information (LoPresti & Harris, 1996; LoPresti, 1997). Three system layers were sonified: system events were represented by electronic tones associated with computers; data objects with real world sounds such as the rustling of pages when scanning a document; and domain attributes with orchestral music or tonal sounds. The Spire system, however, did not follow a hierarchical structure; it used the visual paradigm of a galaxy, with documents placed so that those with similar themes were closer together. Any formal user evaluations of the audio-visual system have not been conducted or published.

Hence the problem of sonifying structured information in a visualization system, such as a hierarchical information system, for sighted users, and formally evaluating it, is fairly novel and merits further research. Extensive research has however been carried out till date in the field of human-computer interaction studies on nonspeech auditory displays in various contexts and environments, from which this research can draw on. Towards this end, the next section provides a review of previous research on the subject.

2.2. Nonspeech Auditory Feedback in User Interfaces

> The neglect of sound in mainstream interface is striking. This can be partially accounted for by the lack of understanding about how sound is processed in the real world, and lack of inspiration about how to use sound innovatively.
>
> (Macaulay, Benyon, & Crerar, 1998, p. 161)

Although there have been significant strides in the field of auditory display since the above quote was made more than a decade ago, it can still be stated as a relevant fact today; computer interfaces are still mostly visual. However, while visual interfaces can provide detailed or high-resolution information, it requires direct focus on the area providing information, whereas auditory feedback can provide more general or background information even outside the periphery of the field of view (Brewster, 2003). Presenting all the possible information that may be required by a user of a system only through the visual medium may lead to visual overload (M. Brown, Newsome, & Glinert, 1989) and may also lead to some information being missed, if the user is focused elsewhere. M. Brown et al. (1989) suggested replacing some of the cues traditionally presented in the visual modality with auditory cues, so as to reduce some of the visual workload. For example, auditory cues indicating the location of objects have been found to improve visual search speed and accuracy, in visual target finding tasks (M. Brown et al., 1989; Kieffer & Carbonell, 2006).

Nonspeech sound used for auditory feedback can be divided into two main categories. Sounds that can be easily attributed to objects or events generating sounds in everyday situations are referred to as *auditory icons* (Gaver, 1986). Abstract sounds, typically synthetic and less identifiable, are referred to as *earcons* (Blattner, Sumikawa, & Greenberg, 1989). Both types of sounds are described in the remainder of this chapter, which reviews previous literature on the use of these sounds in various tasks and applications. Following this, studies addressing audio-visual integration and evaluations of both types of auditory feedback and how they affect task performance or user experience, are discussed. Descriptions of several tools and applications that have used non-speech auditory feedback are also provided. A published version of this review can also be referred to in (Absar & Guastavino, 2008).

2.2.1. Auditory icons. The concept of auditory icons was introduced by Gaver (1986) as emulations or caricatures of naturally occurring sounds in everyday life. Gaver (1993) suggests that humans perceive everyday sounds in terms of the sources, materials and actions that made them, rather than the individual sound attributes such as pitch and timbre. This view has been validated by several studies in everyday sounds perception (Ballas, 1993; Guastavino, 2006, 2007).

Hence, an auditory icon is a sound that is intended to provide information about an event or object in the interface by representing the desired data using properties of the sound's source, rather than properties of the sound itself (Gaver, 1986). Another important property of such everyday sounds is that they convey information about the sound source (e.g. size and material), and the interaction (e.g. force applied or action). These features can be useful in providing multi-faceted information in human-computer interfaces (Brewster, 2003).

One of the earliest applications using auditory icons is the SonicFinder (Gaver, 1989). Gaver used nonspeech real-life sounds as auditory feedback to interface events that can be intuitively mapped to the respective sounds as analogies to the actions performed (directly or metaphorically). For instance, selecting a file was mapped to the sound of an object being tapped, with the type of object indicated by the object material, and size of the file represented by the size of the struck object. An example of a form of auditory icon used as additional feedback to visual events in desktop interfaces today are sounds like the metallic crunching that accompanies the action of placing an object in trash (deletion).

2.2.2. Earcons. Earcons are abstract sounds in the user-computer interface that provide feedback to the user (Blattner et al., 1989). The sounds used are synthetic combinations of parameters of sound, such as the timber, pitch and rhythm. Manipulating these parameters allow the representations of hierarchies to create both simple and complex forms of auditory feedback to encode information at the interface level.

While auditory icons that represent real world sounds may have the advantage of being easier to remember and learn as they sound more familiar and relatable, abstract earcons have a number of advantages working for them as well: systematic, well-defined building blocks can be used to create larger sets of earcons more easily (Blattner et al., 1989); families and hierarchies can be created out of basic audio signals, unlike auditory icons, which cannot be easily manipulated.

2.2.3. Evaluations of non-speech sounds.

2.2.3.1. *Methods.* The methods used in experiments to evaluate nonspeech auditory feedback are discussed by enumerating the following aspects of the studies: the types of tasks that were designed, the independent variables and the dependent variables, including the types of subjective ratings.

Types of tasks designed for evaluations included:

- Finding and selecting the lowest level menu items or recalling menu levels in a hierarchical menu structure, guided by earcons, to evaluate if earcons are effective in communicating menu structures (Barfield, Rosenberg, & Levasseur, 1991; Brewster, 1997a, 1998b; Brewster, Räty, & Kortekangas, 1996).
- Navigating and localizing in a room-based simulation guided by nonspeech audio, having to follow a predetermined route and making sure everything is working properly using the background sounds, to evaluate if nonspeech audio helps in navigational support systems (Skantze & Dahlbck, 2003).
- Identifying picture categories made of line drawings by classifying the pictures as animals or non-animals, to investigate how task performance is affected by auditory icons and earcons. Some conditions employ a dual-task of including a mental addition task for greater cognitive load (Bussemakers & Haan, 1998, 2000; Bussemakers, Haan, & Lemmens, 1999; Lemmens, Bussemakers, & Haan, 2000, 2001).
- Monitoring the simulation of a factory using background sounds (Gaver, Smith, & O'Shea, 1991).
- Listening to and describing everyday sounds (Fernström & Brazil, 2004; Mynatt, 1994), or listening to and selecting the audio cues that best match a function or computer event (Lucas, 1994; Sikora, Roberts, & Murray, 1995).

- Braking when an accident seems imminent in a vehicle-collision avoidance system using nonspeech audio cues and warnings (Graham, 1999).
- Mobile service notifications (Garzonis, Jones, Jay, & O'Neill, 2009).

Table 2.1 summarizes and outlines these methods, tasks and the independent and dependent variables involved together with the references. The results are discussed in the following sections.

Table 2.1: Outline of experimental tasks and variables

Tasks	**Independent variables**	**Dependent variables**	**References**
Finding and selecting the lowest level menu items or recalling menu levels using earcons	Type of menu, icons, commands; Presence or absence of sound; Earcons as musical timbres, simple tones or sounds with no rhythm	No. of correct menu level identifications; error numbers and rates; ease of use	Barfield et al. (1991); Brewster (1997a, 1998b); Brewster et al. (1996)
Navigating in a room-based simulation using the background sounds	Feedback as auditory icons, earcons, or no sound	No. of correct object recognitions; ease of use; annoyance factor	Skantze and Dahlbck (2003)

Table 2.1: Outline of experimental tasks and variables

Tasks	**Independent variables**	**Dependent variables**	**References**
Identifying picture categories of animals or non-animals	Presence of auditory icons, earcons, no sound; relevance of cues	Response time	Bussemakers and Haan (1998, 2000); Bussemakers et al. (1999); Lemmens et al. (2000, 2001)
Monitoring the simulation of a factory using background sounds	Presence or absence of auditory icons; Task difficulty	Time for completion; no. of correct object identifications; recall performance; no. of errors	Gaver et al. (1991)
Listening to and describing everyday sounds, or mapping audio cues to a function		Recall performance; confidence in mapping; pleasantness; appropriateness	Fernström and Brazil (2004); Fernström, Brazil, and Bannon (2005); Mynatt (1994); Lucas (1994); Sikora et al. (1995)
Braking when an accident seems imminent in a vehicle-collision avoidance system	Feedback as auditory icons or earcons	Response time; error rate; preference	Graham (1999)

Table 2.1: Outline of experimental tasks and variables

Tasks	**Independent variables**	**Dependent variables**	**References**
Mobile service notifications	Feedback as auditory icons or earcons	Intuitiveness, learnability, memorability and preference	Garzonis et al. (2009)

2.2.3.2. *Evaluating auditory icons.* Gaver et al. (1991) developed the ARKola simulation of a soft drinks bottling factory to evaluate the effectiveness of auditory icons. This type of auditory feedback was used to represent some of the objects and events within the simulation interface. Pairs of users controlled the factory to test if the audio feedback allowed more efficiency in running the plant and if it affected collaboration efforts. It was observed that the feedback formed a combination of several sounds intermixed with one another, much like the ecology of sounds we use in everyday life, and led the users to be more effective in monitoring the status of ongoing processes. Users were not as efficient in the purely visual condition. The addition of auditory icon feedback also better allowed collaboration between users since one user could not always see the other user's part of the plant, but could hear and identify the relevant feedback.

Mynatt (1994) investigated how well people identify auditory cues by asking participants to describe a collection of short everyday sounds. Results showed that people identified a sound only 15% of the time (a sound is said to be identified if it is mapped to the recorded action or source) in the absence of context. She also found that certain sounds were systematically identified as objects (such as cameras, printers, doors) and some as actions producing a sound (such as closing, tearing, locking). This supports the guideline that some sounds should be selected to represent interface objects, and others for actions.

Similar kinds of listening tests and free-text identification were carried out by Fernström and Brazil (2004); Fernström et al. (2005). Since the design of auditory icons requires intuitive mapping to the computer interface model, Fernstrom *et al* explored what people hear to develop an understanding of people's perception of auditory events and identify mappings for actions in the interface. These responses and categorizations were gathered to investigate how accurately people can identify sounds, and suggest possible mappings and metaphors to the human-computer interface. They found that a fairly high percentage of objects and actions were correctly identified (about 70% of the time). Similar to (Mynatt, 1994), they also found that hearing sounds without context can be confusing, and the order of presentation of the sounds affects the way it was identified. Hence, auditory feedback has to be designed relevant to the function, with care being taken not to have loss of context or high ambiguity.

Graham (1999) on the other hand, evaluated the use of auditory icons for in-vehicle collision avoidance applications. He compared auditory icon warnings with earcons and speech warnings as well. He measured the braking reaction times, number of inappropriate responses and subjective ratings of participants. Results showed that although auditory icon warnings gave faster reaction times and were also rated higher subjectively, they did result in a higher number of inappropriate responses. This meant that the perceived urgency and inherent meaning of such everyday sounds can be easily misinterpreted and care needs to be taken to design these sounds as warnings for such critical applications.

2.2.3.3. *Evaluating earcons.* Earcons differ from auditory icons in that they have no natural link or mapping to the objects or events they represent, and hence, have to be learned.

Brewster, Wright, and Edwards (1992, 1993) performed a series of detailed experiments based on compound and hierarchical earcons to examine their effectiveness. Participants were presented with earcons representing families of icons, menus or both, and had to identify them when played back. This study also investigated whether musical ability affected recall performance and it was found that earcons were recalled equally well by both musicians and non-musicians. However, training the participants in familiarizing themselves with the sounds used was an important factor in recall performance. They conclusively found that earcons were more effective than unstructured bursts of sound and that musical sounds were more effective than simple tones, which differed significantly from the design principles proposed by Blattner et al. (1989). A richer design based on more complex musical timbres gave even better results in communicating information in computer interfaces, leading to the conclusion that complex sounds should be used to design earcons rather than simple tones (Brewster et al., 1992).

Brewster has also shown that structured earcons can reduce information overload by improving usability and task performance, e.g. by reducing the time to recover from errors (Brewster, 1997b). This study describes two experiments where nonspeech sounds were added to graphical buttons and scrollbars. Results showed that participants had a strong preference for the sonically-enhanced widgets over the graphical ones and the sounds also led to faster error recovery from errors such as button slip-off errors or scrollbar positioning errors. They also concluded from the experiment's annoyance ratings that if sounds provide useful information, they will not be perceived as annoying to the user.

Barfield et al. (1991) studied whether earcons can help to represent and recall depth in a menu structure and found that sound did not improve recall performance of depth in menu structures and users found them distracting. However, in a later study, Brewster et al. (1996) also investigated whether earcons effectively provide navigation cues in a menu hierarchy and found different results. Earcons were created for a hierarchy of menu levels and nodes, and participants had to identify their location in the hierarchy using these earcons. Results showed over 80% accuracy, providing evidence that earcons afforded an efficient method for menu localization cues. This study was further extended by Brewster (1998b) with a larger hierarchy, more types of earcons and with a test of recall of earcons over time. Recall over time had good results, but they did, however, find that the type of training had significant results on recall performance. In (Brewster, 1998b), it was found that lower sound quality lowered recall of earcons - in this case, CD quality sound over the lower quality of sound played over the telephone in telephone-based interfaces.

Lemmens et al. (2000) studied the effect of earcons in picture categorization tasks of animal and non-animal line drawings with auditory cues containing redundant information. The drawings were presented either with relevant information via auditory cues of sounds of animals or objects matching the picture, or non-relevant cues in the incongruent condition. In one of the experiments, participants had to carry out an additional mental addition task for greater cognitive load. Results showed earcons containing relevant redundant information helped reduce errors in both the single and dual-task environments.

In a similar study including picture categorization tasks (Bussemakers & Haan, 1998), mood cues in major and minor chords were used with the pictures to see if they affected performance. It was hypothesized that earcons in minor chords suggest a negative emotion and hence should favor a negative answer, whereas those in major chords should favor a positive answer. According to this hypothesis, when the answer to the picture categorization task was positive (yes for animal), major chords should help in response time. However, the auditory mood cues seemed to delay responses in these tasks, leading to the conclusion that the auditory modality together with the visual modality was not always appropriate for these tasks, which they referred to as the "modality appropriateness hypothesis". It is doubtful, however, whether this conclusion can be drawn in all similar situations, since the validity of the use of the auditory mood cues in these tasks is open to interpretation.

2.2.3.4. *Comparisons and combinations.* Both auditory icons and earcons have been found to be effective in communicating information in the human-computer interface through audition. However, each method has its own advantages and disadvantages, and no single method has been conclusively shown to be superior to the other.

Lucas (1994) evaluated the two types of nonspeech auditory feedback, and compared them with speech cues as well. Participants had to listen to and select which audio cues from the three best represented an action or object in the interface. An explanation of the design of the cues were given to half the subjects and both halves were tested again one week later to see if design knowledge helps recall performance. It was found that this prior knowledge did help retain information on the cues. Results also showed that after speech, auditory icons were most accurately associated with the correct action or object.

Bussemakers and Haan (2000) investigated whether redundant auditory icons used with visual information influence the performance on picture categorization tasks on line drawings, and they compared the results with experiments using redundant earcons. Results illustrated that response times are faster in conditions with auditory icons than the silent condition, and that response times were even slower with earcons than the silent condition. Lemmens et al. (2001) performed two more similar experiments, one with a dual-task requiring a mental addition task with the picture categorization, and one experiment using intermixed auditory icons and earcons with the picture categorization task. These experiments confirmed the previous results: although the dual-task slowed reaction times, auditory icons still led to faster response times than earcons. Hence, auditory icons seemed to have a facilitatory effect in picture categorization tasks of this kind, while earcons seemed to have an inhibitory effect.

A navigational support approach in a building maintenance system using a room-based metaphor was evaluated by Skantze and Dahlbck (2003). It used auditory icons for auditory feedback and was compared with the system using earcons and no sound. It was found that none of the subjects preferred the earcon condition in this experiment, while 67% preferred auditory icons. Also, auditory icons allowed a significantly better recall performance for navigation in the rooms.

Sikora et al. (1995) designed auditory feedback in a graphical user-interface for business communication using either musical sounds (earcons) or real world sounds (auditory icons). Users mapped the sounds to functions and rated their confidence in the functional mapping, its pleasantness and appropriateness. Real world sounds mapped most predictably to functions, although musical sounds had higher ratings for pleasantness. For the business application, no auditory icons were selected. Hence, preference does not always reflect the best functional mapping. The authors also concluded that real world sounds may be less appropriate for actual workplace applications. Edworthy (1998) tried to determine if sound helps people work better with machines and suggested that real world sounds may be more suited for auditory feedback on monitoring tasks via background sounds, while abstract sounds may be better suited for warnings and alarms as they tend to attract our attention more effectively.

However, in a more recent study, Garzonis et al. (2009) evaluated auditory icons and earcons as mobile service notifications, by comparing them on the four factors of intuitiveness, learnability, memorability and user preference. A longitudinal evaluation involving two lab experiments, a field study and a web-based experiment showed that auditory icons performed significantly better in all four factors.

2.2.3.5. *Summary.* Table 2.2 summarizes the more relevant or definitive findings of the evaluative studies discussed in this section. Factory-monitoring tasks were facilitated by adding auditory icons; recall of level and location during navigation in hierarchical menus were improved by adding earcons. However in picture categorization tasks and room-based navigation tasks, auditory icons were found to be more effective and preferred over earcons. This is in contrast with business applications, where earcons were given higher subjective ratings, and in-vehicle collision avoidance systems, where earcons gave rise to fewer errors than auditory icons. However, it should be noted that the methodologies used for each study differ from one another significantly and often do not agree.

2.2.4. Contexts of use. In this section, some of the applications that have used nonspeech sounds for enhancement and feedback are discussed. Examples also include a few non-visual and audio-haptic interfaces as well.

2.2.4.1. *Desktop applications.* One of the first desktop interfaces developed using auditory icons was the SonicFinder (Gaver, 1989), mentioned in Section 2.2.1, where real-life sounds were mapped to different common interface objects and events for intuitive auditory feedback. For example, selecting interface objects made sounds of tapping a material depending on the type of object, e.g. files gave a wooden sound, applications a metal sound and folders a paper sound. Copying actions were aurally illustrated using a pouring analogy - the sound of how full the receptacle indicated the progress of the copy action (with increasing pitch). The challenge, however, is

TABLE 2.2. Summary of evaluative and comparative studies

Tasks or Environment	**Type of sound**	**Results**
Factory monitoring systems	Auditory icons	Increased efficiency, increased collaboration
In-vehicle collision avoidance system	Auditory icons	Low responses time, high number of inappropriate responses, high subjective ratings
	Earcons	Less number of inappropriate responses
Hierarchical menus	Earcons	Highly effective in recall of menu level and location
Picture categorization tasks (line drawings)	Auditory icons	Facilitatory effect
	Earcons	Inhibitory effect
Navigational support system in room-based simulations	Auditory icons	Higher subjective rating, higher recall
	Earcons	Lower subjective ratings and recall
Business applications	Auditory icons	Better functional mapping, low subjective rating
	Earcons	High subjective ratings

finding representative sounds for all actions and events, since some events at the interface level are abstract and difficult to portray with a real-life sound.

Brewster (1998a) developed earcons for desktop use and performed detailed evaluations on different types of earcons, as discussed in Section 2.2.2.

2.2.4.2. *Complex systems.* Some applications that have utilized nonspeech audio have integrated it into much more complex environments than the desktop interface. One such study mentioned previously in Section 2.2.1 was the ARKola simulation, which used an ecology of auditory icons in a complex soft drinks factory simulation to convey information about the current state of the factory and its components and help improve collaboration efforts (Gaver et al., 1991).

Skantze and Dahlbck (2003) described another such complex environment portraying a navigation support approach based on auditory icons for navigating in room-based designs. The prototype system simulated a buildings maintenance support system using a room-based metaphor. It was found that users responded positively to the use of auditory icons, rather than earcons, in this environment.

Mynatt, Back, Want, Baer, and Ellis (1998) designed a more complex system that provides continuous serendipitous information to users via background auditory icon cues in the workplace. The Audio Aura system provides information to the user even when away from his desk, so that users do not have to be confined to their office space at all times. The auditory peripheral cues are meant to be ambient and provide information which can be ignored if not required. For example, the sound of surf represented the amount of new e-mails received by the user, with a higher number of e-mails being characterized by increasing surf. They used an electronic tag and networking system and wireless headphones linked to each person in the workplace for tracking and notifying purposes.

2.2.4.3. *Mobile devices.* In today's world where communication in a mobile environment is critical, mobile devices interfaces have to be designed well to compensate for the lack of screen space and low-resolution visual data. Hence, Brewster, Leplatre, and Crease (1998) proposed the use of nonspeech sounds to improve interaction without the need for more screen space. Later, Leplatre and Brewster

(2000) described a framework for integrating such nonspeech audio to mobile phone menus where visual feedback is constrained. The hierarchical menu structures were enhanced using earcons, and evaluations showed significant performance benefits in navigating the menus from the sonifications.

In (Holland, Morse, & Gedenryd, 2002), a prototype audio user interface for a GPS system is designed so that users can carry out location tasks on mobile computers while their attention and hands are occupied elsewhere. The interface uses a simple form of spatial audio, rather than speech audio, and was shown to be effective and inexpensive for location tasks.

A similar application is the Nomadic radio, a wearable computing platform for managing voice and text-based messages in a mobile environment (Sawhney & Schmandt, 2000). It uses an auditory user-interface for navigational and notification purposes among messages. Speech audio and spatial auditory icon cues are continuously played in the background to provide peripheral awareness of the system status. Evaluations showed that users preferred this type of auditory awareness to speech-based navigation systems.

2.2.4.4. *Applications for the visually impaired.* One of the most important uses, and the most widely studied areas, for nonspeech audio is in computer applications for visually-impaired users (for a recent review, see Murphy (2007)). Since speech audio takes time to be played out and listened to, and hence is not the most efficient method of communication, nonspeech audio can be effectively replaced as some of the feedback in such applications. Mynatt (1997) developed a methodology for transforming graphical interfaces into nonvisual auditory interfaces by converting the salient components of the graphical interfaces into auditory components. Auditory icons are used to convey these interface objects, based on a hierarchical model of a graphical interface, providing visually-impaired users many of the benefits of graphical user interfaces (GUIs). Mynatt and Weber (1994) also compared two different applications for converting GUIs to nonvisual interfaces: *Mercator* replaces the spatial graph display with a hierarchical auditory interface, while *GUIB* translates the screen contents into tactile information based on the spatial arrangement of the GUI. User evaluations showed that auditory cues as used in Mercator were very effective for nonvisual interfaces.

Morley, Petrie, O'Neill, and McNally (1999) designed an auditory system for visually-impaired users to enable efficient navigation on the web or hypermedia. This interface uses nonspeech sounds to identify links and provide information and feedback about text and commands to improve usability. They incorporated naturalistic auditory icons where appropriate, to engage blind students, and simple earcons for other situations. Evaluations showed that participants liked these sounds and found them easy to remember. The auditory feedback allowed them to work faster and more efficiently than conditions without feedback. Goose and Möller (1999) also designed a web browser using spatialized 3D audio to convey the structure of the hypermedia document. It provides audio structural surveys, positional audio feedback of links and anchors, progress indicators and meta-information of new links, improving browsing experience for both sighted and visually-impaired users.

Another such tool for web access was developed by Murphy, Kuber, Strain, McAllister, and Yu (2007). They designed a plugin for web browsers that provides auditory feedback and haptic cues to enable visually-impaired users to spatially localize themselves on web pages, and build a mental model of the spatial structure of the web document to enable effective navigation of web pages. The plugin generates audio feedback to indicate links, images, and other such web objects, and also aurally indicates when the user crosses the boundaries of the page.

2.2.4.5. *Immersive systems.* Auditory feedback also plays a very important role in contributing to the feeling of presence in immersive virtual environments (Slater & Usoh, 1993). Grohn, Lokki, and Takala (2003) carried out a navigation test in a spatially immersive virtual environment that simulated a game-like experience to test this. The system used both auditory and visual cues. It was found that audio-visual navigation was more efficient and immersive, than only visual or auditory navigation in a 3D virtual environment. Multimodal gameplay using gestural input in a virtual environment which incorporated both audible and visual cues as feedback was also shown to foster a sense of immersion and increased user preference (Benovoy, Zadel, Absar, Wozniewski, & Cooperstock, 2008).

Auditory feedback has been added to virtual assembly environments and studies have been performed to evaluate task performance in such environments. Zhang, Fernando, Xiao, and Travis (2006) presented an approach for the integration of 3D auditory feedback into virtual assembly environments and evaluated the resulting system. They reported that the addition of auditory feedback improved task performance and that audio-visual integration gave the best results, when compared to any individual modality feedback alone. Edwards, Barfield, and Nussbaum (2004), on the other hand, studied whether the inclusion of auditory cues or force feedback to an immersive virtual environment improved the performance of an assembly task. Results showed that the addition of force feedback slowed completion times and increased errors in some users, while auditory feedback had no such negative performance effects.

2.2.4.6. *Summary.* Applications and tools that apply auditory feedback to improve usability of desktop applications, web interfaces, or more complex environments simulating real-life situations are discussed. The applicability of auditory feedback to mobile devices is discussed to improve usability by reducing visual clutter and amount of screen space required to communicate information. Applications using auditory feedback for better access by visually-impaired users are also described. Applying auditory enhancements to immersive virtual environments has also been found to increase the sense of presence and improve performance in virtual assembly tasks as well as navigation tasks. Adding auditory feedback to navigation tasks in the above applications was found to be beneficial to user performance or experience.

2.2.5. Menu navigation using audio feedback. Since one of the main purposes of this proposal is to look into the navigation of information spaces, an area closely related is the use of auditory feedback to assist in menu interaction.

Brewster has worked extensively in the use of nonspeech audio in menu navigation. As was mentioned in Section 2.2.3, Brewster et al. (1996) investigated earcons as navigation cues in a menu hierarchy. Results showed high accuracy, proving that earcons afforded an efficient method for menu localization cues. This study was further extended by Brewster (1998b) with a larger hierarchy and more earcons. A test of recall of earcons over time showed good results on recall performance, depending on the type of training received. Subsequently, Vargas and Anderson (2003) also showed that combining speech and earcons in spoken menu systems to assist in navigation improved task performance.

Yalla and Walker (2008) presented design concepts for auditory menus, where auditory feedback is received during menu navigation to aid sighted or visually impaired users. They examined different types of auditory scrollbars for an auditory menu and detailed how visual menu concepts can be applied to auditory menus.

S. Zhao, Dragicevic, Chignell, Balakrishnan, and Baudisch (2007) designed and evaluated an eyes-free menu technique called EarPod that combined both auditory feedback and haptic touch input. They found that it was potentially a reasonable eyes-free menu technique for general use, especially for use in mobile devices.

2.2.6. Discussion. Some of the advantages offered by nonspeech sounds besides reducing workload on users' visual system (M. Brown et al., 1989) are that they can provide complementary information to vision (Brewster, 2003); they can also reduce the amount of information needed to be displayed on screen, hence optimizing screen space and reduce the demand on visual attention. Sound is also attention-grabbing and can be used for peripheral awareness and ambient audio (Arons & Mynatt, 1994).

There are, however, some disadvantages to using nonspeech sounds (Brewster, 2003) which need to be addressed if audio is to be successfully incorporated into human-computer interfaces. Presenting either abstract data or absolute data using sound is often difficult. Sounds can be used to portray relative differences in values, but to get an absolute value, users typically require looking at a number or graph. Another issue is that audio is a transient medium which disappears after it is presented, and has to be replayed if not remembered. Stationary visual data, on the other hand, can be referred back to whenever required. Many audio parameters are unsuitable for high-resolution display of information. And finally, auditory feedback can cause annoyance in users if not designed appropriately.

However, the previous research discussed in this chapter prevalently show that nonspeech audio is an effective means of communicating information to the user in the computer interface, be it via auditory icons or earcons, in a multitude of applications. The results of the studies discussed here are promising for audio-visual integration in computer interfaces, as relevant auditory feedback tends to enhance task performance in the specific modalities.

Auditory icons have the advantage of being easy to learn and remember as they are natural and relatable to our everyday lives. Auditory icons with good mappings and metaphors can make for a very effective feedback system for most users. However, the disadvantage for this type of feedback also arises from this issue. All computer interface functions and objects do not have real world equivalents and it may be hard to find a metaphor to represent such functions without being faced with issues of ambiguity, loss of context and even annoyance factors in users.

While earcons have the converse disadvantage of having to be learned and remembered since they have no natural intuitive link to the interface action or object, they have the advantage of being highly structured. As such, it is easier to follow structured design principles to create families of earcons, so that users typically can learn to recognize them by remembering their common characteristics and attributes. Auditory icons, on the other hand, have to be remembered individually, as it is not easy to connect them in structured families.

This section has highlighted the fact that auditory icons and earcons are each more effective than the other in different environments and task situations. Preliminary conclusions and deductions that can be drawn from the literature reviewed has been categorized and summarized in Table 2.3 and Table 2.4, for auditory icons and earcons, respectively, while keeping in mind that the methodologies in each study are vastly different.

Table 2.3. Applications and tasks where the addition of auditory icons have been found to facilitate tasks

Desktop interfaces	*Navigation*; picture categorization; hypermedia and web interfaces
Complex systems	Monitoring tasks; collaborative tasks; peripheral awareness cues, ambient sound; navigation tasks
Immersive virtual environments	Localization and *navigation* tasks; assembly tasks

It can be seen from the two above tables that one of the tasks that have been benefited by both auditory icon and earcon feedback is navigation (shown in italics in the tables), be it in the context of desktop interfaces, immersive systems or mobile

TABLE 2.4. Applications and tasks where the addition of earcons have been found to facilitate tasks

Desktop interfaces	Sonically-enhanced widgets; *navigation* of menu hierarchies; business and workplace applications; graphs and tables; *navigation* of hypermedia and web interfaces
Alarms and warning systems	Vehicle-collision detection
Immersive virtual environments	Assembly tasks
Mobile systems	*Navigation* in mobile phone menus

systems. Hence, it is well worth investigating the effects of adding auditory feedback in navigational tasks specifically related to information seeking in information systems. Thus, the design and evaluation of such a system using auditory feedback as navigational and structural cues to improve user experience merits further research and is one of the primary goals of this dissertation.

CHAPTER 3

CONCEPTUAL FRAMEWORK

3.1. Models of Navigation

Navigation is human behavior intended to make sense of an "information space", as has been defined by Newby (1992). Newby also defined an *information space* as the set of concepts and relations among them possessed by an information system. Previous conceptualizations of frameworks for the process of navigation have been in the context of real physical environments, where navigation requires using concepts such as landmarks and routes. With the advent of virtual electronic spaces, it has become necessary to develop new frameworks for navigation, as the concepts involved in navigating in virtual and abstract worlds is essentially different from those involved in physical space (Wittenberg, 1997).

However, since most information is laid out spatially, we are dealing with spatial models of navigation. Dourish and Chalmers (1994) classify two main areas of application for spatial models: *inherently spatial* ones, such as seen in computer-based maps of physical spaces, or *semantic* models, in which an underlying semantic relationship between information objects is mapped onto a spatial layout. Systems using the latter model, group objects according to similarity or some semantic relationship between objects and render these relationships as spatial dimensions. Hence, navigation entails not only purely spatial movement, but semantic navigation performed in spatial terms.

At a workshop on Navigation in Electronic Worlds (Jul & Furnas, 1997), Darken proposed a framework for the navigation process in an information space, shown in Figure 3.1. This model involved four main steps: the task being performed, the strategy used for navigation, the movement or action required to complete this and the evaluation of the user's progress, resulting in either going back to reformulate the strategy or repeat the action. It did not explicitly consider the formulation of an internal cognitive model of the space.

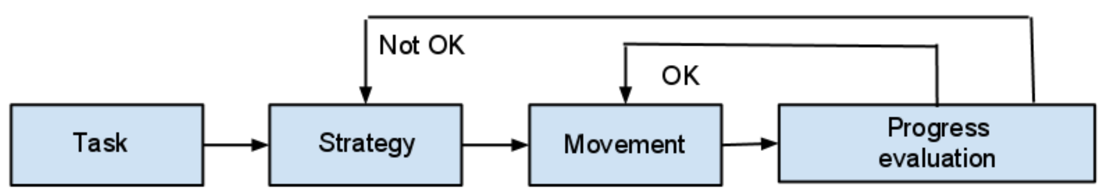


Figure 3.1. Darken's model of navigation (Juls & Furnas, 1997)

However, in this same workshop (Jul & Furnas, 1997), another framework for navigation was proposed by R. Spence (1997), which was modified by the participants to make it more comprehensive (shown in Figure 3.2). Here, the significance of the formulation of a mental model was realized, as can be seen from the figure.

This model was extended and improved on in Spence's 1999 model for navigation, in which the explicit representation of a goal was taken out since this is not always a part of browsing an information space, shown in Figure 3.3. This research

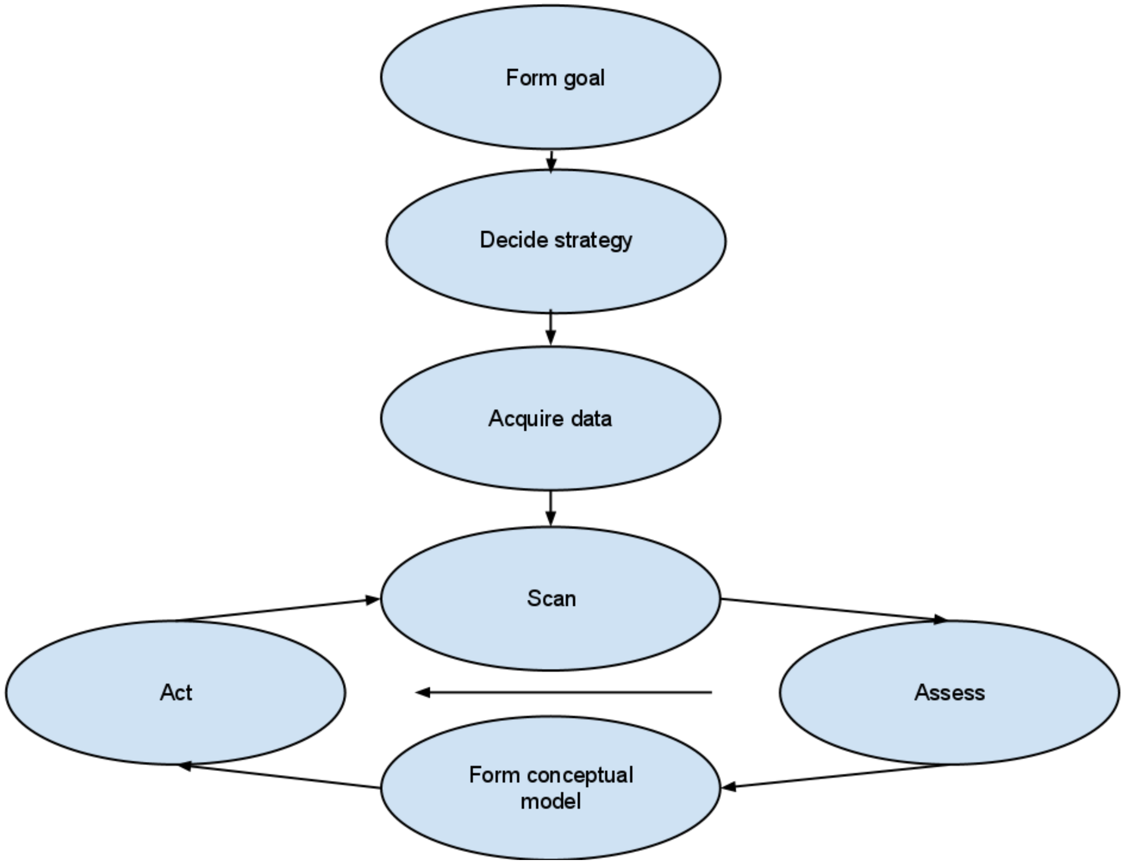


FIGURE 3.2. Spence's (1997) model of navigation (R. Spence,1997)

will mainly refer to the following framework as a navigational model. The framework defines navigation as "the creation and interpretation of an internal model" (R. Spence, 1999). It is composed of four cognitive activites: browsing, formation of an internal model or cognitive map, interpretation, and formulation of a browsing strategy (Figure 3.3).

The *browsing* activity as defined in this model signifies the assessment of the contents of the space, where the result of perception is briefly held in sensory storage (e.g. the act of scanning through a newspaper before deciding what to read). There may be no specific goal in mind while browsing. The content that was assessed in the browsing activity is then integrated and constructed into the internal model, comprising the *formation of the internal model* step. The *interpretation* step entails interpreting the available internal model and displayed data to form a decision on how the navigation process should proceed, resulting in the *formulation of the browsing strategy*. A sequential and iterative traversal of these cognitive activities comprises the *navigational process* as defined by this model (R. Spence, 1999).

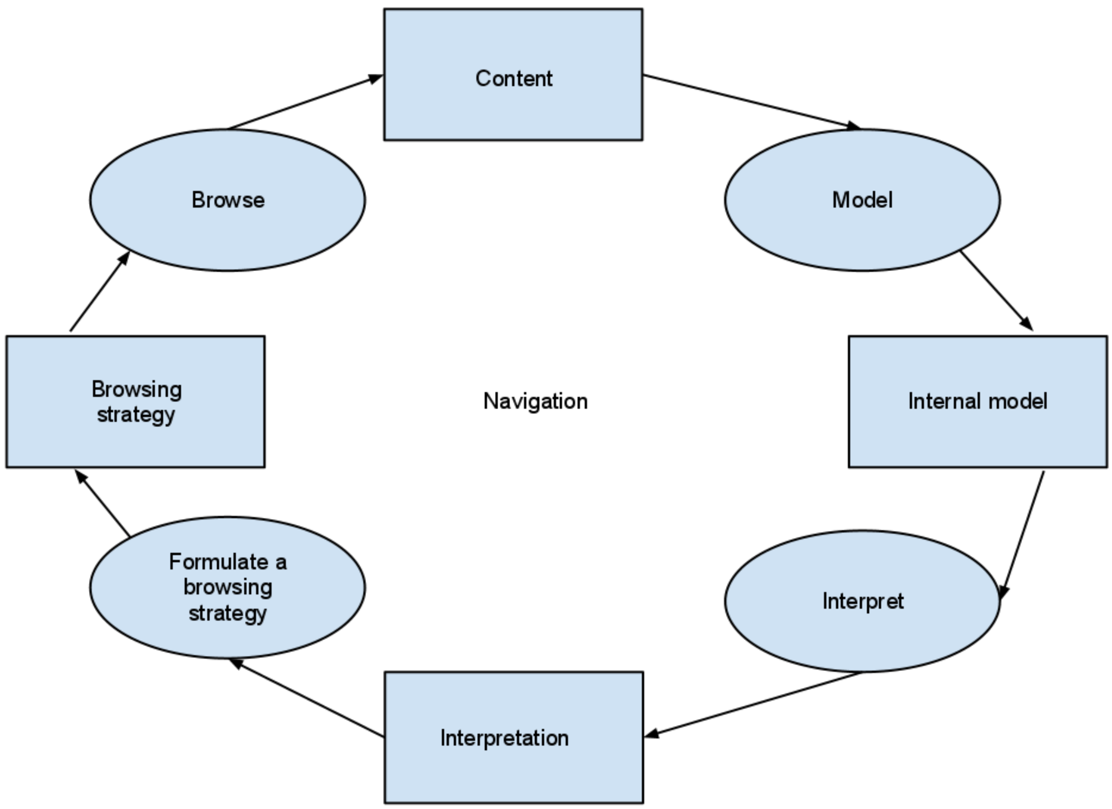


FIGURE 3.3. Spence's (1999) model of navigation (R. Spence, 1999)

To be able to effectively navigate through a world with minimal prior knowledge of its layout, it is required that there be cues or signs presenting information that will help the navigator make their next navigational decision. The proposed research will extend the local views that present visual cues to the navigator by introducing auditory cues to aid in navigational decision making. As Spence recognized, the modelling step is profoundly affected by the way the raw data is displayed and the affordances it provides the navigator (R. Spence, 1999). The aim is to focus on the formation of an internal model or "modelling" part of the navigational framework, as illustrated in Figure 3.3. Adding auditory feedback would add to the information "Content" of the information space, by adding structural and navigational information. This would lead to the "browsing" or simultaneous scanning of both visual and auditory content, which in effect may lead to more effective "interpreting" of the information structure, and consequently to improved "modelling" of the space.

It is hypothesized that the presence of auditory feedback will reduce the cognitive load imposed on the navigator to create the internal model of the information structure, although this may be shown by improvements in the experiential or affective aspects for the navigator of the system more significantly than the quantifiable performance. The next section describes the literature in theoretical frameworks for technology acceptance or user satisfaction in using new technologies. These frameworks will help understand how the evaluation of the user experience can be conceived and a framework for measuring the user experience influenced by the auditory enhancement of the hierarchical information system in the current research is proposed.

3.2. Models of Technology Acceptance and User Satisfaction

> *"User experience is about technology that fulfils more than just instrumental needs in a way that acknowledges its use as a subjective, situated, complex and dynamic encounter."*
>
> (Hassenzahl & Tractinsky, 2006, p. 95)

In the past two decades, there has been extensive research done on user experience and what it encompasses. User experience is a multi-faceted concept, and as conceptualized by the authors of the above quote, it is a consequence of a user's internal state, such as predispositions, motivation, expectations, mood, etc.; the system's characteristics, such as usability, functionality, complexity, etc.; and the context of use, e.g. for work, entertainment, etc. (Hassenzahl & Tractinsky, 2006). These facets of user experience is illustrated in Figure 3.4.

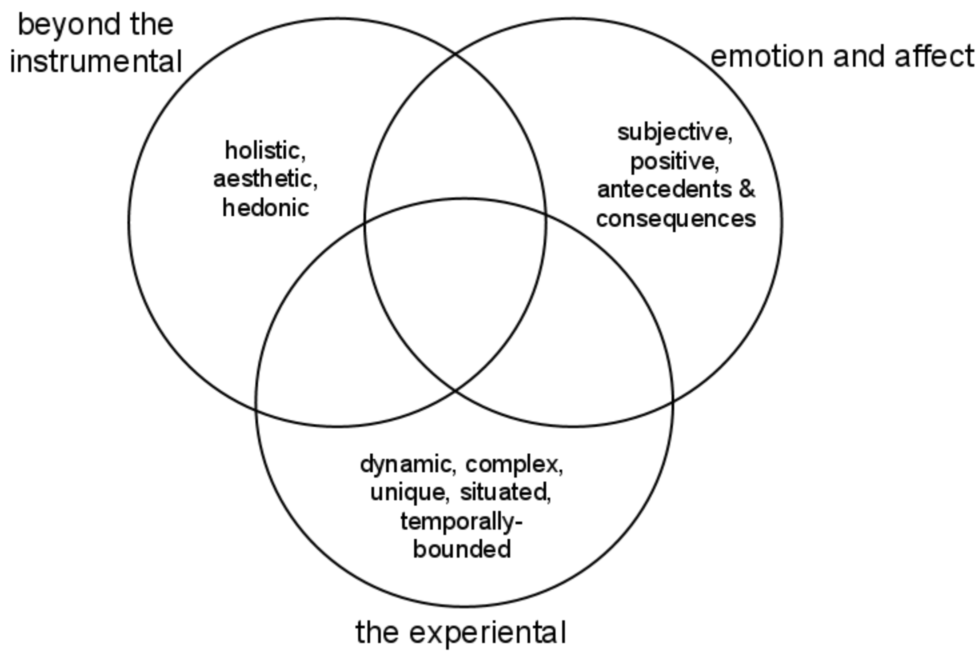


FIGURE 3.4. Facets of User Experience (Hassenzahl & Tractinsky, 2006)

There have been mainly two streams of research that examine user perceptions of information technology and how it affects usage: the *technology acceptance* approach and the *user satisfaction* approach.

Davis (1985) proposed the first conceptual model for the Technology Acceptance Model (TAM) in his doctoral thesis. He proposed that the use of a system can be predicted by a user's motivation, which is in turn predicted by the features and capabilities of that system, as shown in Figure 3.5.

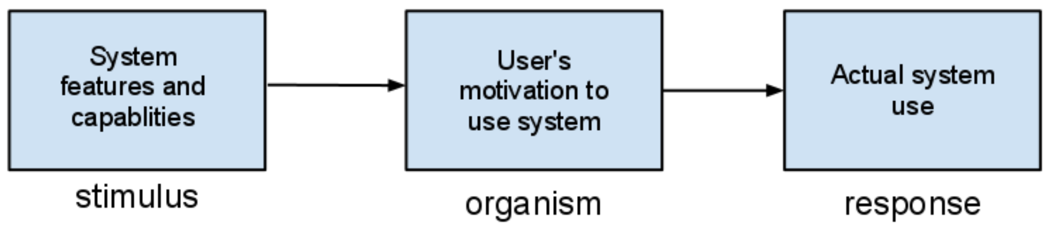


FIGURE 3.5. Conceptual model of technology acceptance (Davis, 1985)

This conceptual model was then refined to the original Technology Acceptance Model shown in Figure 3.6. In this revised model, user motivation is described by three determinant factors that predict the consequent acceptance or rejection of a new technology or system. The user's *perceived usefulness* and *perceived ease of use* influence the *attitude toward using*, which in turn determines the actual usage of the system. And these two main perceptions were influenced by the system design and capabilities (X1, X2, etc. in the Figure 3.6). Davis' measurement scales for perceived usefulness and ease of use (Davis, 1989) are included in Appendix D. Several of these measurement scales have been incorporated into the post-test questionnaires in the current research, described further in Chapter 5 and summarized in Table 5.6.

Over time, the TAM has become a very popular basis for research in predicting use of information systems and has generated several proponents of the theory (e.g. Hendrickson, Massey, & Cronan, 1993; Sharp, 2007), as well as criticisms (e.g. Bagozzi, 2007; Legris, Ingham, & Collerette, 2003). The original TAM has thus been refined or extended with other dimensions and variables (Venkatesh & Davis, 2000); some popular extensions are displayed in Figure 3.7. Several researchers have also applied the model or used it to consolidate their results in different applications. A

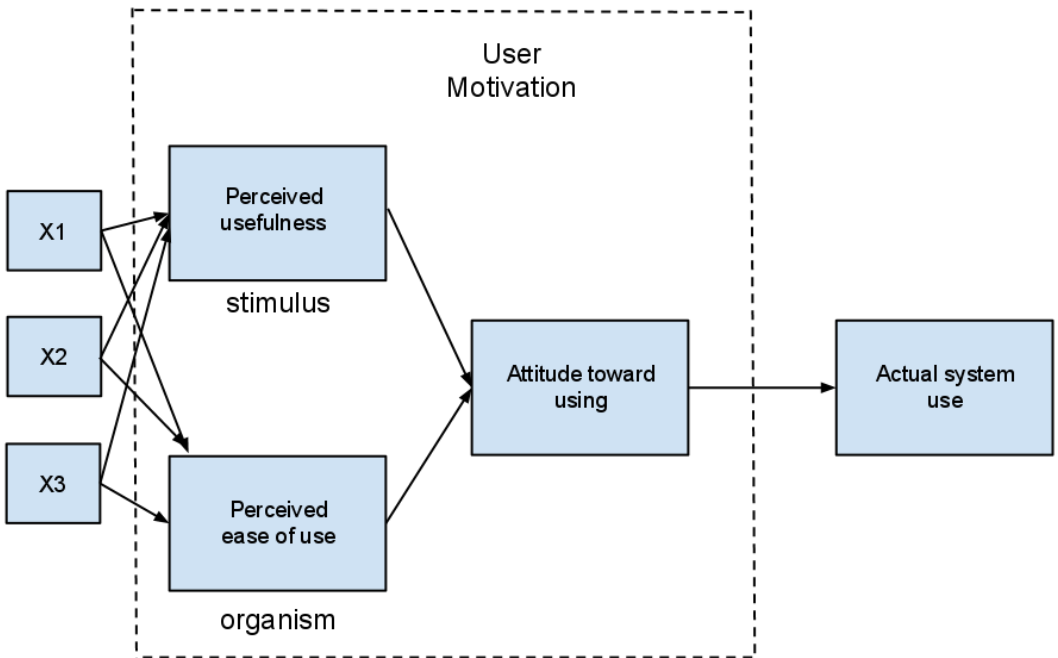


FIGURE 3.6. Original TAM (Davis, 1985)

review or meta-analysis of 145 such articles can be found in (Yousafzai, Foxall, & Pallister, 2007a, 2007b).

However, to go beyond the two main key concepts of the TAM limited to perceived usefulness and perceived ease of use, we now review the second research stream dealing with user response to technology: User Satisfaction. The concept of user satisfaction emphasizes more on system characteristics and design attributes, rather than behavioral beliefs, as in the TAM (Wixom & Todd, 2005). User satisfaction has been defined as a complex construct comprising several affective components such as emotion, expectation, likeability and usability (Lindgaard & Dudek, 2003). In

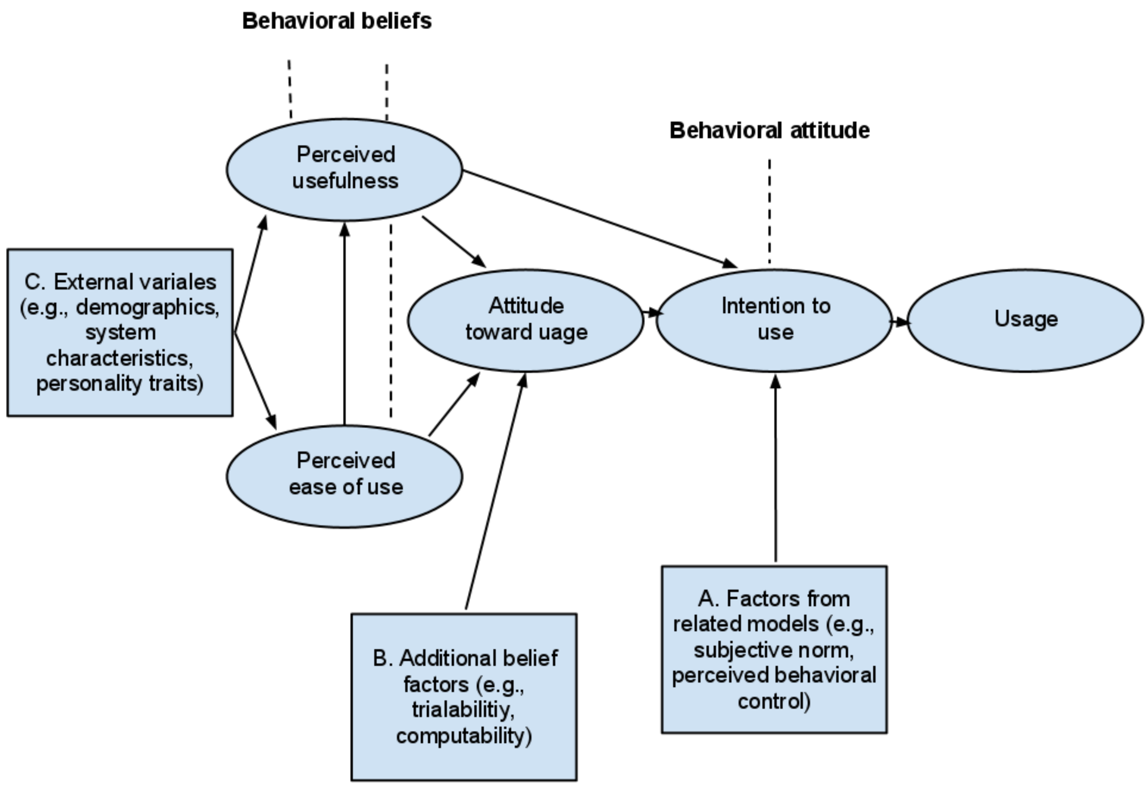


FIGURE 3.7. Three popular extensions of TAM (Wixom and Todd, 2005)

the context of information science, a large amount of research has been devoted to establishing a standard instrument for measuring user satisfaction in using information systems. Two of the most popular scales include those developed Bailey and Pearson (1983) and Doll and Torkzadeh (1988). These scales are still used today and have been applied and validated repeatedly in the literature (Gelderman, 1998; Gatian, 1994; Xiao & Dasgupta, 2002). A more comprehensive review of literature dealing with user satisfaction in the context of information systems can be found in (Griffiths, Johnson, & Hartley, 2007).

Bailey and Pearson (1983) defined user satisfaction as the sum of one's positive and negative reactions to a set of factors. They identified a scale of 39 factors comprising the domain of satisfaction, such as means of input, convenience of access, accuracy, timeliness, reliability, format of output and relevancy. Similarly, the instrument developed and validated by Doll and Torkzadeh (1988) consisted of five factors: content, accuracy, format, ease of use and timeliness (Figure 3.8), quite a few of which were in common with Bailey and Pearson's instrument. Doll and Torkzadeh's one was however regarded as more comprehensive by subsequent studies since it was based on a review of previous work (Xiao & Dasgupta, 2002) and also included a component for "ease of use" and two global measures for perceived overall satisfaction. The global measures for satisfaction were incorporated into the measurement scale for the current research and described in more detail in Chapter 5 (Section 5.2.6). The original questions from Doll and Torkzadeh's instrument are included in Appendix D.

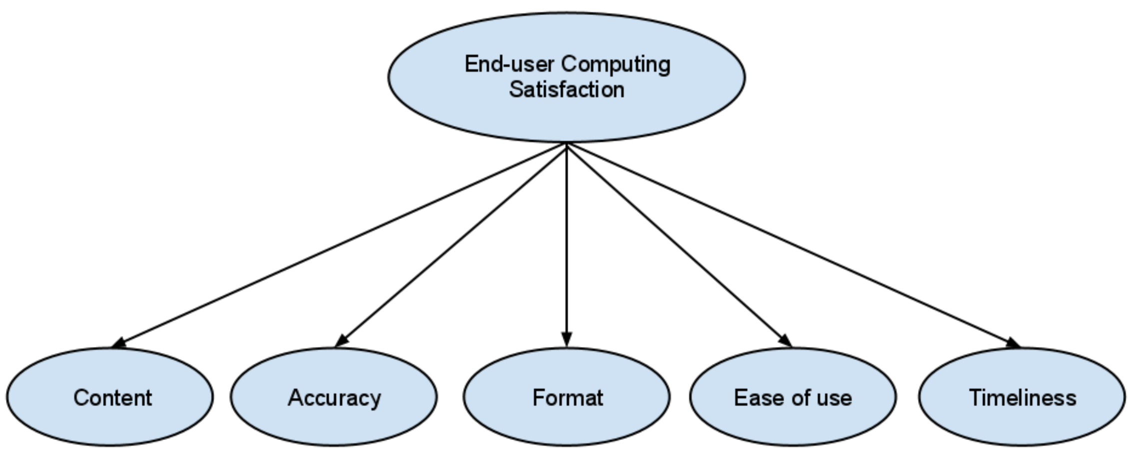


FIGURE 3.8. The EUCS model (Doll and Torkzadeh, 1988)

An attempt to integrate the two models or research streams has been undertaken in (Wixom & Todd, 2005). Their proposed integrated research model is shown in

Figure 3.9 (survey items attached in Appendix D). Lund (2001) has also integrated the key concepts from the two models and developed a questionnaire that measures usability using primarily three factors: usefulness, satisfaction, and ease of use. Using factor analyses on several studies, it was found that as predicted in the TAM model, perceived ease of use and perceived usefulness influence one another, such that improvements in perceived ease of use improve ratings of usefulness and vice versa. It was also found that satisfaction was strongly related to actual or predicted usage. Lund's questionnaire also includes scale items on the perceived ease of learning, as a component of the ease of use variable. This questionnaire can also be found in Appendix D, and several of the items have been used in the experimental design, described in Chapter 5.

Besides the two beliefs of perceived usefulness and perceived ease of use present in the TAM, a further extension to the TAM included a third belief called *perceived enjoyment* (Venkatesh & Davis, 2000). This is defined as "the extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated". van der Heijden (2004) classifies information systems that appear to be accepted more because of their perceived enjoyment and perceived ease of use, rather than their perceived usefulness, as *hedonic* systems, where *hedonism* denotes pleasure or happiness. Examples of these systems include systems used in the home for leisure or entertainment, gaming systems, the World Wide Web, etc. These are in contrast to mainly *utilitarian* information systems which aim to provide instrumental value to the user and increase the user's task performance while encouraging efficiency.

Hedonic systems aim to provide a pleasurable experience or one that is "fun to use". Heijden surmises that this may entail seeking sensations on multiple sensory channels, the inclusion of hedonic content, such as graphics, a focus on colors and sounds, and aesthetically appealing layouts. Adding hedonic features can increase

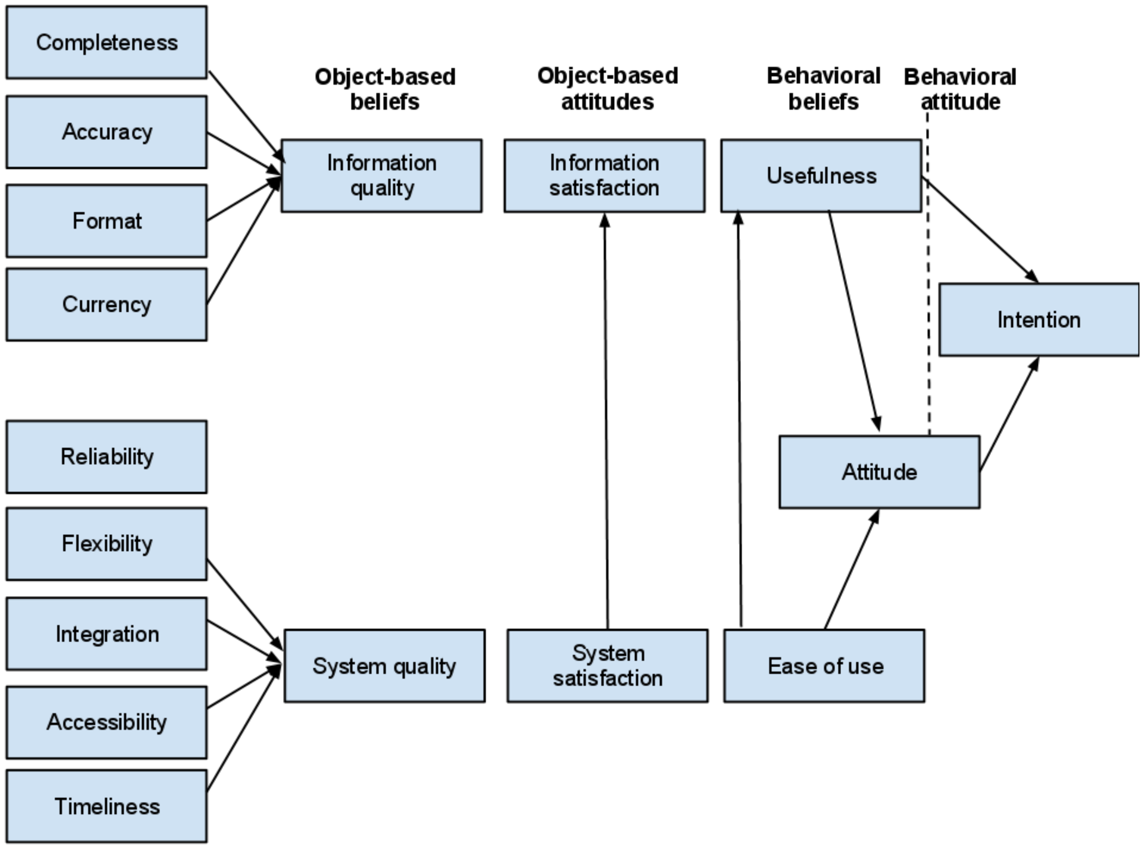


FIGURE 3.9. Wixom and Todd's proposed integrated model of user satisfaction and technology acceptance (Wixom and Todd, 2005)

acceptance of otherwise utilitarian systems (van der Heijden, 2004). Hence, it is important to include this extension of the technology acceptance model to measure the effects of auditory enhancement of the information system.

Another body of literature that studies the measurement of how enjoyable, pleasurable or *engaging* a technology is involves measuring the *user engagement* with the system and provides one way of measuring the hedonic value of a system. HCI studies have established a shift in the paradigm that looked primarily at ensuring the usability of systems to allow for the design of more engaging systems (Hassenzahl &

Tractinsky, 2006). The characteristics of engaging systems that can be found in different studies include those of interactivity, sensory engagement, and sensory fidelity through the use of multiple channels such as text, graphics, and sound (Laarni et al., 2004). Sensory appeal and the format of presentation of the application has also been linked to aesthetics, which in turn has been found to be important to engagement (Overbeeke, Djajadiningrat, Hummels, Wensveen, & Frens, 2003).

Recent studies have deconstructed the term engagement as it applies to user experience with technology (O'Brien & Toms, 2008). O'Brien and Toms conceptually and operationally defined the term through an extensive review of multidisciplinary literature and an exploratory study of users of four areas of applications: Web searching, Webcasting, online shopping and gaming. Using previous literature and the four exploratory studies as the basis, they defined the different attributes that contribute to engagement to be characterized by challenge, aesthetic and sensory appeal, feedback, novelty, interactivity, perceived control and time, awareness, motivation, interest and affect.

They subsequently developed a survey to measure user engagement consisting of 31 scale items (O'Brien & Toms, 2010a) and assessed its reliability and validity with two large-scale survey studies on online shopping applications. This User Engagement Scale (UES), included in Appendix D, corresponds to six attributes of engagement. In a later study, where they applied the scale on an interactive information retrieval system, they used factor analysis to refine the attributes to five attributes: Aesthetics, Focused attention, Perceived Usability, Novelty and Endurability. A subset of these scale items were used in the experimental design of this dissertation and is further described in Section 5.2.

Hence, the current study draws on the literature and concepts derived from three conceptual frameworks - Technology Acceptance, User Satisfaction and User Engagement (Figure 3.10). No single framework and no individual questionnaire

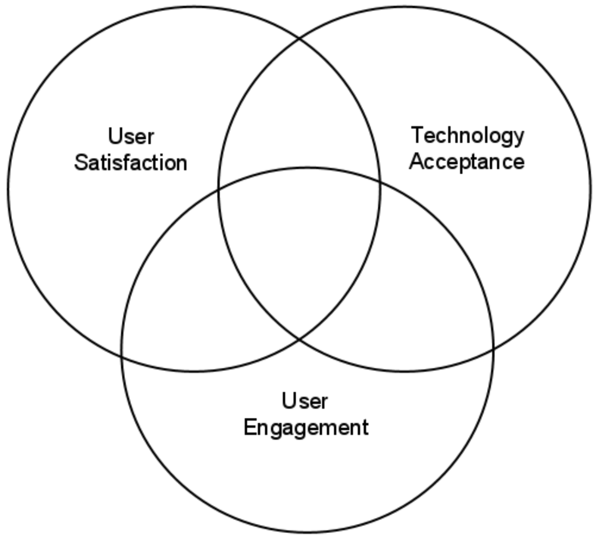


FIGURE 3.10. The conceptual framework drawn from three bodies of literature

scale proposed by the studies discussed is appropriate by itself for the current research, since they each deal with different environments, contexts or applications. Furthermore, none of the above studies address the presentation of information to an additional sensory modality in a system and measure the effects on user experience. Hence, it is necessary to combine concepts from each framework and adapt the scale items from different measurement scales to make it relevant and appropriate for this study.

A more detailed model for the measurement of user experience can be seen in Figure 3.11, which shows the main variables specific to the current research. The

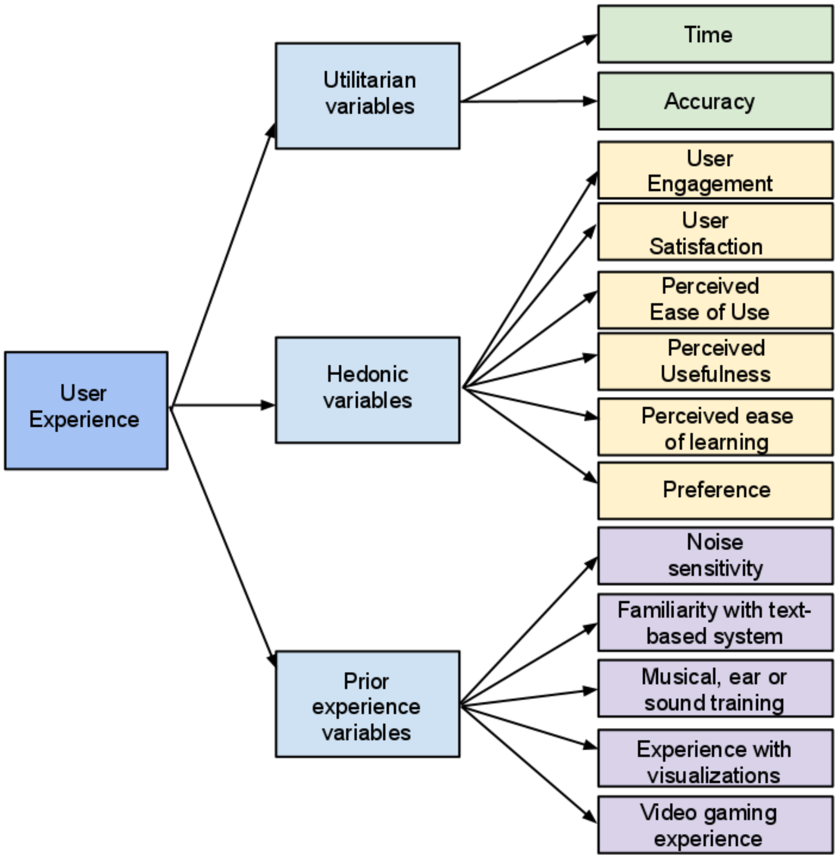


FIGURE 3.11. The proposed integrated model for measuring user experience of auditory enhancement in current system

"utilitarian" variables were drawn from the User Satisfaction literature and the evaluation studies of auditory feedback described in Chapter 2. These are the only variables that are not self-reported by the user, but are evaluated from system readings. The "prior experience variables", which may influence user acceptance, are based on the "external variables" addendum found in the extension of the TAM (refer to Figure 3.7). The "hedonic" variables are a combination of concepts derived from all three schools of thought, including User Engagement. The scale items corresponding to each variable on the right-hand side of Figure 3.11 is described in detail in Chapter

5, Section 5.2. It is expected that analysis of the results of the controlled evaluation will also show any correlations or relationships that exist between these variables.

CHAPTER 4

STUDY 1: SOUND DESIGN

4.1. Sound design guidelines and methodologies

The field of auditory interface design lacks formal or structured guidelines and methodologies that can be found abundantly in most other fields, such as graphical or visual interface design. Hence the design of auditory cues for interfaces is often based on the personal preferences of the designers, available technology choices, or on an ad hoc basis and is rarely based on theoretically analysed methodologies (Pirhonen & Murphy, 2008). However, some guidelines can be extracted from HCI research for theoretical support to design effective visual-auditory interfaces.

A few brief guidelines that will be kept in mind for designing the sounds are:

- Sounds should be aesthetically pleasing (James, 1998).
- Sounds should be short, interruptible and present relevant information early (James, 1998), (Leplatre & McGregor, 2004).
- The density of a sound or multiple sounds should be limited where possible (Leplatre & McGregor, 2004).

- The sounds should all be aesthetically homogeneous to the interface, although each have to be different to each other to convey appropriate information (Leplatre & McGregor, 2004).
- The total number of sounds should be kept to a minimum (James, 1998).
- Multiple sounds overlaying each other should be avoided as much as possible (James, 1998).
- Commonly occurring events should be designed with less obtrusive sounds (James, 1998).

An extract from Walker and Kramer (2004, p. 149) which summarizes the approach to auditory display design by suggesting the division into three steps provides three valuable guidelines:

> One approach to this topic is to focus on the task of interacting with an auditory display to extract the meaning the designer intended. In this task-oriented approach, we can consider three general types of subtasks, each of which depends on the psychoacoustics research community for design recommendations. First, there is the simple perception of the sounds in the listening environment. If you cannot hear it or discern how it is changing, you cannot extract meaning from it. Second, there is the subtask of parsing the auditory scene into sound sources, or streams and distinct variables such as pitch, loudness, and a host of timbral variables such as brightness or woodiness. Finally, there is the subtask of associative and cognitive processing that result in deriving meaning from the sounds in line with what the sound designer intended. All of these stages are required for successful communication of information via an auditory display.

The above quote is particularly relevant since this research will require designing sounds that convey structural information, such as hierarchical depth. This will require the design of families of earcons or auditory icons in which the change in structural information is apparent, such as a change in depth level. The first guideline suggests using changes that are discernable to the human ear, be it in loudness, pitch or any other parameter. The above guidelines also suggest that whichever sounds are used to represent a function should be meaningful to the users in the same way as it is to the designers.

One of the goals of this research is to use a user-centred design approach. As such, a semiotic design method was selected for the sound design process, which included end-users during the entire process. This corresponds to the first research sub-question posed in Section 1.3. *Participatory design* is a set of theories, practices, and studies that view end-users as full participants in activities leading to the design of software and hardware computer products and computer-based activities (Muller, 2003; Schuler & Namioka, 1993). In this research, the participation of users would lead to the design and selection of sounds that are expected to be closer to what end-users would prefer and be able to interpret in meaningful ways, rather than basing them solely on the designer's ideas.

Other design methodologies that involve end-users as collaborative co-designers include Co-operative Inquiry (Heron, 1996), Informant Design (Scaife, Rogers, Aldrich, & Davies, 1997), Learner-Centered Design (Quintana, Carra, Krajcik, & Soloway, 2001) and Contextual Design (Holtzblatt, 2001). The thread of commonality amongst these methodologies is that each involves participants or potential users during the design stage in different ways. These methods have been commonly used in the fields of computer interfaces, information systems, digital libraries and learning technologies. However, as previously observed, the availability of structured, theoretically-analyzed, and formally evaluated design methodologies for sounds is sparse. There has been no such sound design method found in the literature that follow the guidelines of the above methodologies specifically for the design of sounds. However, Murphy, Pirhonen, McAllister, and Yu (2006) developed a participatory design method, with a theoretical framework strongly based in both structural and musical semiotics [1].

Hence, this design approach was specifically selected to be used here. It was proposed by (Murphy, 2007; Murphy et al., 2006) and formally evaluated by applying the method on a web browser plugin for visually-impaired users (more details on the application can be found in Section 2.2.4.4). They developed a method for sound design in which the wider context of the sounds to be designed are taken into account, since the context of use is considered as important as the individual sound properties (Murphy, 2007). The basis of their sound design method is the "rich use scenario" that is presented to a panel of participants, comparable to a story or a radio play of a character using the system. This method is akin to the traditional use scenario methods used in HCI, in which a use scenario is defined as a formal description of the use an application (Carroll, 2000). However, in this approach the members of a design panel can use the story as a tool to trigger creative sound ideas, instead of being concerned with technical application details (Section 4.2).

Erickson (1995) had previously outlined methods in which stories can be used as a tool for designing human-computer interactions, since they provide a natural way of initiating and continuing dialog with users. Barrass (1996) also explored the use of stories to design auditory interfaces. As Barrass describes, stories can convey principles and methods in a concrete manner which is easy to understand, assimilate and emulate, hence facilitating knowledge transfer through examples, in this case, between the user and the designer. Franinovic, Hug, and Visell (2007) explored sonic interaction design of everyday sounding objects using a participatory

[1] *semiotics* is defined as the "the study of signs".

approach. Moffatt, McGrenere, Purves, and Klawe (2004) used a participatory design method to design an electronic daily planner, which was enhanced with images and sounds, for speech and language impaired users, although not through the use of stories.

The following sections describe the specific case study and the outcomes of the iterative design panel sessions. Recommendations are then derived on how to modify and extend the design methodology for designing auditory cues for familiar or novel user interfaces to effectively convey structural information. This extended methodology and the results can also be referred to in (Absar & Guastavino, 2011).

4.2. The design methodology

The focal point of the methodology is defined as a rich use scenario describing a unique character and his surrounding environment (Murphy et al., 2006). Use scenarios (formal descriptions of the usage of an application) have been used in the field of HCI where they are presented to groups of users or designers (Bodker, 1999). Pirhonen and Murphy (2008) suggest using panels of four or five members, who do not need to be experts in sound design, nor familiar with the application or its usage. Panellists were used to describe and evaluate non-speech sound functions for different task descriptions in three iterative design sessions. The methodology has been followed here to create a rich use scenario, describing a person in an unique situation using the application; in this case, we described a university student using the information system for an assignment. Gaps occur at appropriate points in the story, where the panellists are asked to suggest sound ideas. A brief outline of the steps for the design method described in Pirhonen and Murphy (2008) with some modifications for this study are given below:

(i) A task description for the sound functions of the application is prepared.

(ii) A user description based on a vision of a plausible user is prepared.

(iii) Based on the previous steps, a short story in which the interaction among the character and application plays an important role is written, from the perspective of the character. In the story, blanks are left for the sounds to be designed.

(iv) A design panel session with four or five panellists is organized. The session is started by reading the use scenario, keeping a brief pause in the place of each sound. Having read the story, it is discussed. Then the story is read again with the sentence that includes a blank space for a sound effect. The panellists are asked to try and describe what kind of sound would be appropriate. This is repeated for each sound. The entire session is recorded.

(v) The panels ideas of the appropriate sounds are implemented.

(vi) A second session with different panellists is organized. The implemented sounds are used when reading the story. Screenshots of the system are shown to the panel after an initial reading of the use scenario to help them form an idea about what the system described in the use scenario looks like. All other steps remain the same as in the first session.

(vii) The reactions and new ideas of the second session are analyzed. The original sounds are modified and new ones are created as suggested by the second panel.

(viii) A third session with a different set of panellists is organized. Using sounds from stage 7, the use scenario is presented to the third panel and participants are asked to choose sounds at relevant points in the story. After the initial reading, a demonstration of the system is presented, illustrating the task descriptions and animations at relevant points in the story.

A description of the system, the sound functions and the design process follows.

4.3. The system and sound functions

A 3D information visualization system was selected as the baseline system for this study. One of the reasons for this was that sonification acts as a natural next step to visualization, by spreading out information over the human perceptual system (as mentioned in Section 2.1). The idea was to move away from traditional text-based systems, so as to be able to create a more immersive and engaging experience via visual and auditory feedback. Since one of the aims of this study is to investigate the effect of sounds on hedonic factors of user experience, including user engagement, an information visualization rather than a more traditional utilitarian information system was more appropriate.

Most studies comparing information visualizations with textual displays do show that the main significant differences found are in user preference and affective reactions, rather than performance measures (Large, Beheshti, Tabatabaei, & Nesset, 2009; Morse, 2002; Sutcliffe, Ennis, & Hu, 2000). Similarly, the 3D visualization system used in this study was based on the McGill University Library online catalogue databases for Science and designed by Julien, Guastavino, Bouthillier, and Leide (2010). The visualization system has been evaluated using a controlled comparative study with a traditional, text-based system (Julien, 2010). The system was found to be superior in terms of such hedonic factors, since it was significantly preferred as well as perceived as more useful than the text-based system. More results of this study as well as detailed descriptions of this system can be found in Section 5.2.1. A screenshot of the hierarchical information system can also be seen in Figure 5.1.

To briefly describe here, the interactive interface of the system integrates the searching and browsing of large category hierarchies with their associated documents using a visual representation of the semantic hierarchy. Each node of the hierarchical tree represents an area or subject. Users can visually inspect subjects on labels hovering over circular areas in each node and can traverse the hierarchy by following branches linked to each subject area. The interface also includes 3D animations to represent the depth and width of the tree. The animations allow the user to zoom into a desired node, or fly around the tree to see different perspectives and help users acquire a mental model of the semantic relationships between the contents of the tree. Adding auditory feedback that complements and reinforces the visual information may further aid in navigating the structure. Descriptions of the various sound functions designed in the course of this study are described below:

- *Sound to distinguish the difference between end nodes and internal nodes*: This information is important for searchers since it signifies where to stop searching in a specific subject area and move to another. Visually internal nodes often look like end nodes, since the branches below are not always shown. The difference is, however, represented by a slight difference in color, which is often difficult to distinguish in a large category hierarchy. To enhance this difference with auditory cues, a distinctive sound is required to differentiate end of branch nodes from internal nodes.
- *An overview sound to display the density of information of a node*: Each node has a number of subjects or text collections under it, which can be an important marker for searchers if discerned quickly. A sound that gives an overview of the density of information under that node would be useful in this way. Shneiderman's (1996) information seeking theory gave rise to the idea of generating an overview before navigation. Several studies have supported the use of nonspeech sounds to display overview information (Kildal & Brewster, 2006; Murphy, 2007).
- *Sound to indicate the hierarchy level number or depth*: Since the hierarchy is so large and spread out, it is often difficult to keep track of the level

of depth traversed by the user. An auditory cue can indicate the relative depth level of a certain node. Brewster et al. (1996) showed that nonspeech audio earcons provide an effective way of representing depth in hierarchical menu structures.

- *Sound to display the selection of a node*: This is an auditory reinforcement of the visual selection of a node to show the contents of that node.
- *Sound to reinforce flying to a node*: This is a supplementary sound to reinforce the visual animation of zooming or flying into a subject area.

4.4. The experiment: design panels and results

4.4.1. Rich use scenario. The use scenario for this study was written from the perspective of a university student preparing for an assignment using an information visualization system. The scenario described a young university student, his mood and his surroundings.

> Sam had an assignment due in two days and he still hadn't touched it, as usual. He ignored the inviting sounds coming from outside his window of people enjoying a sunny day, and decided to procrastinate no further. He logged on to his favorite information search system - he preferred to use a visualization system rather than an online library catalogue for his assignments.

The scenario also briefly described the system he was using. The design methodology (Pirhonen & Murphy, 2008) suggested that overly detailed descriptions of the technical aspects should be avoided, so as not to hamper the creative process. The system was based on the library catalogue for Science and Engineering faculties, thus the descriptions were targeted for panellists who would be students in Science or Engineering. Hence, the description of the hierarchy looking like a "tree" was used to simplify the explanation and make it sound less "technical" as suggested.

Recommendations are, however, later made in Section 4.5 regarding the choice of words.

> The system looked like a top-down tree, in which each of the leaves of the tree were a subject area. He could browse the tree by subject to find areas of interest or even search for any item he wanted.

It went on to describe the activities he carried out on the system, with gaps in the regions of the story that required sound effects.

> He started browsing through the tree, by moving the mouse over the labels of the tree leaves to find anything that caught his interest. When he found something he wanted to look at, he clicked on it [sound 1]. There was an animation that showed the zooming in or flying into that part of the tree accompanied by [sound 2]. From this subject, he went deeper and deeper into the tree by clicking on subjects branching out from above ones. The deeper he went, he still could keep track of his depth, as there was a sound that showed him his relative depth level [sound 3]. And once he reached the end of a branch (no more subjects under that) he would hear a specific sound [sound 4] that clearly told him he had come to the end of that particular branch.
>
> He also liked the option of being able to right-click on a subject area and hearing [sound 5] that told him the density of information present under that area (such as how many books and papers are available under the subject).

4.4.2. Panellists. The design methodology (Pirhonen & Murphy, 2008) suggested using an iterative sequence of three panels, with each panel consisting of a different set of four to five members. Hence, thirteen panellists were recruited in total for the three design panels, with two panels consisting of four panellists, and one panel consisting of five. The demographics and criteria for the panellists are given below. They were:

- aged between 18 to 35 years (mean age 24.1, standard deviation 3.2);
- required to have unimpaired hearing;
- fluent English speakers;
- University graduate students in, or with a background in, Science or Engineering;
- comprised of 6 males and 7 females, with each panel consisting of at least one panellist in each gender;
- at least one panellist in each panel had experience or knowledge in information systems;
- two panellists in each panel had at least 5 years of formal musical training or sound design experience;

The exact composition of the panels are given in the following sections describing each of the three panels.

4.4.3. User panel 1. The first panel session was made up of four graduate students, who were two females and two males. Two of them had experience in sound engineering or music and one had experience in information systems design. They were presented with the use scenario, with spaces at relevant points in the story where they were asked to input their sound ideas. In the use scenario, the hierarchical system was described as a top-down tree, with labels on leaves, which may have lead to the groups identification of forest sounds as a metaphor for auditory cues.

Based on the sound ideas described by the panel in Table 4.1, sounds were selected from the McGill Multimodal Interaction Lab (MIL) audio resources and

TABLE 4.1. User panel 1 reactions

Task description	**Sound description suggestions**
Clicking on a label	Generic clicking sound
Zooming or flying into a part of the tree	A whooshing sound
Indication of current depth level in tree	Ticking sound, quicker with increased depth (e.g. a Geiger counter). However they realized it would be annoying.
	Change in pitch: deeper levels should have lower frequency.
	Ambient sounds of leaves rustling, more intense the deeper the level.
	Opposite idea: Leaves rustling with less intensity for increased depth coupled with birds chirping, with louder chirping for more depth.
Indication of reaching a leaf (end of branch)	Generic error sounds
	Sound of a lock, a thud or a door closing
	A distinctive bird chirp
Overview of the density of information of a subject	An applauding crowd, louder with more information (again they realized it would be annoying)
	A falling thud: heavier thuds for denser information
	Page flipping: small number of pages for less dense information, flipping through a book for more information.

online open-source sound resources (*The Freesound Project*, n.d.), all recorded at 44.1 kHz with a bit-depth of 16 or 24 bits.

- Sound 1: 4 clicking sounds were selected: 2 single clicks and 2 double-clicks.
- Sound 2: 2 whooshing sounds were selected and edited to different lengths to match animation times by changing the speed or tempo.

- Sound 3: 3 options were selected: 1) 3 choices for birds at 3 depth levels, achieved by choosing parts of the sound files where the birds sounded more intense for deeper levels. The difference in average intensity level between each depth level was scaled to 6 dB. 2) The same procedure was followed for 2 choices in wind or leaves rustling. 3) 2 choices were given for the combined sounds of birds and leaves rustling. As the depth level increased, the sound of leaves rustling was reduced by 6 dB while that of the birds were increased by 6 dB in each step.
- Sound 4: 2 options were selected and scaled to the same RMS (root-mean square) level: 1) 5 short distinctive bird chirps 2) 2 choices of the sound of a door or lock.
- Sound 5: 2 options were selected: 1) The sound of a book dropping was scaled up in 3 incremental files with 6 dB steps. 2) The sound of page flipping was edited to create 5 levels: the first level was one page flipping, the second of two pages, and the fifth of several pages flipping rapidly.

4.4.4. User panel 2. Panel 2 consisted of five graduate students, two males and three females. Two of them had more than 5 years formal musical training, and two had experience in information system design. The implemented sound ideas from Panel 1 were presented to Panel 2 at the relevant parts of the use scenario. They discussed these sounds, their preferences, or came up with new sound ideas. Panel 2 was also presented with system screenshots, to allow them to visualize the system described in the use scenario.

Based on Panel 2's sound suggestions in Table 4.2, the original sounds from Panel 1 were modified and additional sounds selected from the MIL audio resources and online resources [15], all recorded at 44.1 kHz with a bit-depth of 16 or 24 bits.

TABLE 4.2. User panel 2 reactions

Task description	**Proposed sounds**	**Sound suggestions and reactions**
Clicking on a label	Clicking sound	2 of 4 variations of single or double clicks preferred by all panellists.
Zooming or flying into a part of the tree	Whooshing sound	Most thought the whooshing sounds were too abrasive; less rough whooshes would be preferred, especially for the longer sounds.
	Watery sounds	Suggested more pleasant sounds like a running brook might work.
Indication of current depth level in tree	Just leaves or wind	Sounds too harsh or stressful at deeper levels. They would not want it to sound stormy when deep in their search.
	Just birds	They preferred it more than just the leaves rustling.
	Combination	They all preferred the combined sounds more; they agreed it was easier to tell the difference in cues.
	Tonal sound	One panellist said she would prefer a tonal sound that changes in pitch to show level changes (higher pitches for higher levels)
Indication of reaching a leaf (end of branch)	Bird chirp	Most thought the chirps would not be easy to distinguish if paired with bird sounds in Depth sounds.
	Thud	They agreed a thud would work better, and liked the sound of a book dropping as a cue here.
Overview of the density of information of a subject	Book dropping or thud	They did not like book dropping volume changes as a cue.
	Page flips	They liked the page-flipping cue, but mentioned the sounds should be the same length, with faster flipping for more pages.

- Sound 1: 3 clicks were selected, by eliminating one sound the panel did not like.
- Sound 2: One whooshing sound remained. A new swishing sound and two new watery sounds were selected.

- Sound 3: The combined sounds of birds and leaves rustling remained. Piano notes, acquired from the IOWA music database (*The University of Iowa music instrument samples database*, n.d.) were added (pitch change of notes from C4 to B4, or F1 to F7 represent depth levels).
- Sound 4: The sound of a book dropping or thud was selected.
- Sound 5: The page flipping option was edited so that the length of each file was almost the same by increasing the speed of the page flips for the longer sounds.

4.4.5. User panel 3. Panel 3 consisted of four graduate students, two male and two female. Two of them had the required musical training, while two had none; one of them had experience with information systems design. The third panel was presented with the modified sounds from Panel 2 and they discussed their preferences for these sound options. Panel 3 was also presented with a demonstration of the system and the animations mentioned in the use scenario, as opposed to screenshots shown to Panel 2. This allowed them to match the sounds with the task descriptions more effectively.

From the above reactions described in Table 4.3, Sound 1 (subject selection) was selected to be the single-click. The new swishing sounds of variable lengths were selected for Sound 2 (the zooming function). The combined birds and leaves rustling, with increased sound of birds and decreased sound of wind for increased depth, was selected for Sound 3 (hierarchical depth level indication). The sound of a book dropping was selected for Sound 4 (end of branch distinction). And the sound of pages flipping was selected as the final overview sound for Sound 5 (information density sonic overview).

After the completion of the panel sessions, natural and synthesized sounds were used and interpolated to create multiple levels in the hierarchy to convey structural

Table 4.3. User panel 3 reactions

Task description	**Proposed sounds**	**Sound suggestions and reactions**
Clicking on a label	Single click	General consensus: single click preferred over the double-click.
Zooming or flying into a part of the tree	Swishing sound	Three panellists liked the new swishing sounds.
	Watery sounds	They thought these were not appropriate in the context, swishing sounds match better with the visuals.
Indication of current depth level in tree	Woods sounds	Woods sounds preferred over to the piano (easier to distinguish and match with the system visuals).
	Piano notes	One panellist with piano training preferred the pianon notes.
		All panellists agreed that for the piano, the more easily distinguishable notes were the F1 to F7 notes, rather than C4 to B4.
Indication of reaching a leaf (end of branch)	Thud	They all agreed the book dropping sound worked well as a cue here.
Overview of the density of information of a subject	Page flips	They all liked the page flipping cue, and agreed they would easily be able to distinguish the changes in information density with this cue.

information based on the results of the design process. Seven levels were created for the hierarchical depth cues as well as seven different levels for the information density overview cues. These sounds were then integrated into the hierarchical system and tested and debugged until system stability was reached. The next step of conducting a controlled evaluation of the system on users for performance effects and affective reactions will help to further verify the effectiveness of the sound design methodology in the current context of hierarchical information systems.

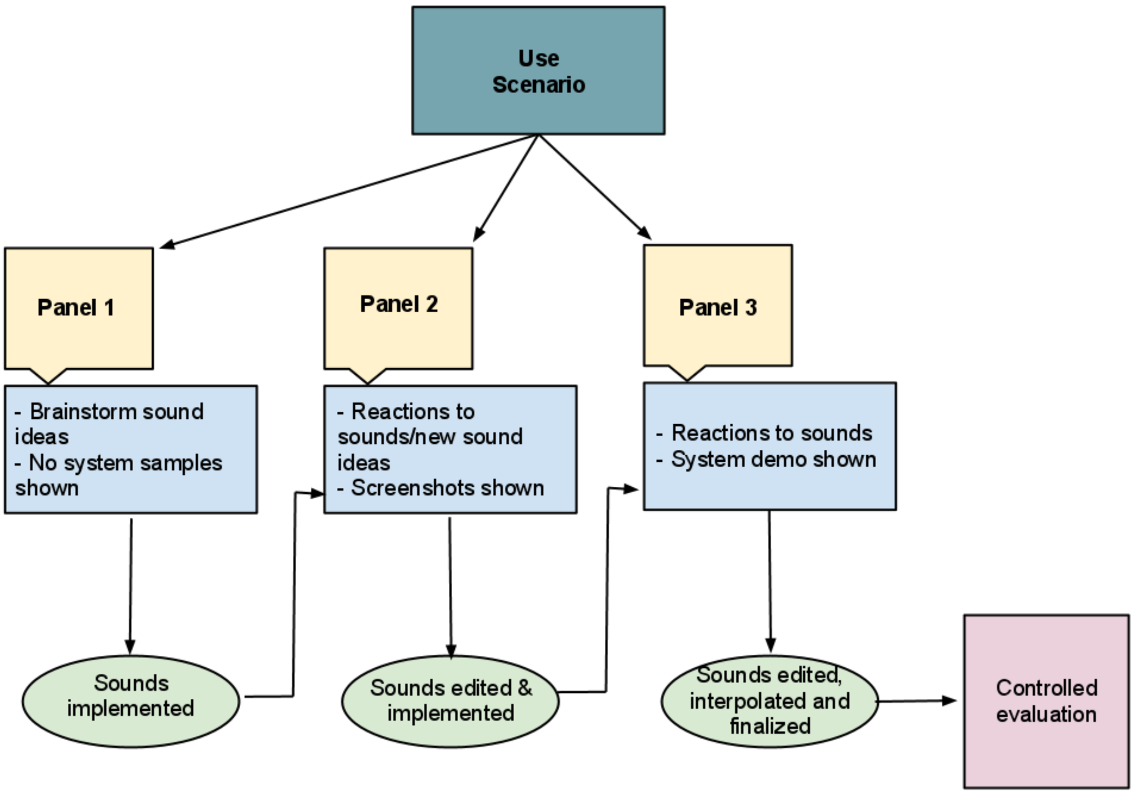


FIGURE 4.1. Overview of the sound design process

Figure 4.1 outlines the entire process of sound design which was followed in this study. A summary of the observations and findings in the study as a result of the sound design process follows.

Summary of observations and findings:

- The user panels help in group confirmation of the design. Creative ideas were successfully generated through group discussion.
- The use scenario helps to trigger creative sound ideas. A story that is easy to connect to for the panel members helps in discussion initiation.

- If the system in the use scenario is not fairly familiar to the panel members, displays or demonstrations of the system should be shown.
- Words in the use scenario to describe the system have to be carefully selected so that the description does not lead the panel members to visualize an inaccurate version of the system and hence result in inappropriate sound choices in the required context.
- The iterative sessions with different panel members help to identify problems throughout the sound design process and lead to more creative input to be processed for each sound.
- Panellists have an almost immediate negative response to sounds that are harsh, abrasive, highly reverberatory, loud, long or busy (attention demanding).
- Auditory icons (environmental or real-life sounds) seem to be generally preferred than earcons (abstract or musical sounds). An exception arises when the panellist is musically trained in the specific musical instrument used in the earcons.

4.5. Discussion

Pirhonen and Murphy (2008) recommended that system specifics should not be described in too much detail to the panel, since it could hamper the creative process. The use scenario is meant to generate creative input rather than focus discussion on the details of the system. In keeping with this suggestion, in the first panel session, the only description of the system given was that in the use scenario; no examples of the application were shown. However, since they were entirely unfamiliar with the novel system, this led to some confusion during the session, as panellists had a difficult time trying to visualize the system, task descriptions and animations described in the use scenario. Before being able to generate creative input, some time had to be spent in answering panellists questions and trying to elaborate on the task descriptions. Hence, instead of less technical details hampering the creative process, it almost had the opposite effect.

Thus, in the second panel session, screenshots of the system were shown after the use scenario was introduced. This led to better understanding of the system, but questions still remained on the animations and the 3D view of the system. Consequently for the third panel, a demonstration of the system and the described animations in the task descriptions was shown. This yielded much better results in the session, consensus was reached much faster, and the group was able to match sounds with the animations much more effectively.

Therefore, while it is feasible to not delve too deeply into the details of the application in environments familiar to users, such as webpage-browsing, as was investigated in (Murphy et al., 2006; Pirhonen & Murphy, 2008), this is not the case in novel systems. In cases where the application is entirely new and unfamiliar, it is necessary to tailor the rich use scenario to allow the panel to be able to visualize the application, while still triggering creative sound ideas by including inspiring details about everyday life.

Another observation was that the way the use scenario is written can influence the type of feedback or sound ideas the panel comes up with. For example, in the use scenario, the words trees and leaves were used to describe the hierarchical system and this may have influenced the panel to lean towards forest-like sounds of leaves rustling and birds chirping. However, when the second panel was shown the screenshots of the hierarchical system, they commented that it did not look like a tree or forest, and did not match the metaphor. The third panel however thought the animations matched the forest feel when shown the demonstration of the system, and hence favored the forest metaphor. Therefore, the way a system is described, or the specific words used to describe the system should be selected carefully, so that it does not lead the panellists to have an inaccurate mental vision of the system and consequently, result in inappropriate sound choices or metaphor identification in the required context.

Thus, depending on the type of application in which auditory feedback is being designed, it is suggested that the use scenario be tailored so as to give the panel an overall feel of the system, while still keeping the essence of a story or a radio play. This is so that, when the scenario is presented, the members of the design panel can use the story to generate creative sound ideas, while not being concerned with parts of the application that they do not understand. Hence, for fairly familiar applications such as web browsers or file systems, it may not be necessary to show them any instances of the application. For 2D applications that have functions that may be difficult to explain in the use scenario, a few screenshots of the system can be shown after the use scenario has been presented and initially discussed. This illustration may bring about new sound ideas from the panel. However for completely new or multimodal applications, with 3D graphics or animations, a demonstration of the system starting from the first panel session is suggested. This would reduce confusion and address questions the panellists may have regarding the task descriptions in the use scenario and should help facilitate the creative process.

CHAPTER 5

STUDY 2: USER EVALUATIONS

5.1. Formative evaluation

After the integration of the finalized sounds from the sound design phase (described in Chapter 4) into the visualization system, an initial formative evaluation was carried out on the integrated system. The main objectives for this study was to:

- test the stability of the overall system integrated with sounds and the web-based answering system that was developed for the experiment,
- acquire user impressions of the system with sounds and evaluate if any changes need to be made to the sounds, such as the lengths, loudness, or any other parameters.
- see if any modifications need to be made to the tasks being tested
- gather if the training provided is adequate
- examine the length of time it takes to carry out the subset of tasks selected for the study, so that the total time required for the main experiment can be judged.

5.1.1. Participants. Five participants, three females and two males, aged between 18 to 30, were recruited for the study, who were all students from Science or Engineering faculties. This was a requirement since the visualization system is based on a Library Catalogue containing Science and Engineering subjects. A familiarity with the subject matter may decrease the learning time for the system and increase participant interest and motivation in the tasks. The other criteria used for recruitment were that they had to be fluent English speakers with no known hearing impairments.

5.1.2. Procedure. Participants were informed that they would be carrying out information seeking tasks on a system using both visual and audio cues. The audio cues were played over speakers. They were also told that a "think-aloud" protocol will be used in which they should describe all that they think or do by elaborating each step they take while carrying out the tasks, as well as any impressions or confusions they have during the study (Fonteyn, Kuipers, & Grobe, 1993). This was to gain maximum feedback from each participant on any issue relating to the system. After the training and the tasks, they will be asked a few questions about their experience in an semi-structured interview. The entire procedure for each participant took about one hour in total and was audio-recorded for further analysis.

5.1.2.1. *Training.* At the beginning of the training session, they were shown a demonstration of the system, illustrating the visualization and how it works. The functions of each visual or audio cue were explained during this demonstration. The outline of the protocol used during this training session is attached in Appendix B. After this, four sample training tasks were carried out by the participant. During this time, the experimenter helped out with any questions or confusion, and prompted the participant to encourage the think-aloud vocalization. The entire training session took about 20 minutes.

5.1.2.2. *Tasks.* A subset of the tasks to be used in the main experiment were tested in the formative evaluation. The tasks are shown in Table 5.1. They consist of hierarchy navigation tasks (Tasks 1, 2), information density comparison (Tasks 3), depth level comparison (Tasks 4), and simple and complex information retrieval tasks (Tasks 5 and 6 respectively).

TABLE 5.1. Tasks tested in Formative Evaluation

Task number	**Task type**	**Task description**
Task 1	Hierarchy navigation	Find a subject which directly belongs to "Operations Research"
Task 2	Subject relations assessment	Find the nearest common most general subject of both "Steelwork" and "Mechanics"
Task 3	Overview	Which of these subjects contains more documents? "Matter, Properties" or "Optical transducers"
Task 4	Depth assessment	Which of these subjects is at a deeper level in the hierarchy? "Power transmission" or "Detergents"
Task 5	Simple retrieval	You are looking for information on Visual Basic programming. Find two promising books for this question.
Task 6	Complex retrieval	How would you go about fixing a broken computer? Find two promising books for this question.

5.1.2.3. *Interview.* After completion of the tasks, the participants were asked a few informal questions. An outline of the questions asked is also given in Appendix B.

5.1.3. Results and conclusions. After each of the study sessions with a participant, the audio recording of the think-aloud procedure and the interview were analysed to find relevant information to edit or improve any part of the system or experimental method. Relevant parts of the audio-recording for each participant were noted down, and summaries of the reactions noted during the procedure can be found in Appendix B. After each session, changes were incorporated and the edited system or method was presented to the next participant. This was continued until no more suggestions of changes or improvements were made. The following changes were those that were made progressively throughout the entire formative study:

- Participant 1 suggested that the zoom sound need not be as conspicuous as the other sounds and "might be distracting if heard repeatedly". Since this cue is only a form of reinforcement of the visual cue, the intensity level of the zoom sound was reduced 3 dB quieter to the previous level for the corresponding sessions.
- Initially, Questions 3 and 4 (Table 5.1) of the task set consisted of comparing the depth levels or information density levels of four subjects. Participant 1 said, "It was hard to remember the sounds when the tasks involved 4 subjects". This meant that one would mostly use the visual cues for comparison, rather than the audio cues, since auditory memory retention is lost more easily. Hence, for subsequent sessions, the number of subjects in these tasks were reduced to two rather than four, to make easier auditory feedback comparisons.
- Participant 2 had a difficult time with Task 2. This was because in the initial format of the question, both the subjects, Steelwork and Automatic Control, were near the bottom of the tree, making it more difficult for the participants to see their link. For the next sessions, one of the subjects in Question 2 was changed, by replacing "Automatic control" with "Mechanics".
- Participant 2 also suggested that it would be helpful if they were given more time to familiarize themselves with the system before starting the training tasks. Hence, a few minutes (about 5 to 7 minutes) were given to each participant to familiarize themselves with the system after the

demonstration, to get used to the system functions before carrying out the training set. All subsequent participants appreciated this time given to them.

- In the initial format of the demonstration, first the visualization was explained and later the sounds were explained. Participant 2 suggested that the sounds should be explained as they come about in the demonstration, and not left till later, since it is easier to absorb then.
- Participant 3 had a hard time understanding Question 2 in the way it was initially worded. Hence, the wording was subsequently changed from "Find the nearest common more general subject of Steelwork and Mechanics" to "Find the nearest common subject that both Steelwork and Mechanics belong to." This seemed to make the understanding of the task much clearer to subsequent participants.

After Participant 3, no more suggestions of changes were given by the next two participants or no more new ideas emerged from the interviews, and it was concluded that data saturation point was reached, and no more participants were tested (Oppenheim, 1992). Table 5.2 below summarizes all the changes made to the system or the protocol of the study, due to the affective reactions. The detailed responses of the participants are attached in Appendix B.

Overall, the participants had positive reactions towards the sounds and said that the mapping of the sounds were appropriate to the various functions. The only sound cue they could not understand upon initial hearing was the auditory feedback denoting the relative depth level of the subject. However, they agreed that as soon as the meaning was explained, it was easy to learn and remember. Most of the participants agreed that given the choice between using the system with or without sounds, they would prefer using the system with the sounds.

Table 5.2. Changes made over Formative Evaluation

Participant impressions	Changes made to the system number	Changes made to the procedure
"Zoom sound might be distracting if heard repeatedly during tasks"	Intensity level of zoom sound reduced by 3 dB	
"It was hard to remember the sounds when the tasks involved 4 subjects"		Number of subjects reduced from 4 to 2 in overview and depth assessment tasks
Difficulty completing task		Subject changed in Subject relations assessment task
"It would be helpful if I had more time to play around with the system before starting the training tasks"		Time given after training demonstration to familiarize participants to the system
"It was confusing to hear some sounds during the demonstration and not being told what they meant right then."		Sounds explained as they came up in the demonstration instead of later
Wording was unclear and needed further elaboration before understanding was reached.		Wording of Subject relations assessment task changed

5.2. Comparative evaluation

The controlled comparative experiment involved formally evaluating the effects of integrating the designed auditory cues to the visual information on user navigation in the information system. The primary objective is to examine whether users benefit either quantitatively or qualitatively from the augmentation with auditory cues. This experiment is described in the following sections.

5.2.1. System used. A 3D visualization system, based on McGill University Library and Laval University Library online catalogue databases and their Library of Congress Subject Headings (LCSH) organizations, has been designed and implemented by Julien et al. (2010), and is referred to as the Subject Explorer 3D (or SE-3D). The design is based on the Cat-a-Cone trees described in Chapter 2, Section 2.1 (Hearst & Karadi, 1997).

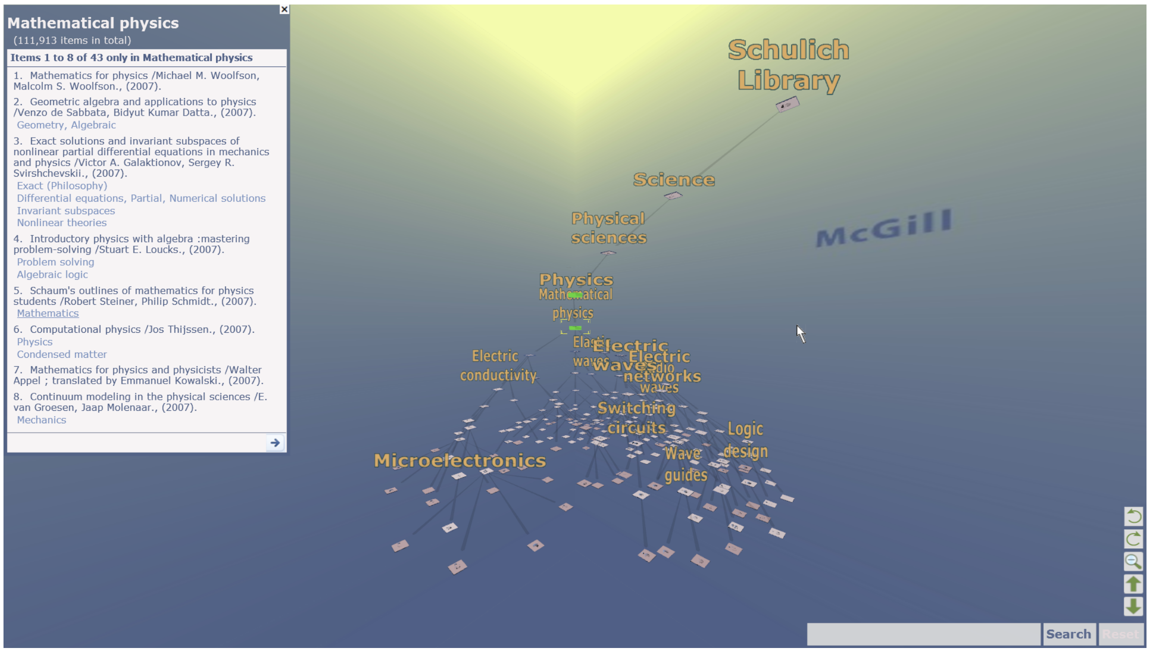


FIGURE 5.1. Screenshot of the SE-3D system

The interactive interface integrates the searching and browsing of large category hierarchies with their associated text collections using a visual representation of the semantic hierarchy. Hence, it is a point-to-move 3D application tool, integrated with keyword searching, which allows users to search and explore a semantic hierarchy and its associated documents by exploring the metaphor of a physical space. Users can visually inspect subjects written on labels that hover over circular areas. The size of the circular areas depend on the number of documents covering the subject. Users can travel up and down the hierarchy of subjects by following branches linked to each subject area. A screenshot of the system is shown in Figure 5.1.

A 3-dimensional hierarchy representation is chosen as opposed to a 2-D one, since 3D hierarchies can make more efficient use of screen space (Robertson et al., 1991), are better suited to convey hierarchical relationships between levels (Bladh & Scholl, 2004), and the third dimension can be used to convey hierarchy depth (van Ham & van Wijk, 2003). Hence, the interface includes 3-dimensional animations to represent the depth and width of the tree. Each node of the hierarchical tree represents an area or field (subject). The animations allow the user to "fly" to any node, "zoom" into a desired node, or fly around the tree to be able to see different views or perspectives of the tree. It is believed that this will help users acquire a mental model of the semantic relationships between the contents of the tree and hence aid in navigation and information retrieval tasks.

This system has been evaluated by comparing to a baseline text-only subject browsing interface (Julien, 2010). It was found that the system provided a significant performance advantage for tasks involving finding the most specific subject, which required repeated evaluations of relations between subjects. This means that the mental model of the semantic structure was formed more accurately for the aforementioned IV system compared to the traditional text interface. It was also found that participants of the study preferred the novel IV system more than the text-based system. However, it is believed that adding a secondary modality such as auditory feedback into the visual interface can further enhance the formation of a mental model of the structure and also improve affective reactions.

One of the reasons this IV system has been chosen as the baseline for the current study is that it is not commercially available yet and previous user experience of the system will be less likely to confound the results. It has been reported that prior knowledge of the baseline system in information visualization studies can put the novel interface at a disadvantage since resultant effects can show the baseline as the favored interface simply because the users were more proficient at using it (Newby, 2002). This effect is aimed to be minimized by using the SE-3D system as the baseline, which will most likely be unfamiliar for the participants of the comparative study.

5.2.2. Comparative evaluation. The 3D visualization system based on a library catalogue described in the previous section was used for the comparative study between the visualization system itself, hereafter referred to as the visual-only system, and that enhanced with the auditory feedback, referred to as the audio-visual system.

The comparative experiment was designed to address the research question posed in Chapter 1 using a repeated measures within subject design. A group of 24 participants were asked to perform equivalent tasks on each of the with and without-audio systems. The goal is to evaluate the effect of the applied auditory enhancement of the system on various performance measures, affective reactions and consequently, the overall user experience.

5.2.3. Participants. This study uses a sample size of 24 participants. *Effect size* or Cohen's d is the most commonly used measure of how much a certain treatment, such as the presence of auditory feedback in this case, affects the dependant variable, which can be the task time, accuracy, or hedonic variables in this study. Cohen (1988) described three different effect sizes, which were small $d = 0.2$, medium $d = 0.5$ or large $d = 0.8$, where the effect for a large d is so large that statistics are often not necessary. *Power* is defined as the probability of rejecting the null hypothesis when it is false. As Kenny (1987) has shown, for the sample size used in this study and a reasonable level of power (such as 0.8), significant differences can be detected assuming a large effect size ($d = 0.8$).

However, this large effect size is sufficient for this study, since in fields like HCI and auditory interface design, the effect or difference caused by the treatment needs to be a relatively large one, that is it should have a large enough impact on user experience, so as to be considered worth pursuing (by making changes to the user interface or adopting new systems). This is often referred to as practical significance, as opposed to statistical significance, where practical significance looks at whether the difference is large enough to be of value in the real world (Kirk, 1996). Hence, comparable sample sizes to the current study are common in typical auditory interface studies (Brewster et al., 1996; Fernstrom & McNamara, 2005; Frauenberger & Stockman, 2006; Garzonis et al., 2009; H. Zhao, Plaisant, Shneiderman, & Duraiswami, 2004).

Furthermore, our study required a number that was a multiple of 8, to allow for the counterbalancing needed between systems and task sets, and hence 24 participants were recruited for the study.

The target population for this study are students from Science or Engineering faculties who are familiar with the subject matter contained in the tested information system. Thus, the 24 participants were between ages 18 to 35, with mean age 23.75 and standard deviation of 4.18, recruited from the University student body. 50% of the participants were male while the other half was female (i.e., 12 males and 12 females). The criteria required for all participants were that they

- be students from Science or Engineering faculties (for familiarity with the subject matter, as mentioned in the formative evaluation)
- have no known form of hearing impairments, and
- be fluent English speakers.

It is planned as a future direction to conduct similar evaluations on user groups with specific prior experience or sensitivities, such as musical training, ear or sound training, video gaming experience, etc. (further described in Section 7.3). Although these demographic data were collected and later analyzed, they were not used to control the sampling process.

5.2.4. Tasks. With the sounds designed and integrated in the system, as described and validated in Section 4.2, the usability of the sonified visualization system can be evaluated by analyzing the performance and experience of participants on a number of information retrieval (IR) tasks. One of the most established taxonomy of tasks that can be used for evaluating information visualization systems was given by Shneiderman (1996), where he gives a list of seven basic tasks:

(i) Overview: The ability to estimate generalized knowledge of the contents of the information collection.
(ii) Zoom: The ability to focus into a more specific area of the search from a more general one.
(iii) Filter: The ability to exclude specified classes of items from the search.
(iv) Details-on-demand: The ability to efficiently inspect details about certain item contents.
(v) Relate: The ability to view relationships among items, such as semantic relationships, co-citations, etc.
(vi) History: The ability to inspect ones own previous search path
(vii) Extract: The ability to obtain and save items for future purposes.

H. Zhao (2005) present an Auditory Information Seeking Principle (AISP) modeled after the visual information seeking mantra described above. They surmise that if information seeking in the auditory modality follows the same pattern as the visual, then the collaboration between visual users and auditory users might become easier. Their proposed principle (AISP) consists of the following :

(i) Gist: Quick grasp of the overall data trends and patterns from a short auditory message.
(ii) Navigate: Fly through the data collection and closely examine portions of interest.

(iii) Filter: Seek data items satisfying certain criteria.

(iv) Details-on-demand: Obtain details of groups or an individual item for comparison.

From the above two taxonomy of tasks, we limit the scope of the tasks to be tested in this experiment to those that are relevant to the auditory feedback here, either by having the option to directly or indirectly use the various auditory cues during the tasks, or by allowing the user to explore the system so as to be exposed to the experience of using all the visual and auditory functions. Two task types were distinguished: hierarchy navigation tasks or information retrieval tasks. The tasks described below are Hierarchy Navigation tasks, which require navigating the information space, either by browsing, searching or a combination of both.

- Subject traversal tasks: These tasks are designed to test the ability of users to effectively navigate through the different hierarchy levels of the tree and acquire information. As such, it involves retaining a mental model of the structure of the information. These tasks are related to the "zoom", "details-on-demand", and "relate" task types from Schneiderman's taxonomy and the AISP. Examples of specific task descriptions are:
  (i) Find a subject directly belonging to "electric conductivity"
  (ii) Which subject does the subject area "C-programming" directly belong to?
- Subject relations assessment tasks: These tasks evaluate the ability to view connections between related items, such as semantic relationships. These tasks are related to the "zoom" and "relate" task types described in Schneiderman's taxonomy. Examples of specific task descriptions are as follows:

(i) Find the closest common subject under which "Bio-informatics" and "Electronics" belong.

(ii) Find the closest common subject which belongs to both "Bio-informatics" and "Electronics".

- Depth assessment tasks: These tasks evaluate the ability to assess the hierarchical depth traversed or compare the depth in the hierarchy between different subjects. These tasks are related to the "details-on-demand" task types from Schneiderman's taxonomy and the AISP. Examples of specific task descriptions are as follows:

  (i) How deep in the hierarchy is the subject "C-programming" in relation to the subject "Physical Sciences"?

- Overview tasks: These tasks are designed to evaluate the ability to assess the density of content information at a certain node or of the whole subject area in general. These tasks are related to the "overview" and "gist" tasks described in Schneiderman's taxonomy and the AISP. Specific task descriptions are as follows:

  (i) Which of the two subjects "electric conductivity" or "C-programming" contain more documents?

Two *information retrieval* tasks were also included, which were search tasks of known or unknown items, referred to as simple retrieval and complex retrieval tasks. Simple retrieval tasks contained terms in the question that matched the target subject, whereas the complex retrieval tasks did not (Pirolli, Card, & van der Wege, 2000; Yee, Swearingen, Li, & Hearst, 2003). Although neither of these tasks directly require the use of the auditory feedback, they are good ways to expose the user to all the different types of sounds that are present in the system while performing the tasks.

Two separate task sets were created for each task category, one equivalent to the other, designed for each condition - the baseline visual-only system and the audio-visual one. The order of presentation of the task sets as well as the visual-only and audio-visual condition were counterbalanced for all participants. The two training and task sets are presented in Tables 5.3 and 5.4.

The equivalence of the two task sets was ensured by selecting subjects that were at the same depth level for each of the hierarchy navigation and relations assessment tasks, and the same difference in depth levels or information density in the depth assessment and overview tasks. The training task set is a subset of four tasks out of the six tasks given in the actual task sets. Both the depth assessment and overview tasks are comparison tasks, so only the overview task was selected for the training. Again, out of the search tasks, only the simple retrieval task was selected for the training, as this displayed the basics of how to conduct keyword searches using the system.

TABLE 5.3. Training Tasks used in comparative evaluation

Task type	**Training Set 1**	**Training Set 2**
Hierarchy navigation	Find a subject which directly belongs to “Electrical Engineering”.	Find a subject which directly belongs to “Mechanical Engineering”.
Subject relations assessment	Find the nearest common subject that both “Electric Motors” and “Remote Control” belong to.	Find the nearest common subject that both “Electric current converters” and “Transducers” belong to.
Overview	Which of these subjects contains more documents? “Dynamics” or “Physical Sciences”.	Which of these subjects contains more documents? “Fluid dynamics” or “Science”.
Simple retrieval	Find what you think would be a promising book on land surveying.	Find what you think would be a promising book on bridge design.

TABLE 5.4. Tasks tested in comparative evaluation

Task type	**Task Set 1**	**Task Set 2**
Hierarchy navigation	Find a subject which directly belongs to "Operations Research".	Find a subject which directly belongs to "Mathematical Physics".
Subject relations assessment	Find the nearest common subject that both "Steelwork" and "Mechanics" belong to.	Find the nearest common subject that both "Electric Motors, Linear" and "Nuclear Physics" belong to.
Overview	Which of these subjects contains more documents? "Matter, Properties" or "Optical transducers".	Which of these subjects contains more documents? "Physical measurements" or "Dynamics".
Depth assessment	Which of these subjects is at a deeper level in the hierarchy? "Power transmission" and "Detergents"	Which of these subjects is at a deeper level in the hierarchy? "Flight Control" and "Wave motion, theory of "
Simple retrieval	You are looking for information on Visual Basic programming. Find 2 promising books for this question.	You are looking for information on internet security. Find 2 promising books for this question.
Complex retrieval	How would you go about fixing a broken computer? Find 2 promising books for this question.	How can you fix a leaky faucet? Find 2 promising books for this question.

5.2.5. Independent variables. There are two conditions in which the tasks were presented: the baseline condition is the purely visual condition of the visualization system without any auditory feedback. This is compared to the combined visual and audio condition, where the visualization has been coupled with auditory cues. Hence, there are two independent variables: one is the system condition which takes 2 values (visual-only and audio-visual); the other is the task at hand which takes 6 possibles values, listed in Table 5.4.

5.2.6. Dependent variables. As posed in the research question in Chapter 1, the objective of the experiment is to evaluate how the addition of auditory cues to support navigation tasks in a hierarchical information system affects the user experience. As defined in Chapter 1, the term user experience encapsulates the entire set of affects that is elicited by the interaction between the user and the information system. This includes *performance*, that is, any performance benefits in using the system, as well as *affect*, that is, any benefits perceived by the user. In the context of information systems, Kourouthanassis, Giaglis, and Vrechopoulos (2007, p. 319) have aptly described user experience as the following:

> In short, the user experience may be viewed as a sum of momentary constructions that grow from the interaction of users with their environments. These constructions may be affected by several strands that include, but are not limited to, compositional, sensory, emotional, spatio-temporal, and interaction-based factors (Battarbee & Koskinen, 2005). In the IS context, the user experience is mostly generated through the interplay of interactions between the system and the user. Depending on design factors such as the user interface, the navigation structure, .. a user may evoke positive experiences for either utilitarian or emotional reasons. The utilitarian aspect of the IS user experience relates to the accomplishment of user tasks in a more efficient or effective way. The emotional aspect of the IS user experience relates to the induction of positive or negative feelings (such as excitement or frustration) during or after using the system.

As such, we divide the dependent variables into those dealing with utilitarian aspects, i.e. those that can be expressed quantitatively as variables in performance, and those dealing with emotional aspects, i.e. perceived satisfaction and other affective reactions. This can also be referred to from the measurement model proposed in Chapter 3, Section 3.2. By building on previous literature that has tried to measure user experience, either by measuring the acceptance of technology by predicting usage (the Technology Acceptance models), user satisfaction or user engagement, the proposed model (Figure 3.11) combines and refines the different concepts, considered to be relevant to the specific study. The following describes the utilitarian, hedonic and external factors introduced in Section 3.2, and illustrated in Figure 3.11.

The **utilitarian variables** measured task performance using the following quantitative dependent variables:

- *Time on task*
  - Conceptual definition: The time needed to complete the given task.
  - Operational definition: The time, starting from when the task is presented until the task is completed. This is recorded by the testing system in a log file: a counter is turned on as soon as the page giving the task is opened, and is closed when the participant clicks on the "Submit" button. The participant is then directed to a page displaying "Click *here* to go to the next question." Clicking on the given link takes the participant to the next task and another time counter is started and recorded in the log file for the corresponding task.
  - Scale of measurement: The total time taken to complete a task will be measured in the ratio scale of seconds, with accuracy up to 1/10th of a second.
- *Accuracy*
  - Conceptual definition: Accuracy is the condition or quality of being true, correct, or exact, as defined by Dictionary.com.

- Operational definition: Each task will have a predefined correct answer or set of answers. An error will entail any other answer other than the set of correct answers for each task. For the hierarchy navigation, relations assessment, depth assessment and overview tasks, the answer will be either correct or incorrect. For the simple and complex retrieval tasks, the answer requires listing 2 books, hence it will be either zero, fifty or hundred percent accurate. The books or answers deemed to be accurate or inaccurate for the simple and complex retrieval tasks were confirmed with a Librarian before analysis. The average accuracy in a certain task will be measured by the percentage of participants who answered correctly by the total number of participants given the task.
- Scale of measurement: The accuracy will be measured in the ratio scale of percentages.

The **hedonic variables** defined by the affective reactions was evaluated using the following dependent variables: perceived ease of use, perceived usefulness, ease of learning, user satisfaction, user engagement, preference of auditory cues in sonified version, and general comments. These concepts have been introduced in Section 3.2 and will be elaborated on in this section. Each of these variables were measured using a combination of closed-ended and open-ended questionnaires.

User satisfaction and engagement are two intertwined concepts, but we have included engagement as another variable as it takes into account attributes such as aesthetics, feedback and sensory appeal through the use of multimedia components including graphics and sound, which is relevant to this research and described in Section 3.2. O' Brien and Toms have developed a survey to measure user engagement, which evaluates each of these attributes for technology use in 31 items of the user engagement scale (O'Brien & Toms, 2010a). They later honed down on 19 items for interactive information retrieval systems (O'Brien & Toms, 2010b) relating to the attributes of aesthetic appeal, focused attention, perceived usability, novelty and endurability. We have used a selection of 11 of these scale items in our questionnaires to measure user engagement. Those items that were considered to be not as relevant or applicable in the context of this system were left out. Table 5.5 gives details of the 19 original attributes from their scale, highlighting which ones were selected and why, as well as the changes that were made to the selected items (✓ denotes items which were included in the final scale, ✗ denotes those not included).

Table 5.5: The scale items for User Engagement

Attribute		**Original scale item**	**Changes to items or reasons for elimination**
Aesthetics	✓	The webcast systems was aesthetically appealing.	"Webcast" replaced with "information"
Aesthetics	✓	This webcast system appealed to my senses.	
	✓	This webcast system is attractive.	
	✗	I found the screen layout of this system to be visually pleasing.	Not applicable for the audio-visual system.
	✗	I liked the graphics and images used in this webcast system.	
Perceived usability	✓	I felt frustrated while using this webcast system.	"Webcast" replaced with "information"
	✓	I felt annoyed while using this webcast system.	

Table 5.5: The scale items for User Engagement

Attribute		**Original scale item**	**Changes to items or reasons for elimination**
Perceived usability	✓	I felt discouraged while using this webcast system.	
	✓	Using this system was taxing.	
	✗	This task was stimulating.	Study consisted of more than one task.
Novelty	✓	I would continue to use this webcast system out of curiosity.	"Webcast" replaced with "information"
	✓	The content of this webcast system incited my curiosity.	
Endurability	✓	Using this webcast system was worthwhile.	"Webcast" replaced with "information"
Endurability	✓	I would recommend that others use this webcast system.	
Focused attention	✗	I blocked out things around me when I was using this system.	Not included, since current study was conducted in laboratory conditions with minimum external distractions, and participants were specifically asked to focus on the tasks.
	✗	When I was using the system, I lost track of the world around me.	

Table 5.5: The scale items for User Engagement

Attribute		**Original scale item**	**Changes to items or reasons for elimination**
Focused attention	✗	I was absorbed in my task.	
	✗	I was so involved in my task that I lost track of time.	
	✗	I lost myself in this experience.	

- *Perceived ease of use*
  - Conceptual definition: Perceived ease of use refers to "the degree to which a person believes that using a particular system would be free of effort" (Davis, 1989).
  - Operational definition: Ease of use was measured from the five scale items presented in Table 5.6. Each of these 5 items were present in Davis's scale for perceived ease of use (Davis, 1989), and later also adopted by scales presented in (van der Heijden, 2004), (Lund, 2001), (Venkatesh & Davis, 2000). The scales were presented to participants in post-test questionnaires, where they self-reported their responses on a 5-point Likert scale, ranging from "Strongly disagree" to "Strongly agree".
  - Scale of measurement: Likert scales are at the ordinal level of measurement, and will be comparatively analyzed using non-parametric statistics such as the Wilcoxon signed ranks test.

- *Perceived usefulness*
  - Conceptual definition: Perceived usefulness is defined as "the degree to which a person believes that using a particular system would enhance his or her performance" (Davis, 1989).
  - Operational definition: Perceived usefulness was measured from the scale item presented in Table 5.6, asking how useful participants' perceived the auditory cues in the system while completing the tasks, adopted from (Davis, 1989) and (Lund, 2001). This was presented to them in a post-test questionnaire, where they self-reported their response on a 5-point Likert scale, ranging from "Strongly disagree" to "Strongly agree". .
  - Scale of measurement: Likert scales are at the ordinal level of measurement, and will be analyzed using non-parametric statistics such as the Wilcoxon signed ranks test.
- *Ease of learning*
  - Conceptual definition: How easy or difficult a user finds it to learn to use the sounds.
  - Operational definition: Ease of learning was measured from the three scale items presented in Table 5.6, about how participants' perceived the learnability of the auditory cues, adopted from (Lund, 2001). These were presented to participants in post-test questionnaires, where they self-reported their responses on a 5-point Likert scale, ranging from "Strongly disagree" to "Strongly agree".
  - Scale of measurement: Likert scales are at the ordinal level of measurement, and will be analyzed using non-parametric statistics such as the Chi-square.

- *User satisfaction*
  - Conceptual definition: User satisfaction has been defined as a complex construct comprising several affective components such as emotion, expectation, likeability and usability (Lindgaard & Dudek, 2003).
  - Operational definition: User satisfaction was measured from the seven scale items presented in Table 5.6, about how satisfied they felt using the overall system or its auditory cues (one scale item for each of the five auditory cues). These items were adopted from the (Doll & Torkzadeh, 1988) EUCS instrument and the (Lund, 2001) usability questionnaire. These were presented to participants in post-test questionnaires, where they self-reported their responses on a 5-point Likert scale, ranging from "Strongly disagree" to "Strongly agree".
  - Scale of measurement: Likert scales are at the ordinal level of measurement, and will be analyzed using non-parametric statistics such as the Wilcoxon signed ranks test.
- *User engagement*
  - Conceptual definition: Engagement is a quality of user experiences with technology that is characterized by challenge, aesthetic and sensory appeal, feedback, novelty, interactivity, perceived control and time, awareness, motivation, interest, and affect (O'Brien & Toms, 2008).
  - Operational definition: Out of the different attributes, a selection of those concerning aesthetics (AE), perceived usability (PU), novelty (NO), and endurability (EN) have been taken to measure user engagement in this study (O'Brien & Toms, 2010b). Endurability refers to the assessment of users' perception of success with a task, and their willingness to use the application in the future or recommend

it to others (O'Brien & Toms, 2010a). The selection of the 11 items out of the 19 is justified in Table 5.5, and the final scale items are presented in Table 5.6, corresponding to each attribute. These items were presented to participants in post-test questionnaires, where they self-reported their responses on a 5-point Likert scale, ranging from "Strongly disagree" to "Strongly agree".

  - Scale of measurement: Likert scales are at the ordinal level of measurement, and will be analyzed using non-parametric statistics such as the Wilcoxon signed ranks test.

- *Preference of system*
  - Conceptual definition: Preference describes whether participants' preferred using one system over the other.
  - Operational definition: Preference was evaluated in two ways. One was a close-ended question, asking if the participant preferred using the visual-only or the audio-visual system. The other was from an open-ended question to explain why they preferred one or the other.
  - Scale of measurement: Preference is based on the nominal level of measurement. Inputs will be categorized into two groups: those that preferred the audio-visual, and those that did not. The mode, which is the most often selected value, will indicate the central tendency. Verbal data taken from the open-ended question will also be used to categorize the overall preference.
- *General comments*
  - Conceptual definition: Any other affective reactions, i.e. those referring to the emotion or feelings produced by the interface, task and cues provided in the experiment, e.g. those of liking, disliking, or the

experience of pleasure or displeasure, were obtained from the general comments.

- Operational definition: General comments or the answers to all the open-ended questions in the questionnaires gave access to a rich source of qualitative data to characterize the affective reactions not captured by the Likert scales.
- Scale of measurement: The comments and observations were coded and categorized into positive and negative feedback for each system.

Table 5.6 lists and summarizes all the hedonic dependant variables and their respective measurement scale items in the post-test questionnaires, together with their references. In the User Satisfaction scale items, each auditory cue corresponds to a separate item in the questionnaire, but are grouped together in the table. Each scale was measured using a 5-point Likert scale ranging from “Strongly disagree” to “Strongly agree’.

Table 5.6: Scale items for hedonic variables on a 5-point Likert scale

Dependent variable	**Scale item**	**References**
User Engagement (AE)	This information system is aesthetically appealing.	(O'Brien & Toms, 2008), (O'Brien & Toms, 2010a), (O'Brien & Toms, 2010b)
	This information system appealed to my senses.	
	This information system is attractive.	
User Engagement (PU)	I felt frustrated while using this information system.	
	I felt annoyed while using this information system.	
	I felt discouraged while using this information system.	
	Using this system was taxing.	
User Engagement (NO)	I would continue to use this information system out of curiosity.	
	The content of this information system incited my curiosity.	
User Engagement (EN)	Using this information system was worthwhile.	
	I would recommend that others use this information system.	

Table 5.6: Scale items for hedonic variables on a 5-point Likert scale

Dependent variable	Scale item	References
Perceived ease of use	The interaction with the system is clear and understandable.	(van der Heijden, 2004), (Lund, 2001), (Venkatesh & Davis, 2000) (Davis, 1989)
	Interaction with the system does not require a lot of mental effort.	
	I found the information system easy to use.	
	I found it easy to get the system to do what I want it to do.	
	I found the different sounds cues and their functions easy to use.	
Ease of learning	I learned the different sounds cues and their functions quickly.	(Lund, 2001)
	I easily remembered the different sounds cues and their functions.	
	I quickly became skillful at using the different sounds cues.	
Perceived usefulness	Overall, I found the different sound cues useful while carrying out the tasks.	(Lund, 2001), (Davis, 1989)
User satisfaction	This system is fun to use.	(Lund, 2001)

Table 5.6: Scale items for hedonic variables on a 5-point Likert scale

Dependent variable	**Scale item**	**References**
User satisfaction	I am satisfied with the “click” / “zoom” / “depth” / “leaf” / “information density” sound in the system.	(Doll & Torkzadeh, 1988)
	Overall, I am satisfied with this system.	

The **prior experience variables** are predispositions, a priori factors or psychological factors that may affect the participant's perception of the tested systems. These include questions asking about the participants' previous experiences or sensitivities, as shown in Table 5.7. Each of these variables will be examined to see if there are any correlations or relationships between the prior experience variable and any of the utilitarian variables or preference, e.g. if there is any correlation between high noise sensitivity and preference for the visual-only system.

Table 5.7: Prior experience variables in post-test questionnaires

Variable	**Question**	**Input format**
Noise sensitivity	In general, I am sensitive to noise.	Strongly disagree / Disagree / Neutral / Agree / Strongly agree

Table 5.7: Prior experience variables in post-test questionnaires

Variable	Question	Input format
Use of computer sounds	Do you use computer interface sounds in your daily life or do you prefer to turn such sounds off? Please specify why.	Free format.
Familiarity with text-based system	Have you ever used the McGill Online Library Catalogue?	Yes / No
	If yes, how often do you use it?	Once or few times a year / Once or few times a term / At least once a month / At least once a week.
	Given a choice between the three, which would you prefer using?	The online library catalogue / The visual-only system / The audio-visual system
Musical experience	Have you ever taken music lessons e.g. instrumental, vocal or music theory, in addition to the regular music curriculum in school?	Yes / No
	If yes, please specify what type and the number of years.	Input boxes.

Table 5.7: Prior experience variables in post-test questionnaires

Variable	Question	Input format
Musical experience	Even if not trained, do you play any musical instruments? If yes, which ones? Please mention the number of years you have been playing.	Free format.
	Have you had any ear training, sound engineering training, music technology or any other relevant training in sound? Please specify the type of training and the number of years trained	Free format.
Video gaming experience	Do you ever play video games (those that include auditory feedback)?	Yes / No
	If yes, which ones do you play?	Free format.
	If yes, how often do you play?	Once or few times a year / Once or few times a term / At least once a month / At least once a week.
Experience in Information Visualizations	Have you ever used any information visualizations? (e.g. visual search engines, data or social network visualization tools, etc.?)	Free format.
	If yes, please specify which ones and comment about your experience with them.	Yes / No

The three post-test questionnaires, presented to the participants after the visual-only session, the audio-visual session, and after the completion of all sessions, are attached in Appendix C.

A summary of all dependent variables in the comparative evaluation are given in Table 5.8, with their definitions and data collection instruments.

Table 5.8: Overview of dependent variables

Dependent variable	**Conceptual definition**	**Operational definition**	**Data collection instrument**
Utilitarian	Time on task	Time from task presentation to submission of task answer	Computer log
Utilitarian	Accuracy	Percentage of errors in tasks	Recorded task answers
Hedonic	Preference	Preference of audio-visual or visual-only system	Post-test questionnaire (nominal categories)
Hedonic	User Engagement	Defined by the attributes of Aesthetics, Perceived Usability, Novelty and Endurability	Post-test questionnaire (5-point scales)
Hedonic	Perceived ease of use	Self-report on the ease of use of the system and sounds	Post-test questionnaire (5-point scales)
Hedonic	Perceived usefulness	Self-report on the usefulness of the sounds	Post-test questionnaire (5-point scales)

Table 5.8: Overview of dependent variables

Dependent variable	**Conceptual definition**	**Operational definition**	**Data collection instrument**
Hedonic	Perceived ease of learning	Self-report on the ease of learning to use the sounds	Post-test questionnaire (5-point scales)
Hedonic	User satisfaction	Self-report on the satisfaction on using the system and sounds	Post-test questionnaire (5-point scales)
Prior experience	Noise sensitivity	Self-report on user's sensitivity to noise level	Post-test questionnaire (5-point scale)
Prior experience	Familiarity with text-based system	Self-report on exposure to original text-based library system	Post-test questionnaire (nominal categories)
Prior experience	Music, ear or sound training	Self-report on training level in music, ear or sound	Post-test questionnaire (nominal categories and free-format comments)
Prior experience	Experience with visualizations	Self-report on exposure to other visualization systems	Post-test questionnaire (nominal categories and free-format comments)
Prior experience	Video gaming experience	Self-report on experience in playing video games	Post-test questionnaire (nominal categories and free-format comments)

CHAPTER 6

ANALYSIS AND RESULTS OF THE COMPARATIVE EVALUATION

This chapter discusses the analysis of the data acquired during the experiment from 24 participants, as described in Chapter 5. Each type of data is described in a different section, i.e. the data for the utilitarian variables of time and accuracy, the analysis of the hedonic variables using the rating scales in the post-test questionnaires, the prior experience variables, and the qualitative analysis of the free-format comments. The next section frames the hypotheses before going into the analyses.

6.1. Hypotheses

Following the conceptual framework and experimental procedure described in Chapter 3 and Chapter 5, the following questions are statistically and qualitatively analysed:

- Are there differences across the audio-visual and visual-only systems in terms of the utilitarian variables (Time on task or Accuracy)?

- Are there differences across the audio-visual and visual-only systems in terms of the hedonic variables (Preference, Perceived Ease of Use, User Satisfaction, and each of the four attributes of User Engagement (Aesthetics, Perceived Usability, Novelty and Endurability))?
- What do participants like or dislike about the sounds (User Satisfaction, Perceived Usefulness, Perceived Ease of Learning and Perceived Ease of Use)?
- Are participants satisfied with each of the individual sounds?
- Is there any correlation between preference and each of the prior experience variables (familiarity with text-based system, musical experience, ear or sound training, video gaming experience, experience with information visualizations and noise sensitivity)?
- What can be concluded from the free-format comments?

Based on the above questions, the following initial hypotheses are listed for the quantitative data (Table 6.1):

Table 6.1. Experimental hypotheses

Number	**Hypothesis**
H1	Time will differ significantly between the two systems.
H2	Accuracy will differ significantly between the two systems.
H3	Preference will differ between the two systems.
H4	Perceived ease of use will differ between the two systems.
H5	User satisfaction will differ between the two systems.
H6	User engagement (aesthetics, perceived usability, novelty and endurability) will differ between the two systems.
H7	Preference and the prior experience variables will be correlated.

6.2. Methods

6.2.1. Participants. 24 student participants from Science and Engineering took part in the comparative evaluation, consisting of 12 males and 12 females. Their ages ranged from 18 to 35 years old (average age of 23.75, median of 23, and standard deviation of 4.18). Their academic majors ranged from Engineering subjects such as Electrical, Civil, Chemical and Materials Engineering, to Science subjects such as Music Technology, Architecture, Computer Science, Biology and Chemistry. The participants were each compensated with $15 CAD for their time and participation.

6.2.2. Procedure. The experiments took place at the Multimodal Interaction Lab, McGill University. The experiments were conducted on a Windows XP operating system, running on a PC with a hyper-threaded quad-core CPU providing a total of 8 computing cores. The CPU speed was 2.67 GHz with 3 GB of RAM, an external Motu 828 MkII sound card and AKG K240 Studio headphones. Dual screens were used for the experiment, with the visualization system displayed on a large wide-screen 32” display on the left-hand side, and the web testing-engine displaying the task forms and questionnaires on a smaller 24” display on the right-hand side.

An experiment with a single participant took no more than 90 minutes, consisting of the training, two experimental sessions and the questionnaires. The steps for the experimental sessions are outlined below. More details can be referred to in the Appendix C, where the forms used for the experiments including the informed consent form, the demonstration guidelines used in the training session, the exact protocol used by the experimenter, as well as screenshots of the post-test questionnaires and sample task forms are all attached.

- Introduction to experimenter and study
- Reading and signing of consent form by participant

- Demonstration of audio-visual or visual-only, depending on system order
- Training session with training task set.
- Experimental session 1 with audio-visual or visual-only, depending on system order, and task set 1 or 2, depending on task set order.
- Given questionnaire where sounds are rated if previous session was with audio-visual. Otherwise, proceeded to next session.
- Demonstration and training session of audio-visual or visual-only, depending on system order.
- Experimental session 2 with audio-visual or visual-only, depending on system order, and task set 1 or 2, depending on task set order.
- Presented with questionnaires: sound ratings (if not presented before), preference form, combined ratings for both systems and post-test demographics form.
- Completion of experiment. Presented with compensation and receipt.

6.2.3. Tasks. An overview of the tasks tested are provided in Table 6.2. The exact tasks in each of the two task sets used in the experiment can be found in Table 5.4.

6.3. Analysis of Utilitarian Variables

6.3.1. Descriptive data statistics. The time data are represented in seconds, while the accuracy data are represented as percentages. Table 6.3 gives the overall statistics for both time and accuracy. The data set contained only one incomplete data entry, in which one participant did not answer Task 6 of Task Set 1. The time recorded for this task was declared as missing, while the accuracy was zero. This can also be seen in Table 6.3.

Table 6.2. Tasks tested in Comparative Evaluation

Task type	Task	Description
Hierarchy navigation	Subject traversal	Find a subject which directly belongs to another subject.
	Subject relations assessment	Find the nearest common subject that two subjects belong to.
	Overview	Which of these two subjects contains more documents?
	Depth assessment	Which of these two subjects is at a deeper level in the hierarchy?
Information retrieval	Simple retrieval	Find 2 promising books on a subject whose keywords are specified in the question.
	Complex retrieval	Find 2 promising books for a topic whose keywords are not specified in the question?

Appendix E contains two more tables that display the means, standard deviations, lower and upper bounds, according to a 95% Confidence Interval, of the two dependent variables for each of the six tasks and for each of the two systems respectively.

Table 6.3. Descriptive statistics

Statistic	Time	Accuracy
Valid	287	288
Missing	1	0
Mean	69.064	87.50
Median	45.000	100.00
Mode	28.0	100
Std. Deviation	71.4300	32.331

6.3.2. Time on task. The results were submitted to a 2 (System) $\times$ 2 (Task Type) mixed ANOVA for Time, where the 2 levels for System are "Audio-visual" and "Visual-only" and the two levels for Task Type are "Hierarchy Navigation" and "Information Retrieval" tasks.

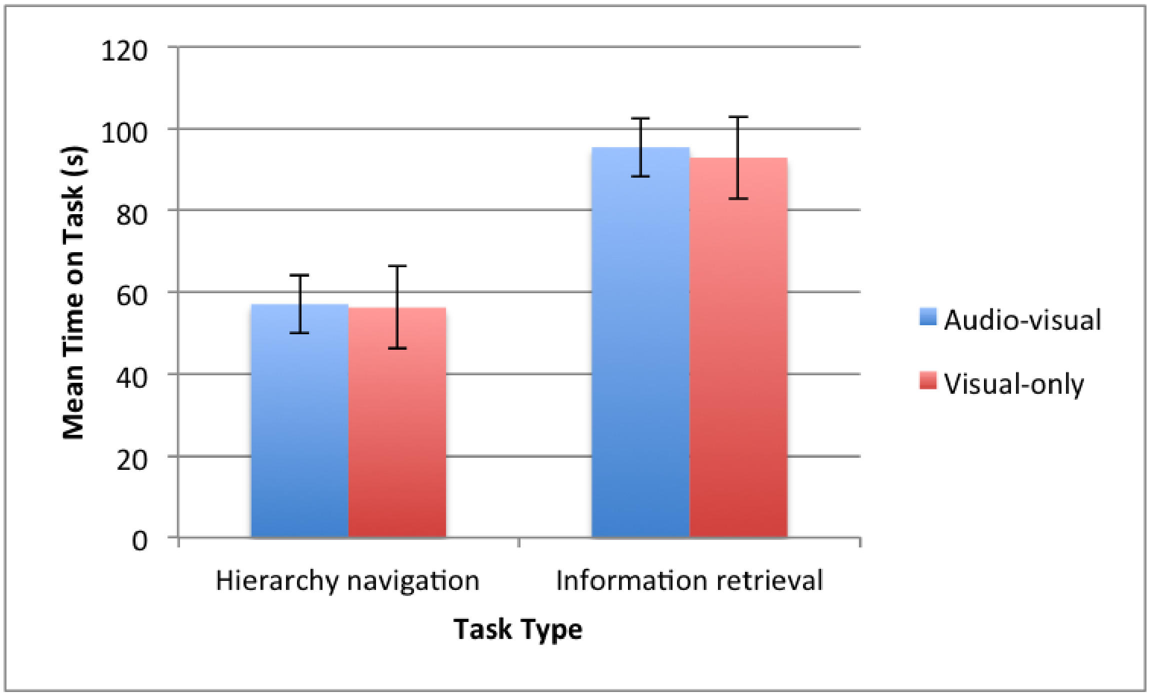


FIGURE 6.1. Bar chart of average completion time per task type for each system

The analysis revealed a main effect of Task Type (F(1, 283) = 18.532, $p < .001$), no main effect of System (F(1, 283) = 0.038, p = 0.845) and no interaction effect for System and Task Type (F(1, 283) = 0.01, p = 0.92).

Figure 6.2 shows a bar chart of the mean time taken per task type (collapsing over all participants within each of the two task types) for each of the two systems, audio-visual and visual-only. The errors bars are also shown in the figure.

The significant effect for Task Type can be explained by the different complexity levels for the two task types. Hierarchy navigation tasks entailed navigating the information system to retrieve data about the relations between different subjects, or compare aspects such as depth level or information density between subjects. Information retrieval tasks required searching for information using known or unknown keyword searches, which were of an entirely different complexity level from the hierarchy navigation tasks, and hence took very different times on the two task types.

A preliminary analysis including all the factors was also conducted to evaluate if there were any other main or interaction effects. A 2 (System) × 6 (Task) × 2 (Task Set) × 2 (Task Set Order) × 2 (System Order) mixed ANOVA was performed, where System, Task and Task Set were within-subjects variables and the Task Set Order and System Order were between-subjects variables. The ANOVA again shows no significant main effect for System on time (F(1, 246) = 0.068, p = 0.795). There is however a significant effect found for Task Set (F(1, 246) = 4.84, p = 0.029) and interaction effects for Task with Task Set (F(5, 246) = 5.049, $p < 0.001$) and System with System Order (F(1, 246) = 13.275, $p < 0.001$).

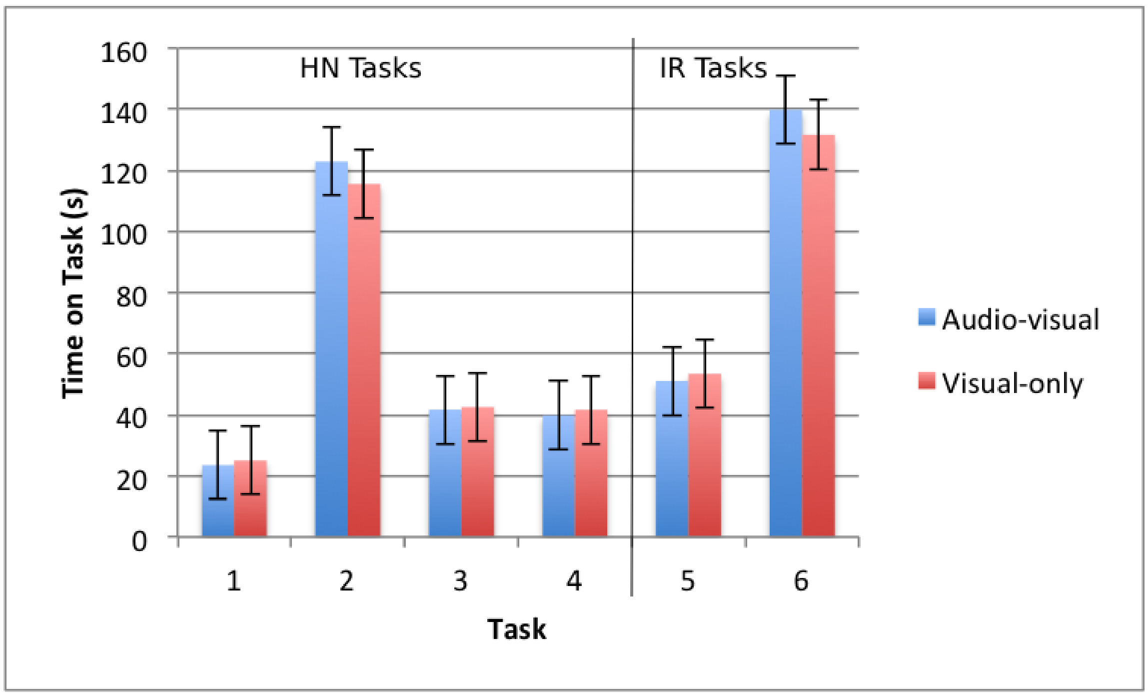


FIGURE 6.2. Bar chart with error bars for average time on task for 4 Hierarchy Navigation (HN) tasks and 2 Information Retrieval (IR) tasks over each system

Post-hoc tests (LSD, Tukey's and Bonferroni test) show Task 2 and 6 to be significantly different across time from Tasks 1, 3, 4, 5. This can be seen in Figure 6.2 since there are peaks at Tasks 2 and 6, which required higher completion times than the other four tasks. Hence a significant main effect for Task can be expected.

However, a significant effect for Task Set means that the two task sets are not equivalent, although equivalence was maintained by selecting subjects that were at the same depth level for each of the hierarchy navigation and relations assessment tasks, and the same difference in depth levels or information density in the depth assessment and overview tasks, for both task sets. Only for the search tasks, Task 5 (simple retrieval) and 6 (complex retrieval), equivalence could not be ensured since search times may vary significantly for similar problems.

Figure 6.3 shows the average time taken per task for each of the two task sets, Task Set 1 and 2, over all participants and systems. As can be seen from the figure, the task completion times between the two task sets do not vary much for the first 5 tasks, but vary more significantly for Task 6, the complex retrieval task.

In Task Set 1, Task 6 asked for books relevant to fixing a broken computer, while that in Task Set 2 asked for books relevant to fixing a leaky faucet (Table 5.4). The latter seemed to cause the time taken to complete the task to significantly increase. This may be due to the fact that all participants had a background in Science or Engineering, hence dealing with computer problems may be more in their domain, rather than solving problems in plumbing. Hence, it was decided to take out Task 6 during the analysis, leaving 5 tasks in each task set. The subsequent results deal with the analysis of the first 5 tasks. Task 6 will be analyzed separately after this.

A repeat of the mixed ANOVA for the 5 tasks shows that the significant main effect for Task Set and the interaction effect of Task Set with Task Set Order is taken out. The significant main effect for Task remains ($F(1, 246) = 84.805, p < 0.0001$), as well as the interaction for System with System Order ($F(1, 246) = 14.799, p = 0.000$).

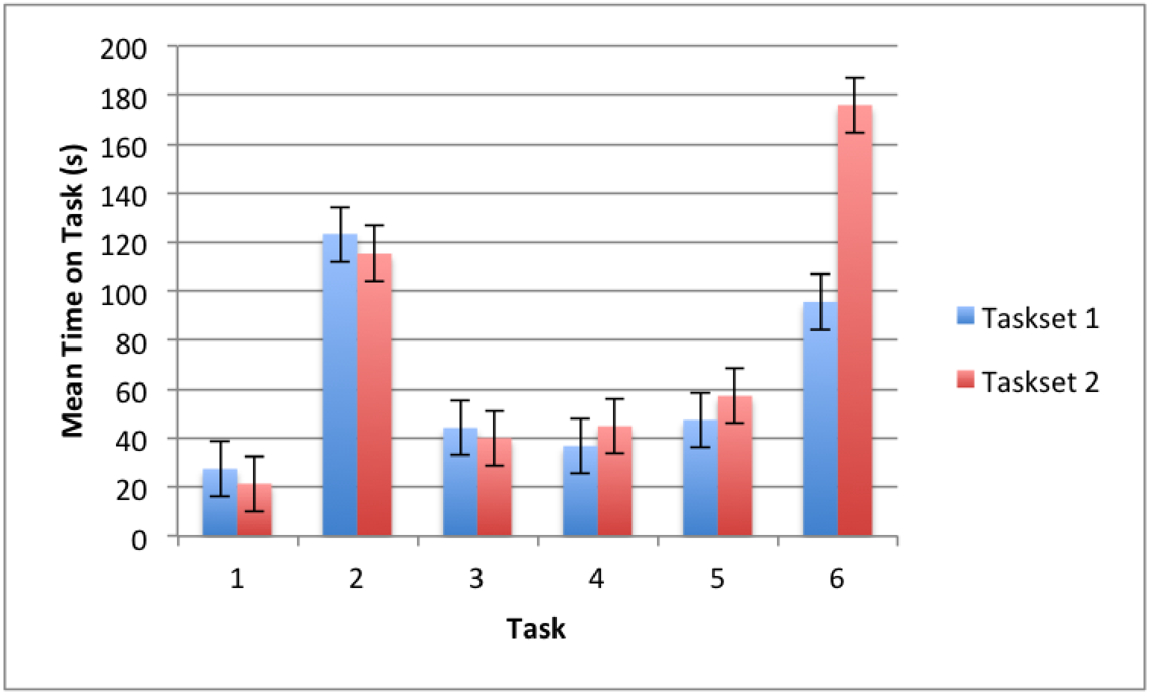


FIGURE 6.3. Bar chart with error bars of average time on task for each taskset

Post-hoc tests (LSD, Tukey's and Bonferroni tests) show Task 2 to be significantly different across time from Tasks 1, 3, 4, 5, which required a higher completion time than the other four tasks. This is due to the relative difficulty of the relations assessment task, where the closest common parent of two subjects had to be found. Hence, again, a significant main effect for Task can be expected.

Figure 6.4 shows the line graph of each system by system order, where 1 denotes the system that was first in the order, and 2 for second. This shows that the average time taken to complete tasks was always less for the second system, no matter which system it was. This can be attributed to a learning effect, so that participants became more proficient at using the system in the second session, especially since the two systems were similar except for the presence or absence of auditory feedback. However, the drop in average completion times for the visual-only system is steeper

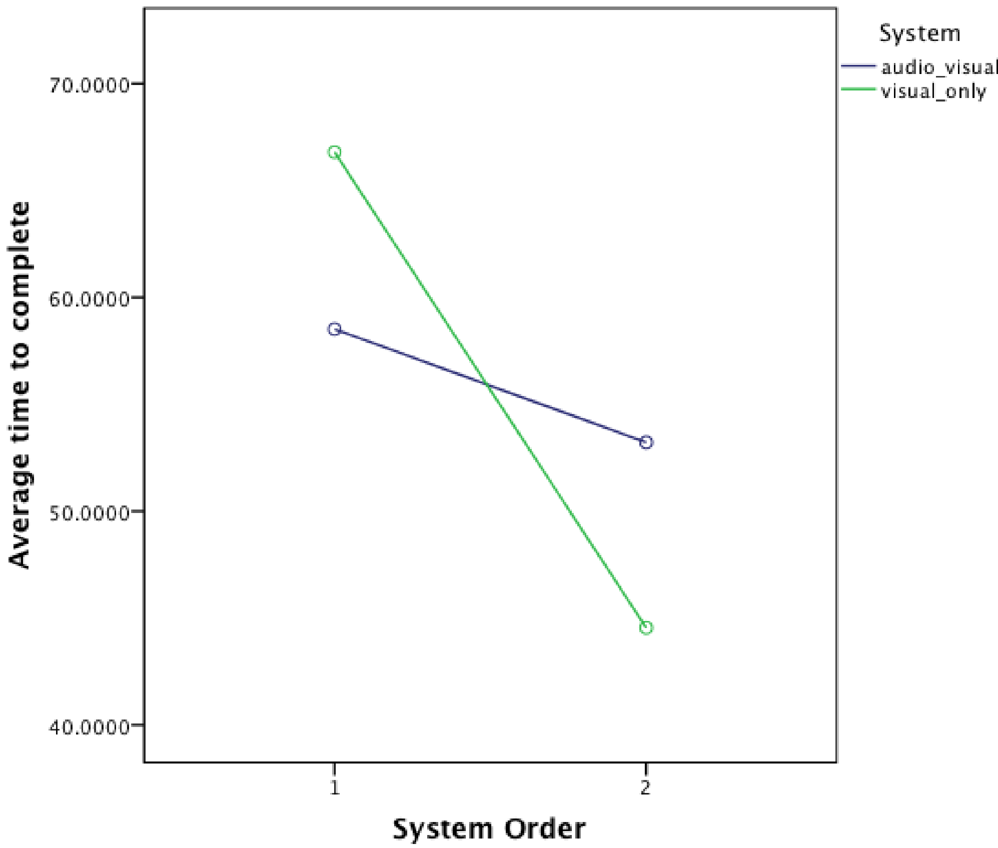


FIGURE 6.4. Line graph of average time on task for each system by system order

than that for the audio-visual, leading to the interaction effect. A reason for this may be that since the visual mode is the predominant mode users are accustomed to, the learning effect for the visual mode is faster than the auditory mode. Further testing would be required to confirm this hypothesis.

Task 6 was analyzed separately, by submitting the task completion times to a 2 (System) x 2 (Task Set) X 2 (Task Set Order) x 2 (System Order) mixed ANOVA, where System and Task Set were within-subjects variables and the Task Set Order and System Order were between-subjects variables. Again, no significant effect for System was found. The only significant effect is for Task Set (F(1,246)=5.864, p=0.025) as was expected from the results of the mixed ANOVA over all 6 tasks.

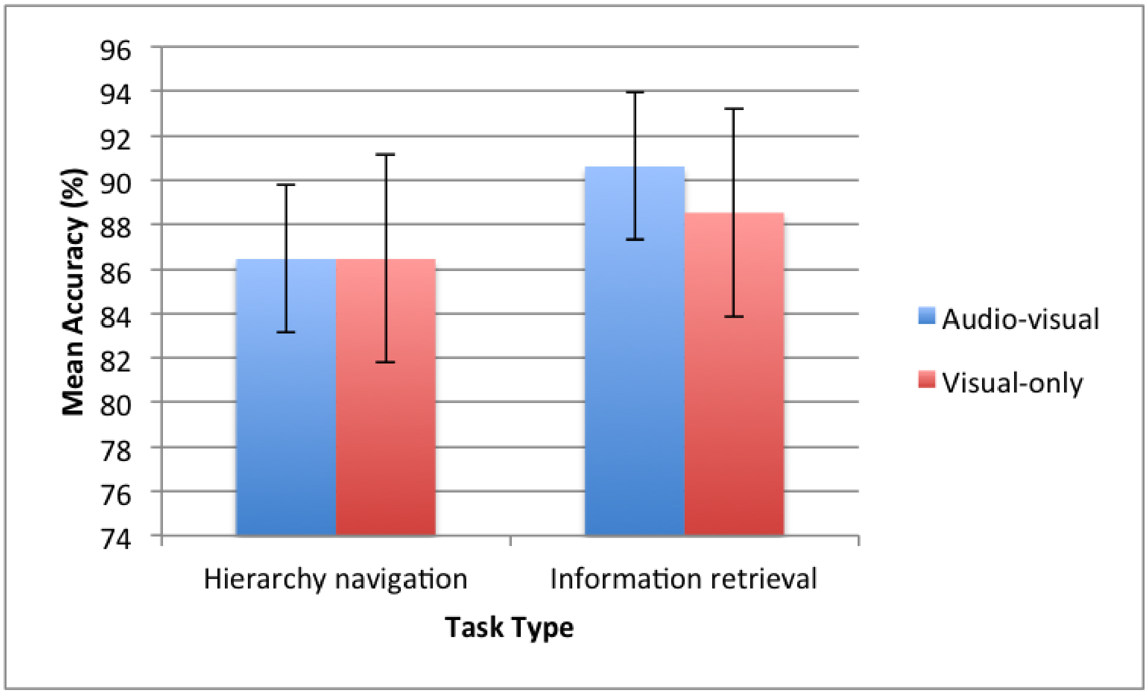


FIGURE 6.5. Bar chart with error bars of average accuracy per task type for each system

6.3.3. Task accuracy. The results were submitted to a 2 (System) × 2 (Task Type) mixed ANOVA, where the 2 levels for System are Audio-visual and Visual-only and the two levels for Task Type are Hierarchy Navigation and Information Retrieval tasks. No significant main effect was found for System on time (F(1, 284) = 0.066, p = 0.798). There is also no significant effect found for Task Type (F(1, 284) = 0.593, p = 0.442) or interaction effect for System and Task Type (F(1, 284) = 0.066, p = 0.798).

Figure 6.2 shows a bar chart of the mean accuracy per task type (collapsing over all participants within each of the two task types) for each of the two systems, audio-visual and visual-only. The errors bars are also shown in the figure.

A preliminary analysis including all the factors was also conducted to evaluate if there were any other main or interaction effects. A 2 (System) $\times$ 6 (Task) $\times$ 2 (Task Set) $\times$ 2 (Task Set Order) $\times$ 2 (System Order) mixed ANOVA was performed, where System, Task and Task Set were within-subjects variables and the Task Set Order and System Order were between-subjects variables. No significant effects were found, except for Task ($F(5, 267) = 22.866, p < 0.0001$), which can be expected due to the different complexity levels of the 6 tasks.

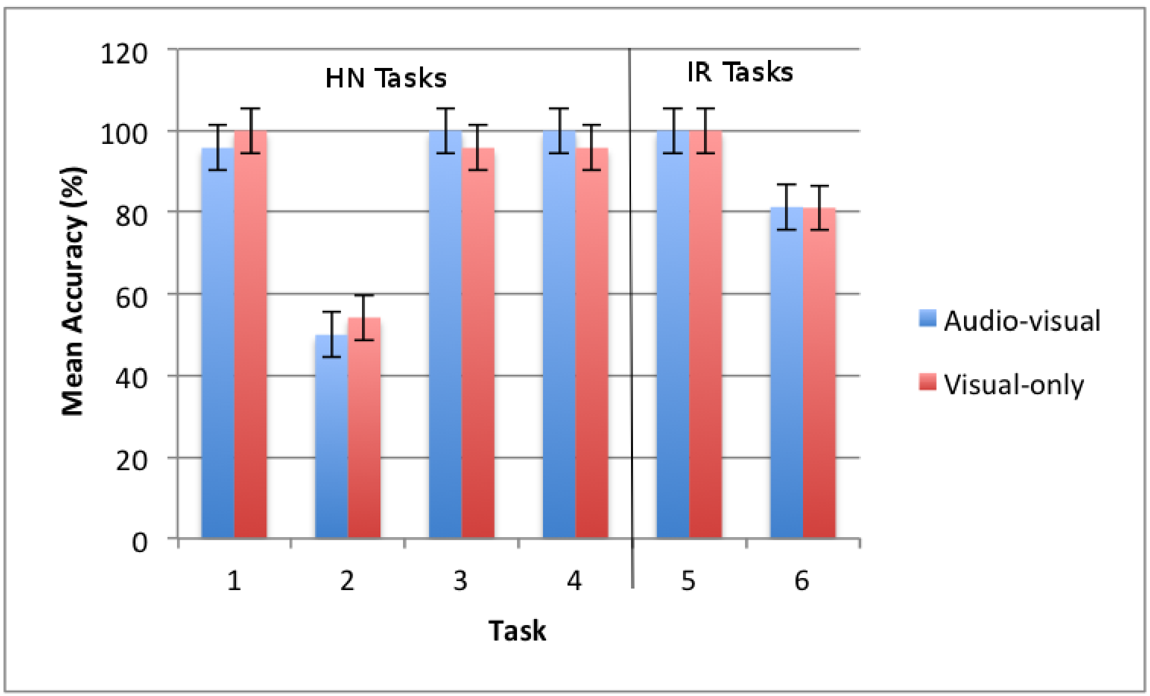


FIGURE 6.6. Bar chart with error bars for average accuracy on task for 4 Hierarchy Navigation (HN) tasks and 2 Information Retrieval (IR) tasks over each system

Figure 6.6 shows the average accuracy with error bars per task for each of the two systems, audio-visual and visual-only, over all participants and task sets.

Similar to the results in the analysis of completion times, post-hoc tests (LSD, Tukey's and Bonferroni test) show Task 2 and 6 to be significantly different across time from Tasks 1, 3, 4, 5. This can be seen in the figure since there is are dips at Tasks 2 and 6, which had higher numbers of errors than the other four tasks due to the relative difficulty in the two tasks. Hence a significant main effect for Task in accuracy can be expected.

6.4. Analysis of Preference and Likert scales

6.4.1. Preference. Participants selected their preference using a close-ended question, which asked if the participant preferred using the visual-only or the audio-visual system. They also explained their choice in a free-format comment. It was found that 19 out of the 24 participants preferred the audio-visual system (79.2%) and 5 preferred the visual-only system (20.8%). Hence, the mode or central tendency lies in the choice of audio-visual. A one-sample Binomial test on the variable shows a significant value ($p=0.007$), so that the Null hypothesis that the two preference categories occur with equal probabilities can be safely rejected.

The participants also provided elaborate explanations for their preferences, which will be discussed in detail in Section 6.6. Worth mentioning here is one of the participants' comments who mentioned that she would switch her preference from visual-only to the audio-visual system, if she had the power to select which auditory feedback sounds she could keep on (given the choice, she would turn off the zoom sound). In that case, the percentages would go up to 83.3% for the audio-visual system and 16.7% for the visual-only.

6.4.2. Combined rating scales. One of the questionnaires presented to the participants asked them to rate their feelings about each of the audio-visual and visual-only systems using a 5-point Likert scale ranging from “Strongly disagree” to “Strongly agree’. Table 6.4 lists the hedonic dependant variables and their respective measurement scale items as an overview of the detailed table in Chapter 5 (Table 5.6). For the purpose of analyses, the table has been divided into two parts, the first dealing with combined ratings for the audio-visual and visual-only systems, while the second, Table 6.5, deals with the ratings on the specific sounds and the overall usability of the auditory feedback (Section 6.4.3).

The scale items for aesthetics, perceived usability, novelty, endurability (the attributes of user engagement (O’Brien & Toms, 2010a)), perceived ease of use, and user satisfaction from Table 6.4 make up 17 different items in total. Since the total number of scale items is large ($n = 17$), each scale item tested by itself in pairwise comparison tests do not yield any significant differences in distributions, as the adjusted p-value for the multiple comparisons is so small ($p = 0.003$ for the Sidák correction and $p = .0029$ for the Bonferroni correction). The scale items for each dependent variable were then grouped together to form one scale for each of the 6 variables.

This grouping is reasonable since each scale item is so similar in nature and asks the participant to rate their reactions to similar concepts. It was also tested to see if the scale items in each group were correlated using correlation analysis (the Spearman rho correlation coefficient for non-parametric distributions of variables with ordinal values). A high degree of correlation (positive association) was found among scale items of the same group for each of the 6 groups or dependent variables (with the coefficient rho values between 0.714 to 0.875, and $p < 0.001$).

Each of the scales for the 6 variables were then pairwise compared between systems using the Wilcoxon signed ranks test (a nonparametric test that allows

TABLE 6.4. Scale items for affective reactions on a 5-point Likert scale for each system

Dependent variable	**Scale item**
Aesthetics	This information system is aesthetically appealing.
	This information system appealed to my senses.
	This information system is attractive.
Perceived usability	I felt frustrated while using this information system.
	I felt annoyed while using this information system.
	I felt discouraged while using this information system.
	Using this system was taxing.
Novelty	I would continue to use this information system out of curiosity.
	The content of this information system incited my curiosity.
Endurability	Using this information system was worthwhile.
	I would recommend that others use this information system.
Perceived ease of use	The interaction with the system is clear and understandable.
	Interaction with the system does not require a lot of mental effort.
	I found the information system easy to use.
	I found it easy to get the system to do what I want it to do.
User satisfaction	This system is fun to use.
	Overall, I am satisfied with this system.

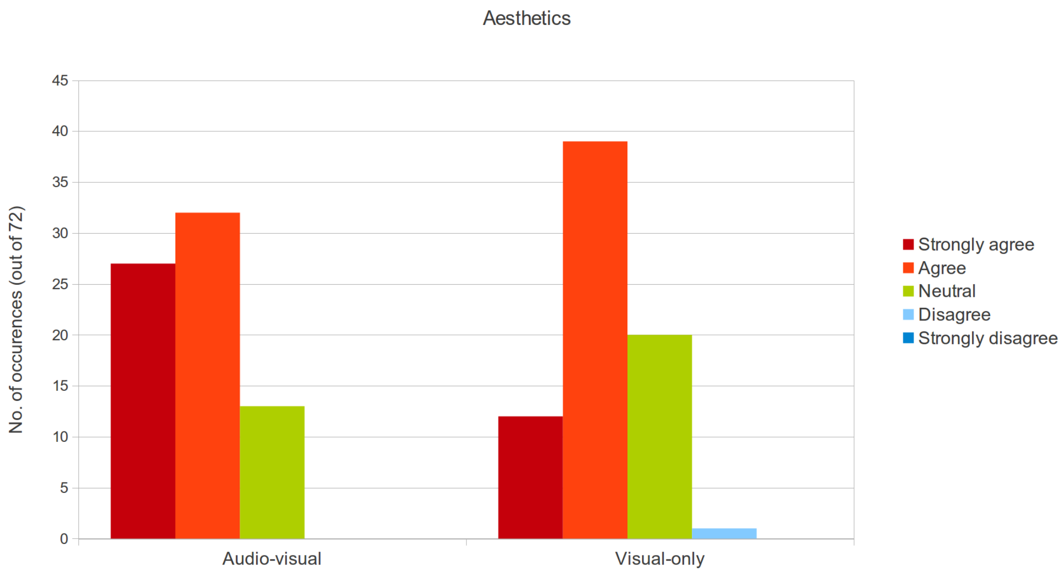
the comparison of 2 sets of scores in the ordinal level that come from the same participants, without assuming normality of distribution). The Sidák correction for the p value was used with $p = 0.0085$, since there are multiple comparisons between 6 variables.

The test showed significance for Aesthetics ($p < 0.0001$), Perceived usability ($p < 0.0001$), and Perceived ease of use ($p = 0.008$), so that the null hypothesis that the ratings for these scales were similarly distributed could be rejected. An examination of the scales themselves show a higher degree of positive rating for the audio-visual scales for these variables (a higher number of "strongly agree and agree" ratings and a lower number of "disagree" and "strongly disagree" ratings for the audio-visual system for Aesthetics, Perceived Ease of Use and User Satisfaction, while the other way around for the negative statements for Perceived Usability). No significant value was, however, found for the Novelty ($p = 0.078$), Endurability ($p = 0.527$) and User Satisfaction ($p = 0.014$) variables.

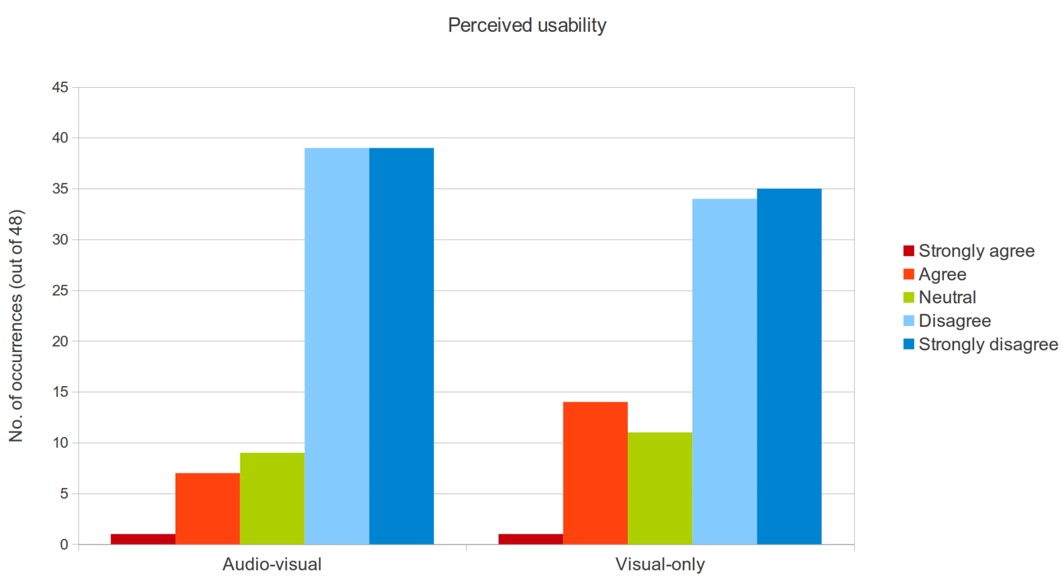
Bar charts for two of the variables that were found to be significantly different (Aesthetics and Perceived Usability) and two which were not (Novelty and Endurability), are illustrated in Figures 6.7 and 6.8 respectively. Perceived usability had scale items stated in a negative way (Table 6.4), hence the "disagree" ratings mean positive feedback.

6.4.3. Rating scales for the sounds in the audio-visual system. This section describes the analysis of the ratings on the specific sounds in the audio-visual system and the overall usability of the auditory feedback. Table 6.5 gives an overview of the scale items used in the questionnaire with a 5-point Likert scale ranging from "Strongly agree" to "Strongly disagree".

Figure 6.9 and 6.10 show bar charts corresponding to these distribution of ratings on the overall sounds and individual sounds respectively for the 24 participants. As can be seen from Figure 6.9, 75% of the participants agreed that the sound functions were easy to use (selected the "agree" or "strongly agree" scales). An even higher number of 87.5% of the 24 participants considered that they learned the sound functions quickly, and 79.2% of them thought they also easily remembered

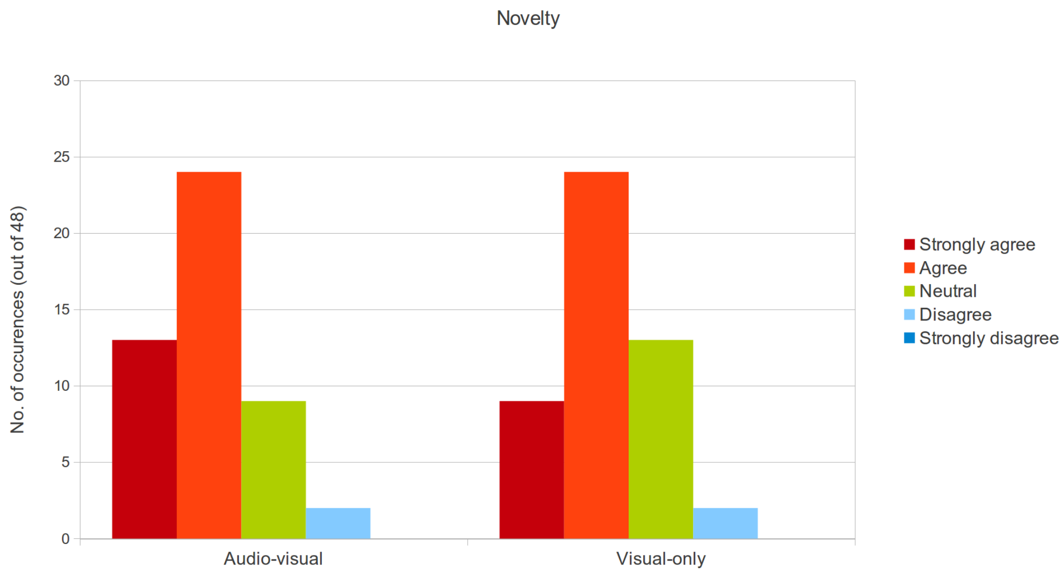


(a) Aesthetics

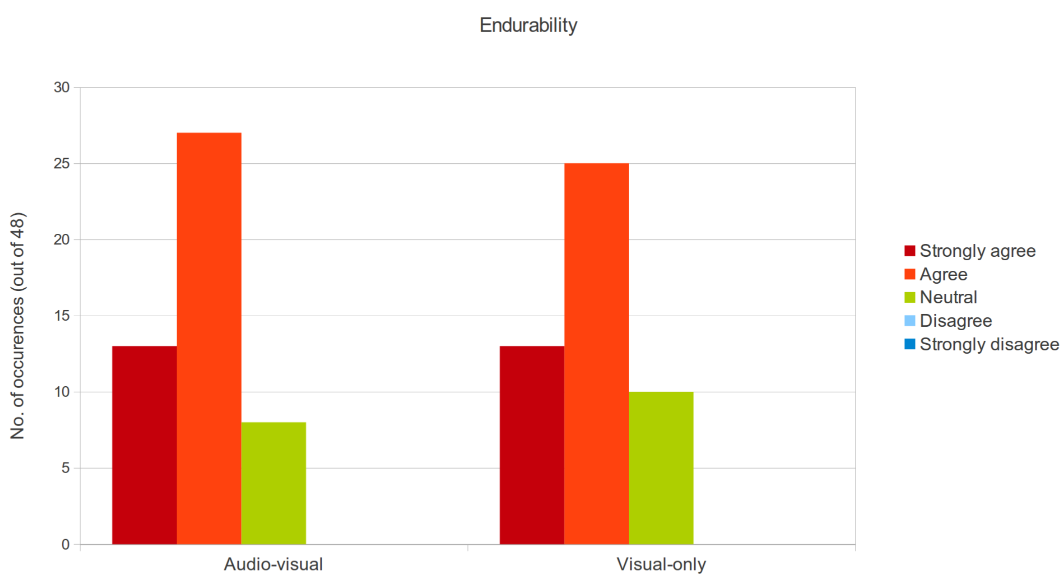


(b) Perceived Usability

FIGURE 6.7. Distribution of Likert ratings for significant variables



(a) Novelty



(b) Endurability

FIGURE 6.8. Distribution of Likert ratings for not significant variables

Table 6.5. Scale items for affective reactions on the sounds

Dependent variable	Scale item
Perceived ease of use	I found the different sounds cues and their functions easy to use.
Ease of learning	I learned the different sounds cues and their functions quickly.
	I easily remembered the different sounds cues and their functions.
	I quickly became skillful at using the different sounds cues.
Perceived usefulness	Overall, I found the different sound cues useful while carrying out the tasks.
User satisfaction	I am satisfied with the "click" / "zoom" / "depth" / "end of branch" / "information density" sound in the system.

the different sounds and their functions. However, a lower number of 54.2% of the participants thought that they became skillful at using the different sounds cues. About 67% of the participants did consider the overall sounds as useful while carrying out the tasks.

When reporting the satisfaction with each individual sounds, 75% of the participants were satisfied with the Click sound, 54.2% with the Zoom sound, 58.3% with the Depth sounds, 91.7% with the End of Branch sound and 79.2% with the Information Density sounds. These percentages show a higher liking for the Click, End of Branch and Information Density sounds, and a lower one for the Zoom and Depth sounds, which is also reflected in their comments, to be discussed in Section 6.6. These values are illustrated in the bar chart in Figure 6.10.

6.5. Analysis of Prior Experience variables

The prior experience variables are predispositions or psychological factors that may affect the participant's perception of the tested systems, including participants'

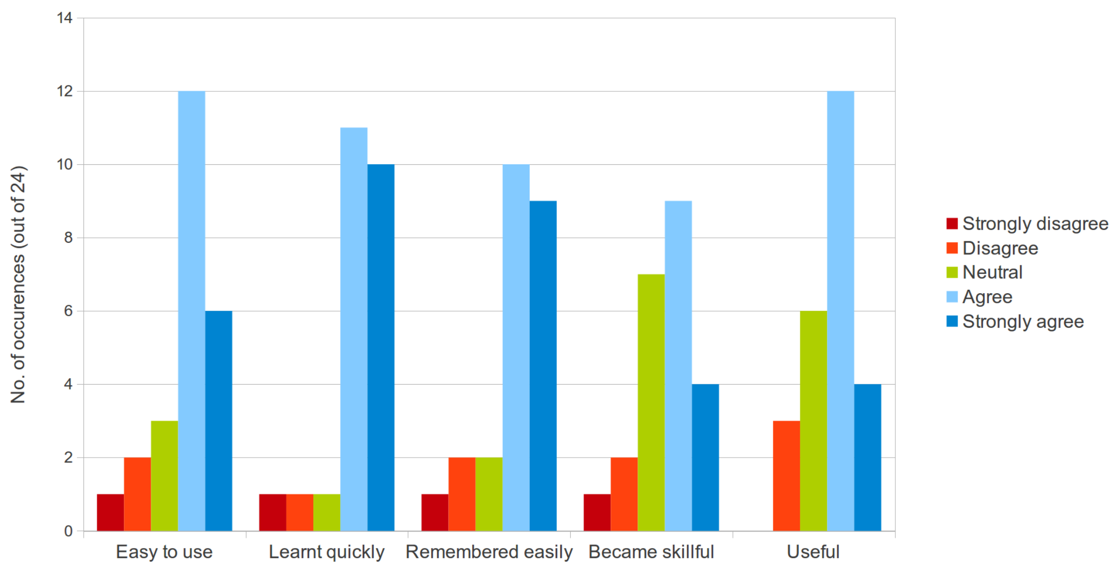


FIGURE 6.9. Bar chart for responses to overall auditory feedback by 24 participants

previous experiences or sensitivities. The variables reported by the participants include:

- Familiarity with the text-based web system
- Musical training
- Ear or sound training
- Experience with visualizations
- Uses computer interface sounds
- Video gaming experience
- Noise sensitivity

The values of each of these variables in the 5-point Likert scale were found to mainly lie in the two extreme ends of the scale. Hence, the values were converted to

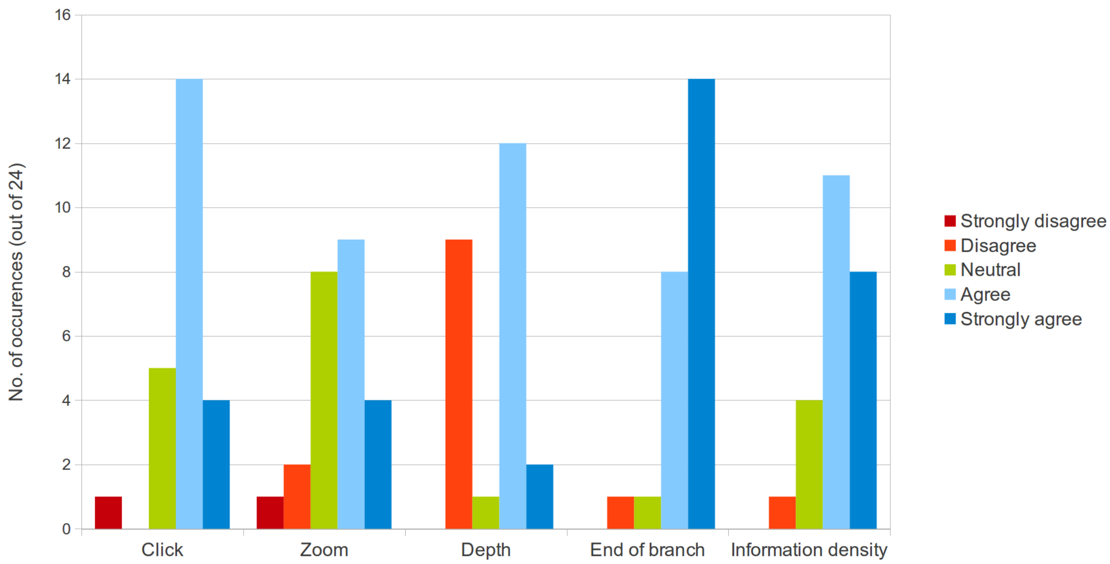


FIGURE 6.10. Bar chart for responses to individual sounds by 24 participants

nominal categories by reducing to yes and no answers. Table 6.6 gives an overview of the variables, the nominal categories and the number of occurrences in each.

These categories then tested for any correlation with preference, which also has nominal levels of audio-visual or visual-only. Familiarity with the text-based information system was not tested since 23 out of the 24 participants were familiar with the original system and only one was not. This would obviously not yield any conclusive results. The correlations with the dependent variables of time and accuracy were also not tested, since no significant effects for system were found for either variable in Section 6.3.

The distributions of each of the prior experience variables with preference are illustrated in the bar charts in Figure 6.11 and 6.12.

TABLE 6.6. Prior experience variables and occurrences of categories

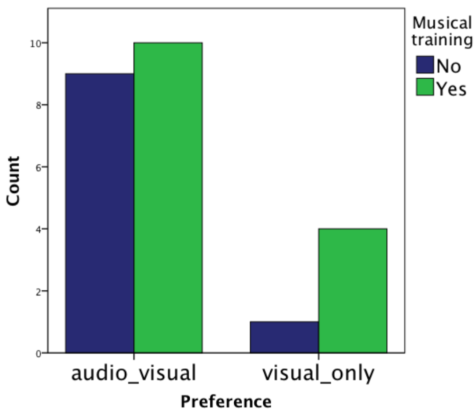
Prior experience variable	Nominal categories	No. of participants
Familiarity with text-based web system	Yes	23
	No	1
Musical training	Yes	13
	No	11
Ear or sound training	Yes	7
	No	17
Experience with visualizations	Yes	3
	No	21
Uses computer interface sounds	Yes	7
	No	17
Video gaming experience	Yes	12
	No	12
Noise sensitivity	Yes	12
	No	12

The correlations of each of the prior experience variables with preference was tested using the Phi Coefficient (correlation coefficient for nominal categories). No significant correlation was found for any of the variables with preference.

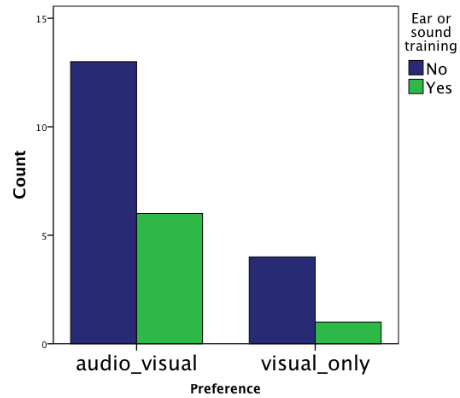
To confirm the correlation test results, the Fisher's Exact Test for the two variables of the prior experience variable with preference was also carried out (Fisher's exact test is used when the chi-square test is inappropriate due to the violation of the expected frequency of 5 or less in each cell). Results again show no significance between distributions of any of the prior experience variables with preference and hence, no conclusions can be drawn from the responses.

6.6. Qualitative analysis of comments

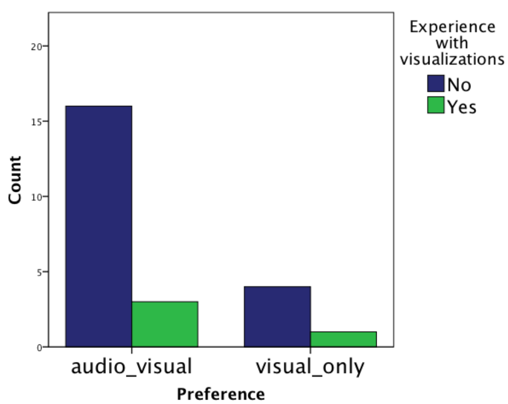
The participants had the option to provide comments or explanations for each of their selections or Likert scale ratings in the questionnaires. These provided a rich source of qualitative data that characterized their affective reactions not acquired



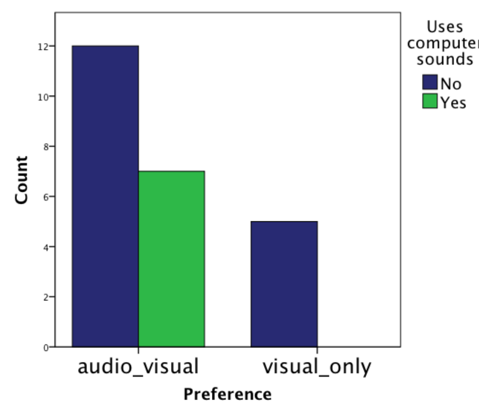
(a) Musical training



(b) Ear or sound training



(c) Experience with visualizations

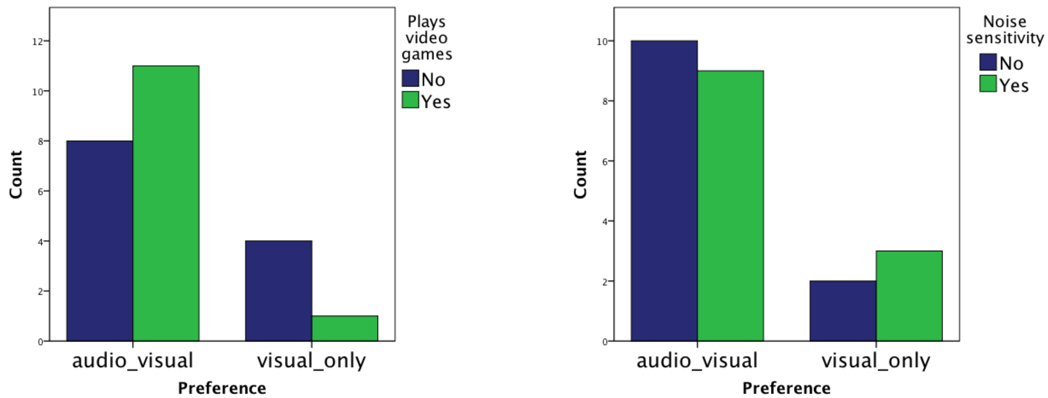


(d) Use of computer sounds

FIGURE 6.11. Distribution of prior experience variables with preference

from the discrete rating scales. All 24 participants provided some form of feedback in the comment fields; they were quite generous with their comments, freely providing explanations and defending their choices in most cases.

All the comments from every participant were gathered together, and any verbal comments recorded by the experimenter during the procedure was also added. A content analysis on this collection of open-ended questions or comments was then conducted.



(a) Video gaming experience (b) Noise sensitivity

FIGURE 6.12. Distribution of prior experience variables with preference

Each comment was examined in a preliminary examination to identify different codes from which subsequent themes or concepts could be generated. For example, codes were any collection of words or phrases that denoted adjectives describing the systems, the sounds or the feelings of the participant, adverbs denoting descriptions of actions, or words denoting actions themselves (e.g. facilitated, reinforced, etc.)

These codes were then divided into three sections: those that were concerned with the audio-visual system only, those that were concerned with the visual system only, and those that were concerned with the specific sounds in the system and not the overall general system (coding scheme presented in Appendix E).

In each of these sections, the codes were then examined to allow the grouping of similar codes to form concepts or themes. Each of these concepts were coded in a table to find the frequency of occurrences of the concepts in the entire set of data. The concepts generated from the codes in the audio-visual section are presented with examples in Table 6.7. The count gives the number of occurrences of each concept and the ratio field gives the percentage or prevalence of the concept in the total number of occurrences in the audio-visual section, which was a total of 123.

Table 6.7: Concepts from audio-visual comments (count out of 123 total occurrences)

Concept	**Count**	**Examples of comments**
Enhanced experience	14	"The audio-visual system together enhanced the experience of browsing." "The sounds added to the experience."
Quicker with sounds	11	"Because the different sounds helped me in answering faster." "Some of the info seemed more quickly available via the audio-visual system."
Engagement	27	"Audio-visual is more engaging." "I preferred the system with sound - more fun!"
Reduced visual workload	13	"The sounds reinforced the perception of navigating through an environment." "It provided additional feedback and helped to orient me within the tree more quickly."
Pleasantness	5	"Sounds facilitated motion and pleasantness of the interface." "sounds are a nice addition to the system"
Usefulness	21	"Sounds are a helpful addition to the system." "Auditory cues were helpful and efficient."
Ease of use	7	"Easy to use with sounds." "I found the system very user-friendly."

Table 6.7: Concepts from audio-visual comments (count out of 123 total occurrences)

Concept	**Count**	**Examples of comments**
Immersiveness	6	"It's simply more immersive." "In general, the audio-visual system seems to be more stimulating and interactive (and fun to use) than the visual-only system."
Prefer continuous sound	3	"I believe a person could get more out of the sounds possibly if the sounds were more continuous." "I'm used to continuous sound or none at all."
Not critical to use	5	"I think the addition of sound to this system enhances this system but is not critical to its usefulness." "Some of the sounds were helpful, but there are other ways to see the information that the sounds indicate."
Trust visuals more	10	"I trust my visual system a lot more than my audio system." "I prefer and am used to visual-only with computers, video games, almost any task."
Limits usage	1	"Although sounds might be a nice addition to the system, they limit its usage (user needs special equipment to use it, like speakers)."

Similarly, the concepts generated from the codes only concerned with the visuals of the system are presented with samples of comments in Table 6.8. The number of comments or concepts generated from the codes were fewer than that for the audio-visuals, with a total number of occurrences being 19.

Table 6.8: Concepts from visual-only comments (Count out of 19 total occurrences)

Concept	**Count**	**Examples of comments**
Would like better view of previous hierarchical levels at a time	5	"I'd like to have a better view of where I am in the tree at any given level. The ability to click up 1 to N levels would be useful." " I wish it were easier to read the parents of the node I'm currently sitting on."
Need more camera control	4	"the zoom out function zooms too much." "It would be nice for the zoom out option to be more gradual, as I often would click it once and be taken all the way back to the starting point."
System latency	5	"The wait time can get old real fast." "The system latency was distracting."
Visuals confusing or obscured	5	" Sometimes the interface was too cluttered.. as in titles overlapping, with the need to circle around them to distinguish." " Slightly confusing when looking for a connecting subject and there are too many branches on the screen, preventing a clear view."

The third section of codes came from all the comments dealing with the specific sounds in the system. These came from the questionnaire where participants rated each sounds themselves and justified their Likert scale selections. These were grouped into two categories, positive feedback and negative feedback. The positive feedback consisted of observations in which the particular sound is said to be good, appropriate, functional, suitable, perfect etc. The negative category consisted of observations in which the particular sound is said to not useful, irrelevant, not appropriate in the context, etc. Both categories together generated a total of 124 occurrences. The categories together with some sample comments for each sound are given in Table 6.9.

Table 6.9: Positive and negative concepts from specific sounds

Sounds	**Positive concept**	**Count**	**Negative concept**	**Count**
Click	"I did feel they added something to the experience of the system." "It was useful, you were certain that you had in fact clicked something."	20	"Not necessary. Irrelevant for search." "It is not particularly useful"	4
Zoom	"It's a cool sound." "Gives feedback as to the action that has been performed."	11	"It is clear what is happening without the sound." "It's okay, gets a little irritating after a while."	10

Table 6.9: Positive and negative concepts from specific sounds

Sounds	**Positive concept**	**Count**	**Negative concept**	**Count**
Depth	"I like the sounds. These sounds in particular made working with the interface very pleasant." "This was useful as it provided an easier indication of the depth of a subject when the rest of the tree was hard to see."	11	"I find it somewhat annoying, perhaps simply because I don't associate research with forest sounds." "I thought this was a little bit more difficult. I like the idea of the birds/leaves, but sometimes its a bit difficult to compare the noises."	12
End of branch	"More useful and immediate than the visual." "This was perfect feedback for hitting the base."	26	"Doesn't match with the bird and leaves."	2
Information density	"This was the most useful sound by far. Could easily extract information quickly and confidently." "Very clear, easy to remember the meaning."	25	"Only helps in differentiating on macro level." "The sound is not indicative of whether the information will be useful to us."	3
Total		93 = 75%		31 = 25%

From the concepts generated from the audio-visual comments (Table 6.7), the themes and ideas highlighted by the participants can be traced back to concepts discussed in the conceptual framework and previous literature discussed in the previous chapters of this thesis.

Concept 1, in which the sounds enhanced or added to the experience, epitomizes the effect of auditory feedback on affective reactions. Participants may not have always felt that the sounds were essential to the use of the system, but the presence of the sounds contributed to an improved user experience and to quote one participant, he “appreciated their existence”. This contributes to the hedonic variables of user experience described in Chapter 3 (Figure 3.11), affecting the user engagement, satisfaction and ultimate preference of the system. A cross-reference of the comments with the participants showed that the comments that were marked with the codes for Concept 1 were from those who preferred the audio-visual system over the visual-only one.

Concept 2, in which the perceived speed of the system or tasks was more than with the visual-only, is an interesting one, given the fact that it contradicts the findings on actual time taken to complete the tasks. No significant difference was found between the times taken for both systems, however, the comments seem to indicate that it was perceived as faster for the audio-visual system. This would contribute to the perceived usefulness part of the user experience model (Figure 3.11). A cross-reference of the comments with the participants showed that the comments that were marked with the codes for Concept 2 originated from those who rated the overall usefulness of the sounds highly, with “strongly agree” or “agree” on the Likert scale (Table 6.5)). Perceived control and time was also one of the attributes described by (O’Brien & Toms, 2008) that contribute to user engagement (described in Section 3.2).

In Concept 3, the sounds in the system or the overall audio-visual system were described with words such as engaging, fun, appealing, interesting, stimulating, fascinating, cool, fresh or new. This can be linked with the literature on user engagement, where some of the attributes that contribute to how engaging an application is are characterized by stimulation, sensory appeal, interest, novelty and affect (Section 3.2). A cross-reference of the comments with the participants showed that the comments that were marked with the codes for Concept 3 originated from those who rated the user engagement scales highly or moderately highly, with "strongly agree", "agree" or "neutral" on the Likert scale. It would also contribute to the scales on user satisfaction, due to the word "fun", and a cross-reference with the word fun showed that those participants rated the fun and overall satisfaction scale highly.

Concept 4, in which the sounds were considered to reinforce perception, give additional feedback to actions and events, help orient the participants, focus their attention, or require less reading from the screen all point towards helping in reducing visual load. This can be linked to the models of information navigation described in Section 3.1, in which a part of the navigation process requires the creation of a mental model of the structure. The ideas in this concept would help in this modelling, since it specifies on helping in orienting oneself in the information space, as well as focus the attention. It would also help in reducing cognitive or visual overload and disorientation problems discussed in Section 1.1.

In Concept 5, the sounds were described as useful, helpful or more informative. This means that even if in some cases the same information was available through the visual medium, the presence of auditory feedback helped the users complete the tasks, be it through a smaller number of steps or a general preference to refer to the auditory feedback rather than visual or textual feedback. This would contribute to the perceived usefulness, and a cross-reference found it to be so.

Concept 6, in which the sounds described as pleasant or nice, suggests that the sounds gave rise to a feeling of goodness or liking when using the audio-visual system. This would contribute to the user satisfaction scales, but no conclusive evidence was found from a cross-reference of the comments with the user satisfaction ratings provided by the respective participants.

Concept 7, in which the system with sounds was easier to use or user-friendly, can be linked directly with the perceived ease of use and perceived usability variables. In previous studies, perceived ease of use was found to be correlated with actual intention to use a system (Davis, 1989). A cross-reference of the comments with the participants showed that the comments that were marked with the codes for Concept 7 was found to originate from those who rated the perceived ease of use scales for the sounds highly.

Concept 8, in which the audio-visual system was perceived to be more immersive or interactive, directly relates to the concepts of user engagement and user satisfaction. This suggests that auditory feedback creates a higher sense of presence and immersion rather than a uni-modal system, and increases the sense of interactivity.

Concepts 9 to 11 described suggestions or factors that caused participants to prefer the visual-only system.

In Concept 9, any phrase or wording which suggest that sounds are not necessary or critical to the usefulness of the system are included. This was often provided by the 5 participants who preferred the visual-only system. This is further explained by the following concept, in which the reason for this is justified.

In Concept 10, participants described how they trust their visual system more, or are more used to and comfortable with using vision rather than hearing. The word "trust" itself was used about 4 times by different participants. Other phrases included "more accustomed to", "used to", or "rely more on". This is evidence of the fact that we live in a visually-biased information society, where the use of the visual medium is so prevalent, it has become a matter of habit.

Concept 11 describes limitations of using sounds in interfaces, such as having to use special equipment like speakers or headphones. It was only provided by one comment by a single participant, but it was worth including as a concept since it makes a valid point against the use of sounds in computer interfaces.

The positive and negative feedback from participants about the individual sounds themselves were also constructive. It can be seen from Table 6.9 that the occurrences of positive concepts were more prevalent, with 75% of the time the participants describing the sounds in a satisfied manner. The negative feedback did highlight what kind of sounds were generally disliked, which will be further discussed in the next section.

6.7. Discussion of results

This chapter dealt with the analysis of all the data acquired during the experimental procedure from 24 participants; quantitative data included completion time and accuracy for 6 tasks in 2 task sets, as well as Likert scale ratings for over 27 items, while qualitative data was composed of a total of 162 separate comments provided by the participants in the post-test questionnaires.

In Section 6.1, a list of questions to be statistically and qualitatively analyzed was provided, followed by the hypotheses to be tested.

Analysis of the time and accuracy data did not show any significant effect of type of system on the results. Hence, the hypotheses H1 and H2 (Table 6.1), which stated that time and accuracy will differ significantly between the two systems, were rejected. Out of the 6 tasks to be tested, only in Tasks 3 and 4 would it be reasonable to expect a difference between systems, since these tasks specifically required the comparison of subjects that had auditory feedback to help in those tasks, but this was not found in the results. One reason for this may be the fact that it is difficult to know if the participant actually used the auditory feedback function while completing the task in the audio-visual system. If there was a way to find the specific participants who did use the auditory feedback during the two tasks, and compare the two systems using that data, it may be possible to derive more conclusive evidence. The other four tasks were there more for qualitative reasons, to enable participants to use the information system in different ways that expose them to the entire set of visual and auditory effects. Hence, no significant difference between time and accuracy was anticipated.

Hypothesis H3 stated that preference between the two systems will differ was supported by the results. 79.2% of the participants preferred the audio-visual system, which is a significant result. This shows that even if performance factors were not enhanced by the presence of auditory feedback, preference and affective reactions of the participants were suitably influenced.

Hypothesis H4 stated that perceived ease of use will differ between the two systems, and the results support this hypothesis, as anticipated. The auditory feedback enhanced the perception of ease of use of the system, and this was also supported from the analysis of the qualitative data. This was not, however, the case for user satisfaction (H5), in which no significant difference was found in the analysis of the Likert scales. In H6, the analysis of the four variables characterizing user engagement, two were found to be significantly different with the audio-visual system being rated higher than the visual-only for aesthetics and perceived usability, but not for novelty and endurability.

One of the reasons for this may be the discrepancy in the number of scale items combined for each variable for analysis. For the novelty, endurability and user satisfaction scales, only two scale items were combined together to give rise to the data for that variable, while the other three combined at least three scale items. This may mean that to get conclusive or meaningful evidence while conducting comparisons of different variables between two systems, there may be the need to ask three or more questions or scale items for each variable. This theory will have to be tested in future work to confirm its validity.

It was expected that relationships may be found between the prior experience variables and preference of the system, for example, higher musical training may lead to a higher chance of preferring the audio-visual system, or a noise-sensitive person may have higher chances of preferring the visual-only system. No such conclusive evidence was, however, found in the correlation analysis. One of the reasons for this may be that the spread of values for the prior experience variables was not sufficient in this group of 24, and also that the two groups in preference, audio-visual and visual-only, were so disparate in size. To find correlation evidence, it may be required to recruit participants from different levels of each variable so as to be able to test this hypothesis adequately.

The overall sounds themselves were rated quite highly, with all scales rated with a majority in the positive group. The only scale item rated slightly lower than the others is the one concerning how skillful participants became while using the sounds, with 54% in the positive group. One of the reasons for this may be that users feel they need more time to get used to or become proficient at using auditory feedback in information navigation and seeking tasks, since sounds are not typical or common in today's information interfaces. This was also confirmed by the interaction effect of System and System Order found in the analysis of time data (discussed in Section 6.3). To quote one of the participants, "If I were given a day with the system, rather than under an hour, I would definitely get more accustomed to the sounds and their functions and be much more proficient and comfortable with the audio-visuals."

Out of the individual sounds, three sounds, the click, end of branch and information density overview, were rated consistently highly in both the Likert scales as well as the free-format comments, while the zoom and depth sounds were not as highly liked (Table 6.9). These ratings and the qualitative data provide evidence that the qualities generally liked by users are those that are short, functional, provide added or feedback information or are complementary to the visual feedback. While using auditory icons, one of the reasons a specific sound may not be liked is if the metaphor used to represent the function is not meaningful to the participant. For example, the depth sound which was represented by the combination of birds and leaves rustling using a forest metaphor, generated quite dichotomous opinions. Participants either liked the metaphor since it was pleasant, soothing and not invasive, or they disliked the forest metaphor in the current context of information search.

Another concept generated by the comments is that of using continuous sounds, instead of sounds that come on and off, which can be "distracting". Especially in the situations where short sounds cannot be used to convey sufficient information, a good idea may be to use continuous sounds instead, in the form of unobtrusive background soundscapes whose parameters change at specific events (an example is the Audio Aura system described in Chapter 2 (Mynatt et al., 1998)). This idea was discussed during the panel sessions in the sound design study, but was discarded by the panel members as it was thought that it may cause more annoyance and distraction if the sounds did not stop. The concept itself was generated in the qualitative data by only three occurrences, but it may be an interesting direction in future research to compare audio-visual systems that employ continuous and intermittent sounds as auditory feedback.

CHAPTER 7
CONCLUSION

7.1. Motivation, contributions and recommendations

As has been repeatedly observed in previous research (Brewster, 2003; Fernstrom & McNamara, 2005) etc., computer interfaces, and hence computer-information interfaces, are heavily visual, resulting in a visually-biased information society. There have been huge strides in the domain of auditory features in recent times, due to the need of auditory feedback in devices such as wearable or hand-held ones, e.g. cell-phones, PDAs, smart-watches, mobile navigation systems, digital e-readers and tablets. Other than the use of non-speech sounds as the typical alerts or notifications of events such as an incoming call or message or as feedback in games, they are also used as simple aids in tasks, such as feedback when typing (in touch interfaces that have virtual keyboards), swiping a page in an e-reader, going over the speed-limit detected by a mobile GPS, etc.

Sounds have always been used by us to consciously or subconsciously understand, monitor, and interpret our surroundings. The above examples serve to highlight how nonspeech sound, specifically designed for such purposes, is also slowly being integrated into the ubiquitous technology around us as well. With the rapidly growing complexity and size of information sources in the electronic space, this use of auditory features merits further research in the field of human-information interaction using technology. The implementation of sound in computer-information interfaces would also allow increased ubiquity for such systems, since it would not require the constant focus of the user's visual system.

This research focused on the investigation of the design and evaluation of auditory feedback on a specific hierarchical information system. It combined approaches from the disciplines of human-computer interaction (HCI), information science (IS) and auditory displays: user-centered sound design, evaluation of an information system and the measurement of the overall user experience. As such, three main contributions to knowledge can be outlined:

- The application and extension of an existing participatory design method for the design of nonspeech sounds for novel interfaces and sighted users.
- The development of a measurement model for user experience that combines measures from 3 bodies of literature: technology acceptance, user satisfaction and user engagement.
- The design of a within-subjects evaluation experiment for the comparison of a visual-only and audio-visual system that can be extended to other types of systems. This included the systematic approach to the analysis of both quantitative and qualitative data.

The first step in this research required an in-depth literature review of ideas in the three disciplines mentioned above (spread out over Chapters 2, 3, 4 and 5).

The next step involved the selection and implementation of the semiotic sound design methodology described in Chapter 4. This followed directly from the first research sub-question posed in Chapter 1.

> **What kind of non-speech auditory cues can be designed to convey information in a hierarchical information system using a participatory design method?**

The sound design approach involved the creation of a use scenario appropriate for the context of the hierarchical information visualization and the sound functions to be targeted. Three panels or groups of potential end-users were used to brainstorm, come up with sound ideas, and in subsequent panel sessions, critically evaluate the sounds and suggest changes or new sound ideas with regard to the use scenario.

This led to a finalized set of sounds to be selected for each of the auditory feedback functions. A study of the sound design methodology confirmed the results found in (Murphy, 2007) that the user panels help in group confirmation of the design and allowed the generation of creative ideas through group discussion and collaboration, to which the use scenario acted as an important trigger. The iterative sessions with different panel members helped to identify problems throughout the sound design process and lead to more creative input to be processed for each sound.

Some recommendations that can be made following this process are:

- If the system in the use scenario is not fairly familiar to the panel members, displays or demonstrations of the system should be shown, starting from the first panel, especially in the case of sighted users.
- Words in the use scenario to describe the system have to be carefully selected so that the description does not bias the panel members to certain

metaphors more than others to represent the sound functions in the required context.

- Panellists have an almost immediate negative response to sounds that are harsh, abrasive, highly reverberant, loud, long or busy (attention demanding). Rough sounds should be avoided, and as much as possible, high quality sounds, recorded at least at 44.1 kHz with a bit-depth of 16 or more bits, should be used. All sounds should also be scaled to a comfortable intensity level for display to the panels. Any sounds more intense than the others will most likely evoke a negative response in the panel, even if the sound itself is thought to be an appropriate metaphor.
- It was found that auditory icons (environmental or real-life sounds) seem to be generally preferred than earcons (abstract or musical sounds). An exception arises when the panellist is musically trained in the specific musical instrument used in the earcons.
- The original methodology proposed that panellists need not be experts in sound design or the usage of the application. However, after conclusion of the sound design process, one recommendation would be to have one member in the panel who has sound design experience or is a sound engineer. It may also help to include another such sound designer or engineer to help the researcher during the session who can find or play sounds during the discussion that match the panel's ideas, so that panellists can further hone down on the ideas being discussed (the creation of sounds during the session was also suggested in the original methodology (Murphy, 2007)). This would lead to concepts in sound design being introduced in the discussion that may not be noted by a lay-person, such as the specific parameters of a sound, etc. The researcher is supposed to hold a neutral and uninvolved position, so as not contribute to or bias the panel in any way. Hence, a

sound designer or engineer in the panel and as a helper would balance the panels' ideas.

The second research sub-question was:

How can the effect of auditory feedback be evaluated for information navigation tasks in such a system, so as to examine the effect on the entire user experience, which combines user performance, preference and affective reactions?

To address this question, previous research on the evaluation or measurement of user experience were studied in different disciplines. In the HCI and auditory display field, evaluation studies consisted mainly of measuring the performance factors of a system. These usually involved the measurements of events such as response times, task completion times, number of correct responses, recall performance, error rates, and in some cases, appropriateness of sound choice, annoyance factor and perceived ease of use (Table 2.1). In the Information Science (IS) field, user motivation to use an information system was defined by the concepts of perceived usefulness and perceived ease of use. These concepts were evaluated by a set of scales, developed and iterated over time to different versions of the Technology Acceptance Model. Other HCI and IS research also looked into ways to evaluate User Satisfaction and User Engagement. All these concepts were combined and refined to match the specifications of the current study to develop the measurement model proposed in Section 3.2 (Figure 3.11).

The third research sub-question asked:

> **What are the differences between a visual-only hierarchical information system and an audio-visual one in terms of the user experience?**

To address the above question, a controlled experiment was conducted to compare the performance, preference and affective reactions of a group of participants between the two systems, using the measurement model described above. Table 7.1 lists the the hypotheses for the statistical analyses and the acquired results.

Table 7.1. Research hypotheses and actual findings

Hypothesis	**Result**
Time differs significantly between the two systems.	No
Accuracy differs significantly between the two systems.	No
Preference differs between the two systems.	Yes
Perceived ease of use differs between the two systems.	Yes
User satisfaction differs between the two systems.	No
User engagement differs between the two systems.	Yes
Preference and the prior experience variables correlated	No

The findings showed that even though the performance of participants did not have any significant difference between the two systems, affective factors such as preference, perceived ease of use and user engagement attributes such as aesthetics and perceived usability were favourably affected in the audio-visual system. More than 79% of the users preferred the system integrated with auditory feedback over the silent system. Other than the statistical analyses of the quantitative data, content analysis of the qualitative data also showed a marked preference for the audio-visual system. These preference results by themselves illustrate the positive effects of non-speech sound augmentation on hedonic factors in users, even if used in utilitarian systems such as information systems. This also confirms the applicability of user-centered sound design methods for such purposes.

The findings of this study are specific to the tasks tested, sounds designed and the restricted demographic of Science or Engineering student participants. However, the results may be extended to students from other faculties, with a library collection under those areas, carrying out similar information navigation and retrieval tasks. Although, the results may not be generalizable to widely varying system types or auditory feedback, this study does present a case about how auditory feedback, perceived as meaningful and engaging, can be designed, used and evaluated in interactive information systems. The research showcases how the addition of a collaboratively designed auditory feedback system for structural and navigational aids in a visual hierarchical information system served to enhance the overall user experience.

Table 7.2 outlines some of the findings of interest that contribute to knowledge in various areas of research.

The results of this research leads to better insights into how the use of modalities other than the visual medium, in interactive information systems, can enhance user experience and engagement. Where the research paradigm in HCI and human-information interaction (HII) is shifting from focusing primarily on usability and efficiency to include the entire user experience (Hassenzahl & Tractinsky, 2006), this research is a step in the new direction.

Table 7.2: Research findings per research area

Research area	**Findings of interest**
Information Science; Information systems; Information navigation; Information retrieval;	- Auditory feedback in information systems helps in the creation of a mental model of the information structure, as defined by R. Spence (1999) and reduces visual workload (M. Brown et al., 1989). This is supported by the emerging concepts of additional feedback from sounds, having to read less on-screen, being able to orient oneself better in the information space, etc. - Using a sonic overview to provide a quick overview of the contents of an information source was found to be helpful and appealing. This is supported by Shneiderman's (1996) information seeking theory and followed up by other studies (Kildal & Brewster, 2006; Murphy, 2007). - Short, meaningful sounds, with easily understandable metaphors, such as the end of branch sound, give rise to the highest popularity with users. - Although auditory feedback may not explicitly aid in simple or complex search tasks, they may still "add to the experience".

Table 7.2: Research findings per research area

Research area	**Findings of interest**
Auditory interface design; Evaluation of auditory displays;	- The extended semiotic sound design approach illustrates how collaborative participation between potential end-users can successfully result in a creative, group-confirmed set of required sounds. - Evaluation of the entire set of sounds as well as each individual sound can be performed by using different variables, such as perceived ease of use, ease of learning and remembering, usefulness and satisfaction. Free-format comments also form an important part of sound evaluation. - Preference of auditory icons as opposed to earcons are suggested in the user-centered design approach.
User experience; User engagement; Measurement scales and surveys;	- A general framework for evaluating IR systems, such as TAM (Davis, 1989) or user satisfaction (Doll & Torkzadeh, 1988) are not adequate for all systems. Other fields relevant to the specific study need to be incorporated into the scales to make it more comprehensive, such as those used in auditory interface or HCI studies for the current research. - A more unified and robust scale needs to be developed with further research and reliability analyses. - Relying mainly on improving utilitarian variables which affect performance is not adequate when trying to enhance user experience with a system. Hedonic variables play an important role in influencing the overall user experience and preference.

7.2. Limitations

As previously discussed, the results of the comparative evaluation showed no significant difference between the overall time on task for the audio-visual and visual-only system. One of the reasons for this may be due to the way the task times were logged; time on task depended only on the time elapsed between the viewing of the task and the submission of the answer, without any knowledge or recording of how the time was spent. Moreover, participants were not specifically instructed to try to complete the tasks as quickly as possible, so as not to interfere with the user experience. In fact, if participants were more engaged with one of the systems, they may spend more time on the tasks due to the enjoyment factor or become so involved with the system that they "lose track of time" (O'Brien & Toms, 2010a), hence causing a counter-effect on the time variable. Also, as was previously mentioned, it was difficult to assess if all participants actually used the auditory feedback functions when completing the tasks in the audio-visual system. More conclusive evidence could be drawn if it was possible to determine the specific participants who did explicitly use the auditory feedback during the tasks, and compare the two systems using the data from those participants. However, it would be extremely difficult, if not impossible, to accurately assess this.

Another problem which became apparent during the analysis of the task times was the difference found between the complex IR task given in the two task sets. One task set asked for books that would help in the problem of repairing a broken computer while the other asked for those in the problem of fixing a leaky faucet. Results showed that the leaky faucet problem caused much higher times than the broken computer, as was discussed in Chapter 6. This was probably due to the Science or Engineering background of the participants, in which case they would be more familiar with solving computer problems than the latter. During the formative evaluation, only a subset of the tasks were used and the leaky faucet one was not one of them. If it had been, it may have led to earlier detection of the discrepancy, before the comparative evaluation.

A limitation of the measurement model of the user experience was that scale items for each hedonic variable had to be selected according to the relevance to the specific tasks and the use of auditory feedback in the system. This led to the fact that some of the variables were measured with as many as five scale items, while some were limited to two. It was found that the number of scale items affected the significance level of the comparison of the two systems, with those variables containing four or more scale items reaching statistical significance more than those with fewer. One of the future directions of this research will be to design a more balanced measurement model, containing a uniform number of scale items per variable, depending on the number determined to be more effective.

The analysis of the prior experience variables, which were expected to influence the preference of the participants, did not show any significant effects either. The fact that the preference groups were so disparate, with one group (those who preferred the audio-visual) containing 79.2% of the total participants, while the other (those who preferred the visual-only) containing only 20.8%, contributed to this result. Since the prior experience variables such as demographics, musical training, noise sensitivity etc. were not used to control the sampling process, this hypothesis could not be conclusively tested, and is another research direction to be pursued in the future.

Another limitation of the study is that this was not a repeated measures study in which each participant is tested at more than one point in time. Hence the effects of familiarity, learning and repeated exposure to the system on the construct of user experience could not be studied. Participants did comment that if they had more time to become skillful at using the audio-visual system, they would be able to use the functionality of the auditory feedback more efficiently. Hence, a longitudinal study designed to study these effects would be another part of future work, as described in the next section.

7.3. Future work

An extension of the current study would be to evaluate the effects of the sonified system on different groups of users to see if user criteria, such as musical training, video gaming experience, noise sensitivity, etc. (the prior experience variables described in Section 5.2) causes the effects to vary across groups. This would entail the comparative study to be conducted between similar size groups of each level of the variable to be studied. It would also be interesting to study the effects of sonifying information visualizations other than a hierarchical information system, such as ones that use clustering paradigms, etc.

An extension to the current sounds itself would be to use a real-time sound synthesis approach, instead of predefined levels of the sounds being used for the different functions. Synthesizing the sounds in real time would allow the sounds to be produced dynamically as the user navigates the system, which may allow a smoother and more synergistic integration with the visuals. Samples of synthetic sounds were displayed to the sound design panels during the design sessions but panellists did not like them and hence they were not part of the final selections.

Another future direction would be to shift from sighted users to visually-impaired users. This would require the design of a wider range of sounds, as well as the addition of speech sounds (to read the contents of the screen when necessary). The addition of speech may, in effect, also be beneficial to sighted users in the situations where the subject labels are visually obscured, since this was a point of objection for a few participants in the study (Table 6.8). It would also include spatialized sounds to enable visually-impaired users to assess the spatial layout of the screen. Spatialization is the use of the localization of sounds by adjusting time differences and levels so as to vary the perceived position of the sounds and give listeners the perception of sound sources within a 3D environment. Hence, this can be used to convey structural information and perceived distance.

Furthermore, spatialization of the non-speech sounds would be a promising direction for research, even in other ways. Previous studies provide evidence, e.g., (Driver & Spence, 1994; C. Spence & Ranson, 2000) for the effective design of multimodal interfaces by illustrating the potential trade-off between visual and auditory arrangements that make it easier to attend to simultaneous relevant information in the two modalities in the same location, or conversely, more difficult to ignore irrelevant information. Hence, it may be beneficial to design multimodal interfaces in such a way that allows sounds immediately relevant to a current task being spatially located near the area of visual focus, and those not relevant, distributed away, in order to facilitate attention or minimize distraction effects. Thus spatializing the sounds in the current information system to study these effects would be an interesting future direction for research.

Spatialization would also be a part of extending the current information system to more immersive environments. For a more immersive experience, the design of 3D soundscapes instead of intermittent sounds would be ideal. The effect of this feeling of immersiveness, which was a concept that generated positive feedback from participants in this study, would be also be a big step in studying how immersive sounds affect the overall user experience.

There is a need to study how multimodal perception can influence the user experience in interactive information systems and how this can be reliably measured using perceptual evaluation scales. One of the future research interests is to conduct a longitudinal study to investigate the effects of multimodal interaction on usability issues and the user experience. Most HCI studies are limited by the fact that they are performed over a short period of time, and hence cannot take into account the long-term effects on user experience. Such a study would require the design of a series of experiments, some to be studied under controlled laboratory settings and some in natural settings, such as the user's own home or workplace, depending on the type of application. Such longitudinal comparative studies between different applications, such as multimodal and unimodal applications, would provide better insight on how multimodal interaction influences the factors of user engagement, satisfaction and the overall user experience.

Another future research direction is to conduct factor and reliability analysis of the current data on the perceptual scales in order to lead to any insights to the development of more robust, reliable and valid measurement tools for the evaluation of the overall user experience. A longitudinal study that can take into account the long-term effects of repeated use, familiarity and recency of prior experience with the studied application may be a more reliable way of achieving this aim. Such a study would allow the examination of how novelty or familiarity of an application affects the construct of user engagement.

REFERENCES

Absar, R., & Guastavino, C. (2008). Usability of non-speech sounds in user interfaces. In *Proceedings of the 14th International Conference on Auditory Display (ICAD '08).* Paris, France.

Absar, R., & Guastavino, C. (2011). Nonspeech sound design for a hierarchical information system. In M. Kurosu (Ed.), *Human centered design* (Vol. 6776, p. 461-470). LNCS, Springer Berlin / Heidelberg.

Ahuja, J., & Webster, J. (2006). Enhancing the design of web navigation systems: The influence of user disorientation on engagement and performance. *Management Information Systems Quarterly*, *30*(3).

Andres, H. P., & Petersen, C. (2002). Presentation media, information complexity, and learning outcomes. *Journal of Educational Technology Systems*, *30*(3), 225–246.

Arons, B., & Mynatt, E. (1994). The future of speech and audio in the interface. *SIGCHI Bulletin*, *26*(4), 44-48.

Bagozzi, R. P. (2007). The legacy of the technology acceptance model and a proposal for a paradigm shift. *Journal of the Association for Information Systems*, *8*(4).

Available from http://aisel.aisnet.org/jais/vol8/iss4/12

Bailey, J. E., & Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science*, *29*(5), pp. 530-545.

Ballas, J. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Experimental Psychology, Human Perception and Performance*, *19*, 250-267.

Barfield, W., Rosenberg, C., & Levasseur, G. (1991). The use of icons, earcons and commands in the design of an online hierarchical menu. *IEEE Transactions on Professional Communication*, *34*(2), 101 - 108.

Barrass, S. (1996). EarBenders: Using stories about listening to design auditory interfaces. In *Proceedings of the First Asia-Pacific Conference on Human Computer Interaction APCHI '96* (pp. 525–538).

Barrass, S., & Kramer, G. (1999). Using sonification. *Multimedia Systems*, *7*(1), 23–31.

Battarbee, K., & Koskinen, I. (2005). Co-experience: user experience as interaction. *Codesign*, *1*, 5–18.

Benovoy, M., Zadel, M., Absar, R., Wozniewski, M., & Cooperstock, J. (2008). Towards immersive multimodal gameplay. In *GAMEON-NA.* Montreal, QC, Canada.

Benyon, D., & McCall, R. (1999). EniSpace: evaluating navigation in information space. In *WebNet '99* (p. 1344-1345).

Bladh, D. C., T., & Scholl, J. (2004). Extending tree-maps to three dimensions: a comparative study. In *Proc. of 6th Asia-Pacific Conference on Computer-Human Interaction (APCHI 2004)* (p. 50-59).

Blattner, M., Sumikawa, D., & Greenberg, R. (1989). Earcons and icons: Their structure and common design principles. *Human Computer Interaction*, *4*(1), 11-44.

Bodker, S. (1999). Scenarios in user-centered design - setting the stage for reflection and action. In *Proceedings of the 1999 Hawaii Internation Conference on System Sciences* (p. 3053-3063). Maui, Hawaii.

Brewster, S. (1997a). Navigating telephone-based interfaces with earcons. In *HCI 97: Proceedings of HCI on People and Computers XII* (pp. 39–56). London, UK: Springer-Verlag.

Brewster, S. (1997b). Using non-speech sound to overcome information overload. *Displays*, *17*(3), 179-189.

Brewster, S. (1998a). The design of sonically-enhanced widgets. *Interacting with Computers*, *11*(2), 211-235.

Brewster, S. (1998b). Using nonspeech sounds to provide navigation cues. *ACM Transactions on Computer-Human Interaction*, *5*(3), 224–259.

Brewster, S. (2003). Nonspeech auditory output. *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications*, 220–239.

Brewster, S., Leplatre, G., & Crease, M. (1998). Using non-speech sounds in mobile computing devices. In J. C. (Ed.), *Proceedings of the First Workshop on Human Computer Interaction with Mobile Devices* (p. 26 - 29). Department of Computing Science, University of Glasgow, Glasgow, UK.

Brewster, S., Räty, V. P., & Kortekangas, A. (1996). Earcons as a method of providing navigational cues in a menu hierarchy. In *HCI '96: Proceedings of HCI on People and Computers XI* (pp. 169–183). London, UK: Springer-Verlag.

Brewster, S., Wright, P. C., & Edwards, A. D. N. (1992). A detailed investigation into the effectiveness of earcons. In *Proceedings of International Conference on Auditory Display (ICAD '92)* (p. 471 - 498). Santa Fe Institute, Santa Fe:

Addison-Wesley.

Brewster, S., Wright, P. C., & Edwards, A. D. N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In *CHI '93: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 222–227). ACM Press.

Brown, L., Brewster, S., Burton, M., Riedel, B., & Ramloll, R. (2003). Design guidelines for audio presentation of graphs and tables. In *Proceedings of the International Conference on Auditory Display (ICAD '03).* Boston, MA, USA.

Brown, M., Newsome, S. L., & Glinert, E. P. (1989). An experiment into the use of auditory cues to reduce visual workload. In *CHI '89: Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 339–346). New York, NY, USA: ACM Press.

Bussemakers, M., & Haan, A. de. (1998). Using earcons and icons in categorization tasks to improve multimedia interfaces. In *Proceedings of the International Conference on Auditory Display (ICAD '98).* Glasgow, UK.

Bussemakers, M., & Haan, A. de. (2000). When it sounds like a duck and it looks like a dog... auditory icons vs. earcons in multimedia environments. In P. R. Cook (Ed.), *Proceedings of the International Conference on Auditory Display (ICAD '00)* (p. 184 - 189).

Bussemakers, M., Haan, A. de, & Lemmens, P. M. C. (1999). The effect of auditory accessory stimuli on picture categorisation: implications for interface design. In *Proceedings of the 8th HCI International Conference on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I* (pp. 436–440). Mahwah, NJ, USA: Lawrence Erlbaum Associates, Inc.

Calvert, G., Spence, C., & Stein, B. E. (2004). *The handbook of multisensory processes.* MIT Press.

Carroll, J. M. (2000). *Making use: scenario-based design of human-computer interactions.* MIT Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conklin, J. (1987). Hypertext an introduction and survey. *IEEE Computer*, *20*(9), 17 41.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems : theory and results.* Unpublished doctoral dissertation, Massachusetts Institute of Technology, Sloan School of Management.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340.

Desmet, P. M. A., & Hekkert, P. (2007). Framework of product experience. *International Journal of Design*, *1*(1), 57-66.

Dillon, A., McKnight, C., & Richardson, J. (1990). Navigation in hypertext: a critical review of the concept. In *Proceedings of Interact '90* (p. 587-592).

Doll, W. J., & Torkzadeh, G. (1988). The measurement of End-User computing satisfaction. *MIS Quarterly*, *12*(2), 259–274.

Dourish, P., & Chalmers, M. (1994). Running out of space: Models of information navigation. In *HCI, glasgow.*

Driver, J., & Spence, C. (1994). Spatial synergies between auditory and visual attention. In C. Umilta & M. Moscovitch (Eds.), *Attention and Performance XV: Conscious and nonconcious information processing* (p. 311-331). MIT Press: Cambridge, MA.

Edwards, G. W., Barfield, W., & Nussbaum, M. A. (2004). The use of force feedback and auditory cues for performance of an assembly task in an immersive virtual environment. *Virtual Reality*, *7*, 112-119.

Edworthy, J. (1998). Does sound help us to work better with machine? a commentary on Rauterberg's paper 'about the importance of auditory alarms during the operation of a plant simulator'. *Interacting with Computers*, *10*, 401- 409.

Ergonomic requirements for office with visual display terminals - guidance on usability [Computer software manual]. (1998). Beuth, Berlin.

Erickson, T. (1995). Notes on design practice: stories and prototypes as catalysts for communication. In (p. 3758). New York, NY, USA: John Wiley & Sons, Inc. Available from `http://dl.acm.org/citation.cfm?id=209227.209231`

Feld, S. (1990). *Sound and sentiment: birds, weeping, poetics, and song in kaluli expression.* Philadelphia: University of Pennsylvania Press.

Fernström, M., & Brazil, E. (2004). Human-computer interaction design based on interactive sonification hearing actions or instruments/agents. In *Proceedings of the 2004 International Workshop on Interactive Sonification.* Bielefeld University, Germany.

Fernström, M., Brazil, E., & Bannon, L. (2005). HCI design and interactive sonification for fingers and ears. *IEEE MultiMedia*, *12*(2), 36–44.

Fernstrom, M., & McNamara, C. (2005). After direct manipulation direct sonification. *ACM Transactions on Applied Perception*, *2*(4), 495499. (ACM ID: 1101548)

Flowers, J. H., Buhman, D. C., & Turnage, K. D. (2005). Data sonification from the desktop: Should sound be part of standard data analysis software? *ACM Transactions on Applied Perception*, *2*(4), 467472. Available from `http://doi.acm.org/10.1145/1101530.1101544`

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430-441.

Franinovic, K., Hug, D., & Visell, Y. (2007). Sound embodied: Explorations of sonic interaction design for everyday objects in a workshop setting. In *Proceedings of the 13th International Conference on Auditory Display.*

Frauenberger, C., & Stockman, T. (2006). Patterns in auditory menu design. In *Proceedings of the International Conference on Auditory Display (ICAD* (pp. 141–147).

The Freesound Project. (n.d.). http://www.freesound.org.

Garzonis, S., Jones, S., Jay, T., & O'Neill, E. (2009). Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (p. 15131522). New York, NY, USA: ACM. Available from `http://doi.acm.org/10.1145/1518701.1518932`

Gatian, A. W. (1994). Is user satisfaction a valid measure of system effectiveness? *Information & Management*, *26*(3), 119–131.

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, *2*(2), 167-177.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human-Computer Interaction*, *4*(1), 67-94.

Gaver, W. W. (1993). What in the world do we hear? An ecological approach to auditory source perception. *Ecological Psychology*, *5*(1).

Gaver, W. W., Smith, R. B., & O'Shea, T. (1991). Effective sounds in complex systems: the ARKOLA simulation. In *CHI '91: Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 85–90). New York, NY, USA: ACM Press.

Gelderman, M. (1998). The relation between user satisfaction, usage of information systems and performance. *Information & Management*, *34*(1), 11–18.

Goose, S., & Möller, C. (1999). A 3D audio only interactive web browser: using spatialization to convey hypermedia document structure. In *MULTIMEDIA '99: Proceedings of the Seventh ACM International Conference on Multimedia (Part 1)* (pp. 363–371). New York, NY, USA: ACM Press.

Graham, R. (1999). Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. *Ergonomics*, *42*(9), 1233 1248.

Griffiths, J. R., Johnson, F., & Hartley, R. J. (2007). User satisfaction as a measure of system performance. *Journal of Librarianship and Information Science*, *39*(3), 142 –152.

Grohn, M., Lokki, T., & Takala, T. (2003). Comparison of auditory, visual, and audiovisual navigation in a 3D space. In *Proceedings of the 9th International Conference on Auditory Display (ICAD '03).*

Guastavino, C. (2006). The ideal urban soundscape: Investigating the sound quality of french cities. *Acta Acustica united with Acustica*, *92*(6), 945-951.

Guastavino, C. (2007). Categorization of environmental sounds. *Canadian Journal of Experimental Psychology*, *60*(1), 54-63.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & Information Technology*, *25*(2), 91–97.

Hearst, M. A., & Karadi, C. (1997). Cat-a-cone: an interactive interface for specifying searches and viewing retrieval results using a large category hierarchy. In *SIGIR '97: Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 246–255). New York, NY, USA: ACM.

Hendrickson, A. R., Massey, P. D., & Cronan, T. P. (1993). On the Test-Retest reliability of perceived usefulness and perceived ease of use scales. *MIS Quarterly*, *17*(2), 227–230.

Heron, J. (1996). *Cooperative inquiry: Research into the human condition*. London: Sage.

Holland, S., Morse, D. R., & Gedenryd, H. (2002). AudioGPS: Spatial audio navigation with a minimal attention interface. *Personal Ubiquitous Computing*, *6*(4), 253–259.

Holtzblatt, K. (2001). Contextual design : Experience in real life. *Computer*, *1*, 19–22.

Isa, B. S., Ogden, W. C., Wolfe, S. J., & Korenshtein, R. (1986). Navigation issues related to menu-based software products. *SIGCHI Bulletin*, *18*(2), 68–69.

James, F. (1998). Lessons from developing audio html interface. In *Proceedings of the 1998 International ACM Conference on Assistive Technologies.*

Jennings, M. (2000). Theory and models for creating engaging and immersive ecommerce websites. In *Proceedings of the 2000 ACM SIGCPR Conference on Computer Personnel Research* (p. 7785). New York, NY, USA: ACM. Available from http://doi.acm.org/10.1145/333334.333358

Jul, S., & Furnas, G. W. (1997). Navigation in electronic worlds: A chi'97 workshop. In *ACM SIGCHI Bulletin* (p. 44-49).

Julien, C. A. (2010). *SE-3D: a controlled comparative usability study for a virtual reality semantic hierarchy explorer*. Unpublished doctoral dissertation, McGill University, Montreal, QC, Canada.

Julien, C. A., Guastavino, C., Bouthillier, F., & Leide, J. E. (2010). Subject Explorer 3D: a virtual reality collection browsing and searching tool. In *Proceedings of the 38th Annual Canadian Association for Information Science Conference (CAIS 2010).*

Keim, D., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006). Challenges in visual data analysis. In *Proceedings of Information Visualization (IV 2006), IEEE* (p. 9-16).

Kenny, D. A. (1987). *Statistics for the social and behavioral sciences.* Little, Brown.

Kieffer, S., & Carbonell, N. (2006). Oral messages improve visual search. In *AVI '06: Proceedings of the Working Conference on Advanced Visual Interfaces* (pp. 369–372). New York, NY, USA: ACM Press.

Kildal, J., & Brewster, S. (2006). Exploratory strategies and procedures to obtain non-visual overviews using tablevis. *International Journal of Disability and Human development*, *5*(3), 285-294.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56*(5), 746-759.

Kourouthanassis, P. E., Giaglis, G. M., & Vrechopoulos, A. P. (2007). Enhancing user experience through pervasive information systems: The case of pervasive retailing. *International Journal of Information Management*, *27*(5), 319–335.

Kramer, G., Walker, B. N., Bonebright, T., Cook, P., Flowers, J., Miner, N., et al. (1999). *Sonification report: Status of the field and research agenda.* Report prepared for the National Science Foundation by members of the International Community for Auditory Display. Santa Fe, NM: International Community for Auditory Display (ICAD).

Laarni, J., Ravaja, N., Kallinen, K., & Saari, T. (2004). Transcendent experience in the use of computer-based media. In *Proceedings of the Third Nordic Conference on Human-Computer Interaction - NordiCHI '04* (pp. 409–412). Tampere, Finland.

Large, A., Beheshti, J., Tabatabaei, N., & Nesset, V. (2009). Developing a visual taxonomy: Children's views on aesthetics. *Journal of the American Society for Information Science and Technology*, *60*(9), 1808-1822.

Laurel, B. (1993). *Computers as theatre* (1st ed.). Addison-Wesley Professional.

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & Management*, *40*(3), 191–204. Available from `http://www.sciencedirect.com/science/article/pii/S0378720601001434`

Lemmens, P. M. C., Bussemakers, M., & Haan, A. de. (2000). The effects of earcons on reaction times and error-rates in a dual task vs. a single task experiment. In P. R. Cook (Ed.), *Proceedings of International Conference on Auditory Display (ICAD '00)* (p. 177183).

Lemmens, P. M. C., Bussemakers, M., & Haan, A. de. (2001). Effects of auditory icons and earcons on visual categorization: The bigger picture. In *Proceedings of International Conference on Auditory Display (ICAD '01).*

Leplatre, G., & Brewster, S. (2000). Designing non-speech sounds to support navigation in mobile phone menus. In P. R. Cook (Ed.), *Proceedings of the 6th International Conference on Auditory Display (ICAD '00)* (p. 190 - 199).

Leplatre, G., & McGregor, I. (2004). How to tackle auditory interface aesthetics? discussion and case study. In *Proceedings of the 2004 International Conference on Auditory Display.*

Lindgaard, G., & Dudek, C. (2003). What is this evasive beast we call user satisfaction? *Interacting with Computers*, *15*(3), 429–452.

LoPresti, E., & Harris, W. M. (1996). loudSPIRE, an auditory display schema for the SPIRE system. In *Proceedings of International Conference on Auditory Display (ICAD '96).*

LoPresti, W. M., E. M.; Harris. (1997). Sonic exploration of thematic information. In *Audio Engineering Society Convention 102.*

Lucas, P. (1994). An evaluation of the communicative ability of auditory icons and earcons. In *Proceedings of the International Conference on Auditory Display*

(ICAD '94).

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability and User Experience*, *8*(2).

Macaulay, C., Benyon, D., & Crerar, A. (1998). Voices in the forest: Sounds, soundscapes and interface design [Technical Report]. In N. Dahlback (Ed.), *Exploring navigation: Towards a framework for design and evaluation of navigation in electronic spaces* (pp. 161–171). Kista, Sweden: Swedish Institute of Computer Science.

Mayer, R., & Moreno, R. (1998). A Split-Attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*(2), 312–320.

Minghim, R., & Forrest, A. R. (1995). An illustrated analysis of sonification for scientific visualisation. In *VIS '95: Proceedings of the 6th conference on Visualization '95* (p. 110). Washington, DC, USA: IEEE Computer Society.

Moffatt, K., McGrenere, J., Purves, B., & Klawe, M. (2004). The participatory design of a sound and image enhanced daily planner for people with aphasia. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems* (p. 407414). New York, NY, USA: ACM. Available from `http://doi.acm.org/10.1145/985692.985744`

Morley, S., Petrie, H., O'Neill, A., & McNally, P. (1999). Auditory navigation in hyperspace: design and evaluation of a non-visual hypermedia system for blind users. *Behaviour and Information Technology*, *18*(1), 18-26.

Morse, L. M. . O. K., E. (2002). Testing visual information retrieval methodologies case study: comparative analysis of textual, icon, graphical and 'spring' displays. *Journal of the American Society for Information Science and Technology*, *53*(1), 28-40.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes,. *Journal of Educational Psychology*, *87*(2), 319–334.

Muller, M. J. (2003). Participatory design: the third space in HCI. *Design*, *4235*, 1–70.

Murphy, E. (2007). *Designing auditory cues for a multimodal web interface: A semiotic approach.* Unpublished doctoral dissertation, Queen's University, Belfast, Ireland.

Murphy, E., Kuber, R., Strain, P., McAllister, G., & Yu, W. (2007). Developing sounds for a multimodal interface: conveying spatial information to visually impaired web users. In *Proceedings of the 13th International Conference on Auditory Display (ICAD '07)* (p. 348-355).

Murphy, E., Pirhonen, A., McAllister, G., & Yu, W. (2006). A semiotic approach to the design of non-speech sounds. In *Proceedings of the 2006 International Workshop Haptic and Audio Interaction Design* (p. 121-132).

Mynatt, E. (1994). Designing with auditory icons: how well do we identify auditory cues? In *CHI '94: Conference companion on Human Factors in computing systems* (pp. 269–270). New York, NY, USA: ACM Press.

Mynatt, E. (1997). Transforming graphical interfaces into auditory interfaces for blind users. *Human-Computer Interaction*, *12*, 7-45.

Mynatt, E., Back, M., Want, R., Baer, M., & Ellis, J. B. (1998). Designing audio aura. In *CHI '98: Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 566–573). New York, NY, USA: ACM Press/Addison-Wesley Publishing Co.

Mynatt, E., & Weber, G. (1994). Nonvisual presentation of graphical user interfaces: contrasting two approaches. In *CHI '94: Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 166–172). New York, NY, USA:

ACM Press.

Newby, G. (1992). An investigation of the role of navigation for information retrieval. In *Proceedings of the 55th annual meeting of the American Society of Information Science (ASIS '92)* (Vol. 29, p. 20-25).

Newby, G. (2002). Empirical study of a 3D visualization for information retrieval tasks. *Journal of Intelligent Information Systems*, *18*(1), 31–53.

O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, *59*(6), 938–955.

O'Brien, H. L., & Toms, E. G. (2010a). 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.

O'Brien, H. L., & Toms, E. G. (2010b). Is there a universal instrument for measuring interactive information retrieval?: The case of the user engagement scale. In *Proceedings of the Third symposium on Information Interaction in Context* (p. 335340). New Brunswick, New Jersey, USA: ACM.

Oppenheim, A. (1992). *Questionnaire design, interviewing and attitude measurement* (2nd Ed ed.). Continuum International Publishing.

Overbeeke, K., Djajadiningrat, T., Hummels, C., Wensveen, S., & Frens, J. (2003). Funology. In M. A. Blythe, A. F. Monk, K. Overbeeke, & P. C. Wright (Eds.), (Vol. 3, pp. 7–17). Dordrecht, The Netherlands: Kluwer.

Oviatt, S. (1997). Multimodal interactive maps: Designing for human performance. *Human Computer Interaction*, *12*(1-2), 93-129.

Oviatt, S., Coulston, R., & Lunsford, R. (2004). When do we interact multimodally? cognitive load and multimodal communication patterns. In *Proceedings of the 6th International Conference on Multimodal Interfaces - ICMI '04* (p. 129). State College, PA, USA.

Parente, P., & Bishop, G. (2003). Bats: the blind audio tactile mapping system. In *Proceedings of ACM South Eastern Conference.* Savannah, GA.

Pirhonen, A., & Murphy, E. (2008). Designing for the unexpected: The role of creative group work for emerging interaction design paradigms. *Journal of Visual Communication Special Issue on Wearable Technology*, *7*(3), 331–344.

Pirolli, P., Card, S. K., & van der Wege, M. M. (2000). The effect of information scent on searching information visualizations of large tree structures. In *Proceedings of International Conference on Advanced Visual Interfaces* (p. 161-172). Palermo, IT.

Plaisant, C. (2005). Information visualization and the challenge of universal usability. In M.-J. K. J. Dykes A.M. MacEachren (Ed.), (chap. 3). Elsevier.

Quesenbery, W. (2003). Dimensions of usability. Content and complexity: Information design in technical communications. In M. Albers & B. Mazur (Eds.), (p. 81-102). Mahwah, N.J.: Lawrence Erlbaum.

Quintana, C., Carra, A., Krajcik, J., & Soloway, E. (2001). Learner-centered design: Reflections and new directions. In *Human-Computer Interaction in the New Millennium* (p. 605-624).

Rabenhorst, D. A., Farrell, E. J., Jameson, D. H., Linton, T. D., & Mandelman, J. A. (1990). Complementary visualization and sonification of multi-dimensional data. In *Proceedings of SPIE - Extracting Meaning from Complex Data* (Vol. 1259, p. 147-153).

Ramloll, R., Brewster, S., Yu, W., & Riedel, B. (2001). Using non-speech sounds to improve access to 2D tabular numerical information for visually impaired users. In *Proceedings of IHM-HCI '01* (p. 515 - 530).

Robertson, G. G., Mackinlay, J. D., & Card, S. K. (1991). Cone trees: animated 3d visualizations of hierarchical information. In *CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 189–194).

New York, NY, USA: ACM.

Rozell, E. J., & Gardner, W. L. (2000). Cognitive, motivation, and affective processes associated with computer-related performance: a path analysis. *Computers in Human Behavior*, *16*(2), 199–222.

Salvador, V., Minghim, R., & Pacheco, M. (1998). Sonification to support visualization tasks. In *SIBGRAPHI '98: Proceedings of the International Symposium on Computer Graphics, Image Processing, and Vision* (p. 150). Washington, DC, USA: IEEE Computer Society.

Sand, D. (1996). *Designing large-scale web sites.* New York: Wiley Computer Publishing.

Sawhney, N., & Schmandt, C. (2000). Nomadic radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Transactions on Computer-Human Interactions*, *7*(3), 353–383.

Scaife, M., Rogers, Y., Aldrich, F., & Davies, M. (1997). Designing for or designing with? informant design for interactive learning environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (p. 343350). New York, NY, USA: ACM. Available from `http://doi.acm.org/10.1145/258549.258789`

Schuler, D., & Namioka, A. (1993). *Participatory design: principles and practices.* Routledge.

Sharp, J. H. (2007). Development, extension, and application: a review of the technology acceptance model. *Information Systems Education Journal*, *5*(9).

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages* (p. 336-343).

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human computer interaction* (3rd ed. ed.). Reading, MA: Addison-Wesley.

Sikora, C. A., Roberts, L., & Murray, L. (1995). Musical vs. real world feedback signals. In *CHI '95: Conference companion on Human Factors in computing systems* (pp. 220–221). New York, NY, USA: ACM Press.

Skantze, D., & Dahlbck, N. (2003). Auditory icon support for navigation in speech-only interfaces for room-based design metaphors. In *Proceedings of the 2003 International Conference on Auditory Display (ICAD '03).* Boston, MA.

Slater, M., & Usoh, M. (1993). Presence in immersive virtual environments. In *Virtual Reality Annual International Symposium* (p. 90-96).

Spence, C., & Ranson, J. (2000). Cross-modal selective attention: On the difficulty of ignoring sounds at the locus of visual attention. *Perception and Psychophysics*, *62*(2), 410-424.

Spence, R. (1997). Towards a framework for organised thought about navigation. In *Position paper for Workshop on Navigation in Electronic Worlds.*

Spence, R. (1999). A framework for navigation. *International Journal of Human-Computer Studies*, *51*(5), 919-945.

Sutcliffe, A., Ennis, M., & Hu, J. (2000). Evaluating the effectiveness of visual user interfaces for information retrieval. *International Journal of Human-Computer Studies*, *53*(5), 741-763.

Teder-Slejrvi, W. A., McDonald, J. J., Di Russo, F., & Hillyard, S. A. (2002). An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. *Brain Research. Cognitive Brain Research*, *14*(1), 106–114.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997, December). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, *3*(4), 257–287.

The University of Iowa music instrument samples database. (n.d.). http://theremin.music.uiowa.edu.

van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS Quarterly*, *28*(4), 695–704.

van Ham, F., & van Wijk, J. (2003). Beamtrees: compact visualization of large hierarchies. *Information Visualization*, *2*, 31-39.

Vargas, M. L. M., & Anderson, S. (2003). Combining speech and earcons to assist menu navigation. In *Proceedings of the 2003 International Conference on Auditory Display* (p. 38-41).

Venkatesh, V., & Davis, F. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, *46*(2), 186–204.

Walker, B. N., & Kramer, G. (2004). Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. In J. G. Neuhoff (Ed.), (chap. 6). New York: Academic Press.

Wittenberg, K. (1997). Navigation and search - what's the difference? In *Position paper for Workshop on Navigation in Electronic Worlds.*

Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research*, *16*(1), 85–102.

Xiao, L., & Dasgupta, S. (2002). Measurement of user satisfation with web-based information systems: an empirical study. In *Proceedings of the Eigth Americas Conference on Information Systems* (pp. 1149–1155).

Yalla, P., & Walker, B. N. (2008). Advanced auditory menus: design and evaluation of auditory scroll bars. In *Assets '08: Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility* (pp. 105–112). New York, NY, USA: ACM.

Yee, K. P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. In *Proceedings of CHI 2003 conference* (p. 401-408). Ft. Lauderdale, Florida, USA.

Yousafzai, S. Y., Foxall, G. R., & Pallister, J. G. (2007a). Technology acceptance: a meta-analysis of the TAM: part 1. *Journal of Modelling in Management*, *2*(3), 251–280.

Yousafzai, S. Y., Foxall, G. R., & Pallister, J. G. (2007b). Technology acceptance: a meta-analysis of the TAM: part 2. *Journal of Modelling in Management*, *2*(3), 281–304.

Yu, B.-M., & Roh, S.-Z. (2002). The effects of menu design on information-seeking performance and user's attitude on the world wide web. *Journal of the American Society for Information Science and Technology*, *53*(11), 923-933.

Zhang, Y., Fernando, T., Xiao, H., & Travis, A. R. L. (2006). Evaluation of auditory and visual feedback on task performance in a virtual assembly environment. *Presence*, *15*(6), 613-626.

Zhao, H. (2005). Interactive sonification of geo-referenced data. In *CHI '05 extended abstracts on Human Factors in Computing Systems* (p. 11341135). New York, NY, USA: ACM.

Zhao, H., Plaisant, C., & Shneiderman, B. (2005). "I hear the pattern": interactive sonification of geographical data patterns. In *CHI '05 extended abstracts on Human Factors in Computing Systems* (p. 19051908). New York, NY, USA: ACM.

Zhao, H., Plaisant, C., Shneiderman, B., & Duraiswami, R. (2004). Sonification of geo-referenced data for auditory information-seeking: design principle and pilot study. In *Proceedings of ICAD '04 Tenth Meeting of the International Conference on Auditory Display.* Sydney, Australia.

Zhao, H., Smith, B. K., Norman, K., Plaisant, C., & Shneiderman, B. (2005). Interactive sonification of choropleth maps. *IEEE Multimedia*, *12*(2), 26–35.

Zhao, S., Dragicevic, P., Chignell, M., Balakrishnan, R., & Baudisch, P. (2007). earpod: Eyes-free menu selection using touch input and reactive audio feedback. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 1395–1404). New York, NY, USA: ACM.

LIST OF ACRONYMS

3D: Three-Dimensional

AE: Aesthetics, an attribute used in the User Engagement Scale

EN: Endurability, an attribute used in the User Engagement Scale

EUCS: End-User Computing Satisfaction

HCI: Human-Computer Interaction

HII: Human-Information Interaction

HTML: Hyper-Text Markup Language

IIR: Interactive Information Retrieval]

IR: Information Retrieval

IS: Information Science

IV: Information Visualization

LCSH: Library of Congress Subject Headings

NO: Novelty, an attribute used in the User Engagement Scale

PU: Perceived Usability, an attribute used in the User Engagement Scale

RMS: Root-Mean Square

SE-3D: Subject Explorer in 3D

TAM: Technology Acceptance Model

UE: User Engagement

UES: User Engagement Scale

UX: User eXperience

VR: Virtual Reality

Appendices

1. Appendix A

1.1. Sound design. The informed consent form for the sound design study is attached in the next page.

Informed consent form - McGill University

WHY ARE WE DOING THIS RESEARCH? Our goal is to establish scientific knowledge about designing appropriate audio cues for information navigation. It has nothing to do with your personality or motivations or intelligence.

PRIVACY. All the responses that we collect will be aggregated for statistical analysis and so your own responses will be averaged with those of many other people. We know that you value your privacy. You will not be identified as an individual in any scientific report of this research, and your name will not be linked to your responses in this study.

DISCUSSION OF RESEARCH IDEAS. We cannot discuss our ideas with you before the experiment takes place, but we will be happy to talk with you about our hypotheses and theories afterwards.

WHAT WILL HAPPEN DURING THE EXPERIMENT? You will be seated in a quiet room, sometimes hear sounds over speakers and participate in a group discussion about sound effects. The experiment will not cause you discomfort. There is no known risks associated with the experiment. You will be free to discontinue your participation at any time without penalty.

The whole experiment will take approximately an hour and you will receive 10$ for your participation.

Participant's Statement:
"I have read the description of the research project and hereby agree to participate. I am aware that the results will be used for research purposes only, that my identity will remain confidential, and that I can withdraw at any time, if I so wish."

Name: ________________________________ Date: ______

Signature: ___________________________

ID #: Experimenter's name:
Receipt #: Room #:

This research is conducted under the supervision of McGill Professor C. Guastavino.
Contact mil.sis@mcgill.ca for more information.
McGill University, 3459 McTavish, Montreal, QC H3A 1B1.
Phone: (514) 398-1530. Fax: (514) 398-7193.

2. Appendix B

2.1. Formative evaluation. The following pages contain the forms related to the formative evaluation:

- The informed consent form
- The experimenter's protocol used for the experiment
- The demonstration guidelines used during training
- The interview guidelines and questions
- A summary of the reactions of each participant

Informed consent form - McGill University

WHY ARE WE DOING THIS RESEARCH? Our goal is to establish scientific knowledge about using audio cues for information navigation. It has nothing to do with your personality or motivations or intelligence.

PRIVACY. All the responses that we collect will be aggregated for analysis and so your own responses will be averaged with those of many other people. The session will be recorded for transcription and analysis purposes only. We know that you value your privacy. You will not be identified as an individual in any scientific report of this research, and your name will not be linked to your responses in this study.

DISCUSSION OF RESEARCH IDEAS. We cannot discuss our ideas with you before the study takes place, but we will be happy to talk with you about our hypotheses and theories afterwards.

WHAT WILL HAPPEN DURING THE EXPERIMENT? You will be seated in a quiet room, hear sounds over speakers and carry out information seeking tasks using both audio and visual cues. The experiment will not cause you discomfort. There are no known risks associated with the study. You will be free to discontinue your participation at any time without penalty. Following the study, you will be asked a few questions in an informal audio-recorded interview. Participation is voluntary and you are free to refuse to answer any question.

The whole experiment will take approximately an hour and you will receive 10$ for your participation.

Participant's Statement:
"I have read the description of the research project and hereby agree to participate. I am aware that the results will be used for research purposes only, that my identity will remain confidential, and that I can withdraw at any time, if I so wish."

Name: ______________________________ Date: ______

Signature: ___________________________

ID #: Experimenter's name:
Receipt #: Room #:

This research is conducted by Rafa Absar, PhD Candidate (PhD5) under the supervision of McGill Professor C. Guastavino of the School of Information Studies, McGill University
Contact mil.sis@mcgill.ca for more information.
McGill University, 3661 Peel, Montreal, QC H3A 1X1.
Phone: (514) 398-1530. Fax: (514) 398-7193.

If you have any questions or concerns regarding your rights or welfare as a participant in this research study please contact the McGill Research Ethics Officer at 514-398-6831 or lynda.mcneil@mcgill.ca

Experimenter's Protocol for Formative Study

Hello, my name is Rafa. Welcome to my information systems study. To begin, I am going to explain the procedure and you can ask me any questions you like. Then I will give you the opportunity to read the consent from and decide whether or not you wish to participate.

So, what you will be asked to do in the study is to use an information system to complete certain tasks. The whole study should not take more than one hour.

If you have any questions now, I would be happy to answer them. I will also be here
throughout the study to answer any questions you might have.

At this time, I would ask you to read the consent form and decide whether you would like to continue your participation. Of course, you may withdraw from the study at any time without penalty. If you decide to participate, please sign and date the consent form, and return it to me. You are welcome to take an unsigned copy for your records. You will receive 10$ for your participation.

HAND IN CONSENT FORM

Now we are ready to start. I will first go through a demo of the information system to give you an idea of how it works. You can ask me any questions you like whenever you have any confusion. You will also have a printout of the keyboard shortcuts for your reference.

Before we begin, you will go through a few practice tasks in a training set. You will be asked to search or browse through a library collection to explore topics or find relevant books. You will carry out the tasks on the information system showing on the left-hand monitor, and the tasks will be described in the right-hand monitor, where you can enter your answers. You will hear sounds related to the information system on speakers.

This is an informal study aimed at collecting feedback about your impression on the system and its usability. Please feel free to discuss any thing during or after the study. We would like to use a "Think-aloud" procedure here to gain maximum feedback. This means you should describe everything you do or every step you take while carrying out the tasks.

The right or wrong answers do not indicate any intelligence or ability; the study is simply to find out what sort of information system best benefits users.

You can take a break whenever you like.

At the end of the experiment, you will be asked a few informal questions about your experience.

Here is a copy of the keyboard instructions.
HAND IN INSTRUCTIONS

Before we begin, do you have any questions? We will first go through a demo of the system.

START DEMO. REFER TO DEMO GUIDELINES.

Would you like a few minutes to play around with the system to get the hang of it? After that, we'll try out some practice trials before we move on to the actual tasks. Remember to describe your activities as you go on. (Give 5 minutes to play around with the system).

If you're ready to go on to the training questions, we'll do that now.

TRAINING SET FOR SE3D (with sound)

RUN TRAINING SET .
http://mil.mcgill.ca/absar/webforms/trainingset_1/question_1.html

Would you like to move on to the main experiment now?

RUN TASKSET.
http://mil.mcgill.ca/absar/webforms/taskset_1/start_form.html

You have completed the tasks. I would like to ask you a few questions now.

OPEN INTERVIEW GUIDELINES

The study is now over. Do you have any questions?
Thank you very much for your participation and time.

Here is your money and your receipt.

WRITE DOWN RECEIPT NUMBER ON THE CONSENT FORM.

Formative Study – System Demo guidelines

- This is the Subject Explorer 3D (or SE-3D) with auditory feedback
- It is an information system where you can find books like in a library or bookstore
- You can move around by using the keyboard arrows, in each direction.
- If you point the mouse at a subject, the color of the subject changes to show you more clearly which subject you are on.
- Suppose I click on "Physics"
- Now what you see is called a "subject map". It shows you all the other subjects that are under Physics, shown in the circles.
    - Notice the book list on the left. It shows books specifically under Physics. The number at the top gives you the total number of items under Physics, while the smaller number gives the number of books only under Physics and not the other subjects under it.
    - Notice how the circles are of different sizes. The bigger the circle the more books the subject contains.
- Now let's click on the "Mathematical Physics" circle.
- I've been flown to the Mathematics physics map and book list has changed.
- Now suppose I want to zoom out back to the view of the tree. Press "Z" on keyboard, or the magnifying glass button on the toolbar.
- You can also press "F" or "G" to twist left or right. "T' or "B" to go up or down, as labeled on the keyboard.
- Notice that some maps are darker than the rest, this means that there are no other maps underneath.
- Now let's click on this leaf here, which is a darker color than the rest. As you may remember, this indicates there are no more subjects under it. The thud sound also indicates this.
- Now lets use the search box by entering keywords to search for.
    - For example, let's search for "Algorithms".
    - Only subjects with books about algorithms are shown (darker red means more results)
    - The book list shows a list of books about algorithms
    - The books in the book list also have links to subject maps
    - Clicking on "Genetic Algorithms" takes me there.
        - Orange subject circles contain books matching the keyword "algorithms"
- Let's zoom out with Z again.

  - Let's click on "evolutionary programming" on the book list again.
  - The book list now shows books with algorithms in evolutionary programming.
- We can also press Reset. This takes out all the orange circles, and the book list disappears, going back to a refreshed state.
- I want to look at it from the top again, so I click on the top "Schulich Library". Then I zoom out with "Z". Any place on you've clicked on before is denoted using the green box.
- Now let's click on "Physical Sciences".
- The sound of the forest indicates the relative depth of the subject in the tree. The more the sound of birds, and the less the sound of leaves, the deeper you are in the hierarchy. For example, let's hear the sound of "Physical Sciences" again. Now let's hear the sound at a deeper level, say "Elasticity". And if we go deeper, say to "Semiconductors", the birds get even louder.
- Now let's right click on this subject. You hear the sound of a page turning. This indicates the relative amount of information or items in that subject. There are 7 levels, of 1 page to 6 pages turning, and sound of a book being flipped quickly, meaning the maximum amount of items. You can opt to right-click and hear this sound when you're zoomed in to a subject map, or even when zoomed out. Aside from the sound, the book list will also update to show you the books found in that subject.
- Let's zoom out and right click on Science. Now let's click on Dynamics. This gives you a sense of how many items are in each of these subjects, relative to each other.
- That's about it. Do you have any questions before we start the training tasks?

Interview Guidelines for Formative Study

What's your impression on the system?

What did you like about the system?
What did you dislike about the system?

Did you find it easy or difficult to use?

Please give me specific examples?

Did you encounter any problems while trying to do the tasks?

Did you find the different sounds helpful while carrying out the tasks? Did any of them help make the tasks easier?

Did you use any of them while doing the tasks?

Did you find any of the sounds annoying?
What did you think of the

- click sound
- zoom sound
- depth sound
- info density sound
- last subject in branch sound

What did you think this sounds indicates?
How well does it convey this meaning, in your opinion?
How easy was it to learn or remember?

Do you have any comments to make about any of them?

Is there any sound you would have liked to leave out? Why?

Do you have any comments about any of the tasks? Or the study in general?

Overall, do you think you would prefer using this information system with sounds or without sounds? Please explain why.

Summary of Participant Feedback for Formative Evaluation

Participant P01

- Session took about 50 minutes
- Said she liked using the system with sounds and would prefer them to be there.
- She did not always use them to do the tasks however.
- It was hard to remember the sounds when the tasks involved 4 subjects, might be easier if reduced to 2 or 3 subjects at most.
- Thought all the sounds worked well in their contexts.
- Was wondering if the real study takes a longer time, the zoom sound might be distracting if heard repeatedly, so maybe it should be toned down.

Participant P02

- Session took about 45 mins
- Said likes the option of having sounds, would also like the option of being able to turn them off.
- Thought all the sounds were appropriate, conveyed meaning well
- Said she particularly liked the depth sound, it matched the visuals, as well as was a useful sound.
- The information density sound was harder to distinguish
- The system is easy to use, but needs getting used to. Maybe have a slightly longer training and be given some time to play around with the system before the real tasks.
- Thought that if one was using the system for a longer time, the depth sound might get repetitive and annoying.
- Got stuck in Q2 and Q6. In q2, could not see the link clearly. In q6, could not find the right keywords to complete the search successfully, had to be given hints at the end.
- Suggested explaining the sounds earlier in the demo, while explaining the visuals, rather than at the end.

Participant P03

- Done in 50 mins
- Really liked the IV system: loved how you could see the structure and links between subjects, etc.

- Said he's more of a visual person, noticed the sounds less than the graphics
- But commented that he definitely liked using the system with the sounds in it.
- Even though he didn't use them much for the task, he thought that with time, he would probably use them more.
- He liked the information density and depth sounds.
- Didn't notice the thud sound that much since it didn't come up in the tasks much.
- Appreciated having the time to play around and get the hang of it.
- He said the system is easy to use, maybe playing around more would get him more used to it.

Participant P04

- First participant to use system with Dynaudio speakers with the MOTU soundcard, instead of the general Yamaha speakers.
- Volume of main out set to -28dB.
- Done in 40 mins
- Was not positive about the IV system from the beginning:
  - Objected on how you could not see all the labels because of obscuring
  - Said she would prefer a list-based system for the subjects
  - Would prefer the system being faster
- Disregarded the sounds, didn't pay much attention to them
- Said she's a science person who needs numbers and facts: the sounds are too qualitative, she needs something quantitative, so she did not use the sounds at all.
- When asked if she would prefer the system with sounds or no sound, she said it made no difference to her either way.
- At the end she said it's a great idea, maybe if she was more used to it, she'd use it more, but she prefers the traditional library system better.

Participant P05

- Positive about system and sounds
- Said didn't use sounds as much but thought would do so with more time.
- Suggested no changes.

3. Appendix C

3.1. Controlled evaluation. The following pages contain the forms related to the controlled evaluation:

- The informed consent form
- The experimenter's protocol used for the experiment
- The demonstration guidelines used during training
- Screenshots of the web post-test questionnaires, in the order they were presented
- A few sample screenshots of the HTML task forms

Informed consent form - McGill University

WHY ARE WE DOING THIS RESEARCH? Our goal is to establish scientific knowledge about information navigation. It has nothing to do with your personality or motivations or intelligence.

PRIVACY. All the responses that we collect will be aggregated for analysis and so your own responses will be averaged with those of many other people. We know that you value your privacy. You will not be identified as an individual in any scientific report of this research, and your name will not be linked to your responses in this study.

DISCUSSION OF RESEARCH IDEAS. We cannot discuss our ideas with you before the study takes place, but we will be happy to talk with you about our hypotheses and theories afterwards.

WHAT WILL HAPPEN DURING THE EXPERIMENT? You will be seated in a quiet room, sometimes hear sounds over headphones and carry out information seeking tasks using both audio and/or visual cues. The experiment will not cause you discomfort. There are no known risks associated with the study. You will be free to discontinue your participation at any time without penalty. Following the study, you will be asked to complete some questionnaires. Participation is voluntary and you are free to refuse to answer any question.

The whole experiment will take approximately an hour and a half, and you will receive 15$ for your participation.

Participant's Statement:
"I have read the description of the research project and hereby agree to participate. I am aware that the results will be used for research purposes only, that my identity will remain confidential, and that I can withdraw at any time, if I so wish."

Name: ________________________________ Date: ______

Signature: ___________________________

ID #: Experimenter's name:
Receipt #: Room #:

This research is conducted Rafa Absar, PhD Candidate (PhD5) under the supervision of McGill Professor C. Guastavino of the School of Information Studies, McGill University
Contact mil.sis@mcgill.ca for more information.
McGill University, 3661 Peel, Montreal, QC H3A 1X1.
Phone: (514) 398-1530. Fax: (514) 398-7193.

If you have any questions or concerns regarding your rights or welfare as a participant in this research study please contact the McGill Research Ethics Officer at 514-398-6831 or lynda.mcneil@mcgill.ca

Experimenter's Protocol for Comparative Evaluation

Hello, my name is Rafa. Welcome to our information systems navigation study. To begin, I am going to explain the procedure and you can ask me any questions you like. Then I will give you the opportunity to read the consent form and decide whether or not you wish to participate.

So, what you will be asked to do in the study is to use an information system to complete certain tasks. The whole study should not take more than 1.5 hours. If you have any questions now, I would be happy to answer them. I will also be here throughout the study to answer any questions you might have.

At this time, I would ask you to read the consent form and decide whether you would like to continue your participation. Of course, you may withdraw from the study at any time without penalty. If you decide to participate, please sign and date the consent form, and return it to me. You are welcome to take an unsigned copy for your records. You will receive 15$ for your participation.

HAND IN CONSENT FORM

Now we are ready to start. I will first go through a demo of the information system to give you an idea of how it works. You can ask me any questions you like whenever you have any confusion. You will also have a printout of the keyboard shortcuts for your reference.

Before we begin, you will go through a few practice tasks in a training set. You will be asked to search or browse through a library collection to explore topics or find relevant books. You will carry out the tasks on the information system showing on the left-hand monitor, and the tasks will be described in the right-hand monitor, where you can enter your answers.

The right or wrong answers do not indicate any intelligence or ability; the study is simply to find out what sort of information system best benefits users.

You can take a break whenever you like.

At the end of the experiment, you will be asked a few informal questions about your experience.

Here is a copy of the keyboard shortcuts. These are labeled on the keyboard as well.

HAND IN INSTRUCTIONS

Before we begin, do you have any questions? We will first go through a demo of the system.
START DEMO. REFER TO DEMO GUIDELINES. (If without sound session, turn off speakers/headphones. If with sound session, turn on speakers/headphones. Check volume setting.)

I will now give you a few minutes to play around with the system to get the hang of it. After that, we'll try out some practice trials before we move on to the actual tasks. Remember to describe your activities as you go on. (Give 5 minutes to play around with the system).

If you're ready to go on to the training questions, we'll do that now.

TRAINING SET FOR SE3D (with or without sound)

RUN TRAINING SET .
http://mil.mcgill.ca/absar/webforms/trainingset_1/question_1.html or
http://mil.mcgill.ca/absar/webforms/trainingset_2/question_1.html

Would you like to move on to the main experiment now?

RUN FIRST TASKSET (with or without sound).
http://mil.mcgill.ca/absar/webforms/taskset_1/start_form.html or
http://mil.mcgill.ca/absar/webforms/taskset_2/start_form.html

Please complete this questionnaire now.
OPEN QUESTIONNAIRE
If session was with sound, open
http://mil.mcgill.ca/absar/webforms/questionnaire_withsound.html

We will now move on to the next session.
(If this session is **with sound**, go through second part of DEMO).
(If this session is **with no sound**, say: This time the information system will be the same as before, but you will have no audio cues, only the visual system).

RUN SECOND TASKSET (with or without sound).

Please complete the questionnaires now.
OPEN QUESTIONNAIRES:
According to system orders (refer to excel sheet)
http://mil.mcgill.ca/absar/webforms/questionnaire_withsound.html

http://mil.mcgill.ca/absar/webforms/preference_form.html
http://mil.mcgill.ca/absar/webforms/combined_withsound.html
http://mil.mcgill.ca/absar/webforms/combined_nosound.html
http://mil.mcgill.ca/absar/webforms/posttest_form.html

The experiment is now over. Do you have any questions?
Thank you very much for your participation and time.

Here is your money and your receipt.

WRITE DOWN RECEIPT NUMBER ON THE CONSENT FORM.

Demo Guidelines for Comparative Study

- This is the Subject Explorer 3D (or SE-3D). It is an information system where you can find books like in a library or bookstore
- You can move around by using the keyboard arrows, in each direction.
- You can also press “F” or “G” to twist left or right. “T’ or “B” to go up or down, as labeled on the keyboard.
    - Each of these keyboard shortcuts have equivalent icons on screen which can be clicked with the mouse.
- If you point the mouse at a subject, the color of the subject changes to show you more clearly which subject you are on.
- Suppose I click on “Physics”.
- Now what you see is called a “subject map”. It shows you all the other subjects that are under Physics, shown in the circles.
    - Notice the book list on the left. It shows books specifically under Physics. The number at the top gives you the total number of items under Physics, while the smaller number gives the number of books only under Physics and not the other subjects under it.
    - Notice how the circles are of different sizes. The bigger the circle the more books the subject contains.
- Now let’s click on the “Mathematical Physics” circle. I’ve been flown to the Mathematics physics map and book list has changed.
    - The depth level of the subject is written on the subject map.
- You can also go back one level by clicking anywhere outside the subject map. This will take you to the subject on top of the previous one.
- Now suppose I want to zoom out back to the view of the tree. Press “Z” on keyboard, or the magnifying glass button on the toolbar.
- Notice that some maps are darker than the rest, this means that there are no other maps underneath.
- Any place which you’ve clicked on before is denoted using the green box.

- Now lets use the search box by entering keywords to search for.
    - For example, let’s search for “Algorithms”.
    - Only subjects with books about algorithms are shown (darker red means more results)
    - The book list shows a list of books about algorithms
    - The books in the book list also have links to subject maps
    - Clicking on “Genetic Algorithms” takes me there.
        - Orange subject circles contain books matching the keyword “algorithms”
        - The book list now shows books with the term “algorithms” in genetic algorithms.
- Let’s zoom out with Z again.
    - You are now seeing the tree pruned to what you had searched for and where you have been.
- We can also press Reset. This takes out all the orange circles, and the book list disappears, going back to a refreshed state.

- Please put on the headphones now. Let's talk about the audio cues now.
- To illustrate, let’s click on “Physical Sciences”.
    - You first hear a click, then a zoom sound and then a forest-like sound.
- The sound of the forest indicates the relative depth of the subject in the tree. The more the sound of birds, and the less the sound of leaves, the deeper you are in the hierarchy. For example, let’s hear the sound of “Physical Sciences” again. Now let’s hear the sound at a deeper level, say “Elasticity”. And if we go deeper, say to “Semiconductors”, the birds get even louder.
- Now let’s right click on this subject. You hear the sound of a page turning. This indicates the relative amount of information or items in that subject. There are 7 levels, of 1 page to 6 pages turning, and sound of a book being flipped quickly, meaning the maximum amount of items. You can opt to right-click and hear this

sound when you're zoomed in to a subject map, or even when zoomed out. Aside from the sound, the book list will also update to show you the books found in that subject.

- Let's zoom out and right click on Science. Now let's click on Dynamics. This gives you a sense of how many items are in each of these subjects, relative to each other.
- Now let's click on this leaf here, which is a darker color than the rest which indicates there are no more subjects under it. The thud sound also indicates this.

Questionnaire for SE-3D with sound

Participant ID:

Please indicate how you feel about the following statements on the sounds in the system.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I found the different sounds and their functions easy to use.	○	○	○	○	○
I learned the different sounds and their functions quickly.	○	○	○	○	○
I easily remembered the different sounds and their functions.	○	○	○	○	○
I quickly became skillful at using the different sounds.	○	○	○	○	○
Overall, I found the different sounds useful while carrying out the tasks.	○	○	○	○	○

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I am satisfied with "the click" sound in the system.	○	○	○	○	○
Please explain why.					

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I am satisfied with the "zooming in" sound in the system.	○	○	○	○	○
Please explain why.					

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I am satisfied with "the relative depth level of a subject" sound in the system.	○	○	○	○	○

Please explain why.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I am satisfied with "the last subject in a branch" sound in the system.	○	○	○	○	○

Please explain why.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I am satisfied with "the amount of information in a subject" sound in the system.	○	○	○	○	○

Please explain why.

Please add any comments or observations you may have had.

Submit

Preference Questionnaire

Participant ID:

Which information system did you prefer using?	Visual-only system	Audio-visual system

Please explain why.

Submit

Questionnaire for both systems

Participant ID:

Please indicate how you feel about the following statements for each of the 2 systems you used.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
This information system was aesthetically appealing.	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
This information system appealed to my senses.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
This information system is attractive.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
I felt frustrated while using this information system.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
I felt annoyed while using this information system.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○

I felt discouraged while using this information system.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
Using this system was taxing.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
I would continue to use this information system out of curiosity.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
The content of this information system incited my curiosity.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
The interaction with the system is clear and understandable.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○
Interaction with the system does not require a lot of mental effort.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No					

	sound					
	With sounds					
I found the information system easy to use.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound					
	With sounds					
I found it easy to get the system to do what I want it to do.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound					
	With sounds					
This system is fun to use.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound					
	With sounds					
Using this information system was worthwhile.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound					
	With sounds					
I would recommend that others use this information system.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound					
	With sounds					

Overall, I am satisfied with this system. Please explain why below.	System	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	No sound	○	○	○	○	○
	With sounds	○	○	○	○	○

Please add any comments or observations you may have had while using each of the information system.

Submit

Post-test Questionnaire

Participant ID:

1. Have you ever used the current McGill Online Library Catalogue?	Yes	No		
	○	○		
2. If yes, how often do you use it? If no, proceed to question 4.	Once to a couple of times a year.	Once to a couple of times a term.	At least once a month.	At least once a week.
	○	○	○	○
3. Given a choice between the three, which would you prefer using?	The current online library catalogue	The information system used in this study (without sounds)	The information system used in this study (with sounds)	
	○	○	○	

4. Gender:

5. Age:

6. Academic Major:

7. Have you ever taken music lessons e.g. instrumental, vocal or music theory, in addition to the regular music curriculum in school?

Yes ○ No ○

**If YES, please complete below; if NO, proceed to #8

What type? # of years of training:

What type? # of years of training:

What type? # of years of training:

What type? # of years of training:

What type? # of years of training:

8. Even if not trained, do you play any musical instruments? If yes, which ones?
Please mention the number of years you have been playing.

9. Have you had any ear training, sound engineering training, music technology or

any other relevant training in sound? Please specify the type of training and the number of years trained:

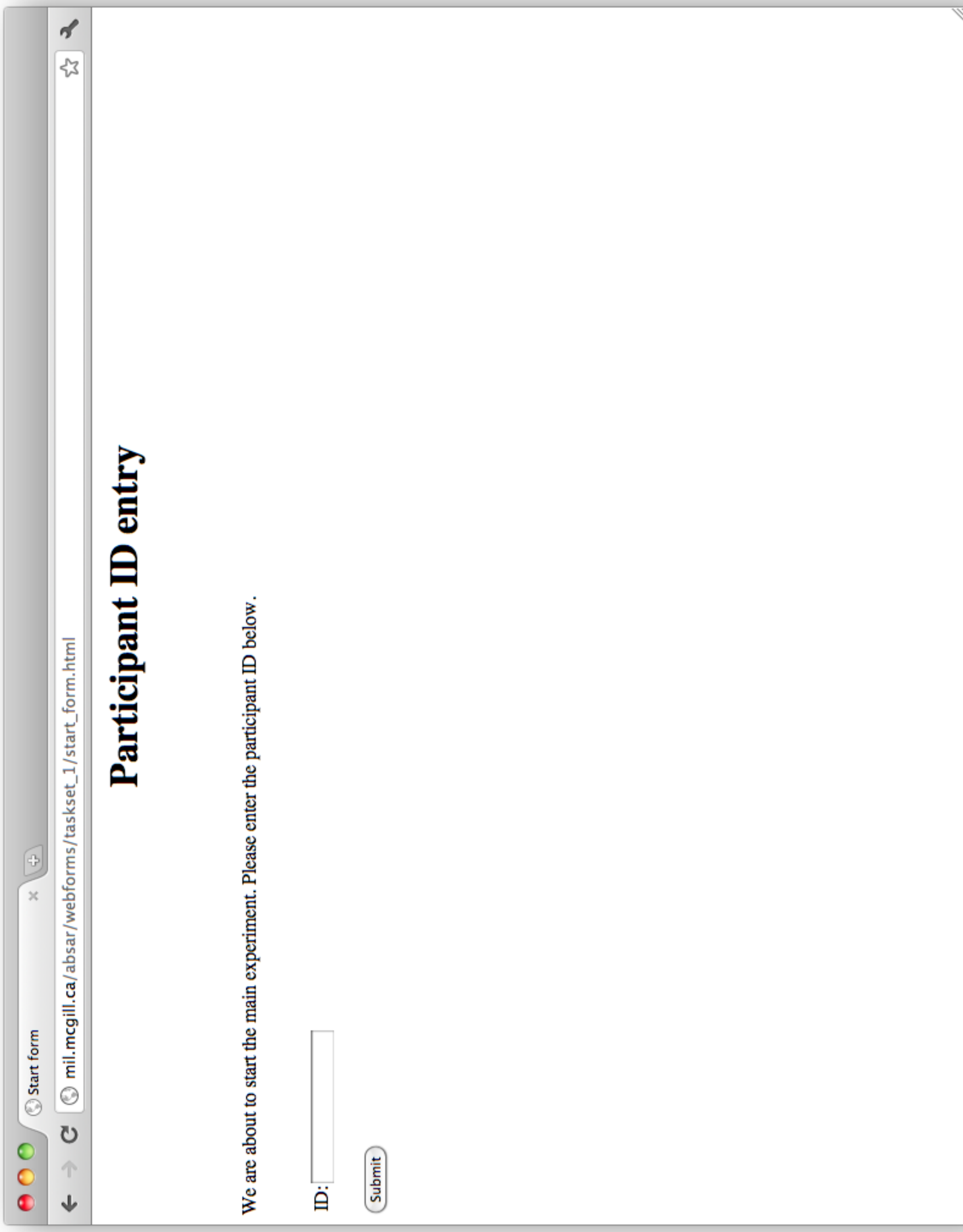
10. Do you use computer interface sounds in your daily life or do you prefer to turn such sounds off? Please specify why.

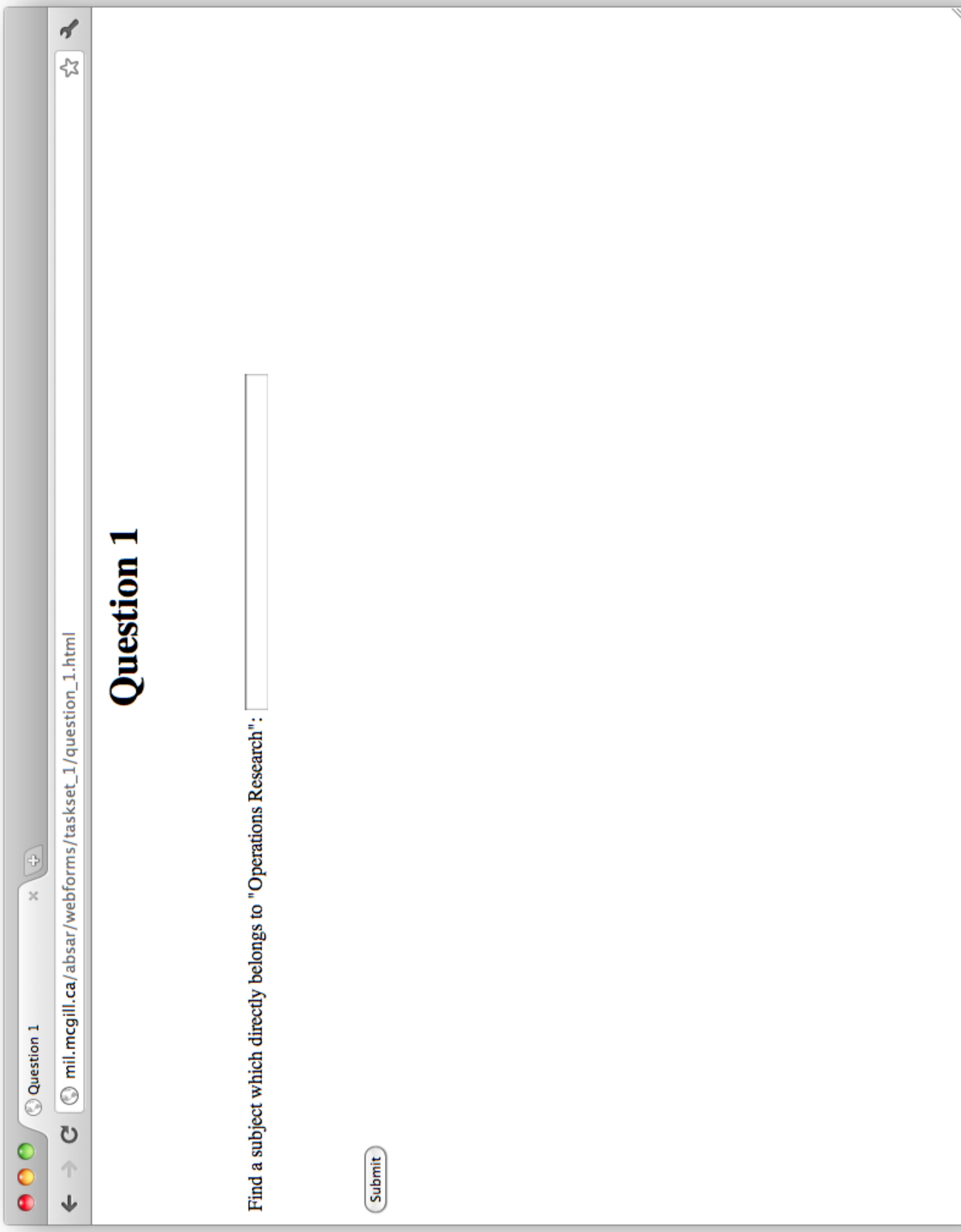
11. Have you ever used any visualizations of any sort? e.g. visual search engines, data or social network visualization tools, etc.? If yes, please specify which ones and comment about your experience with them.

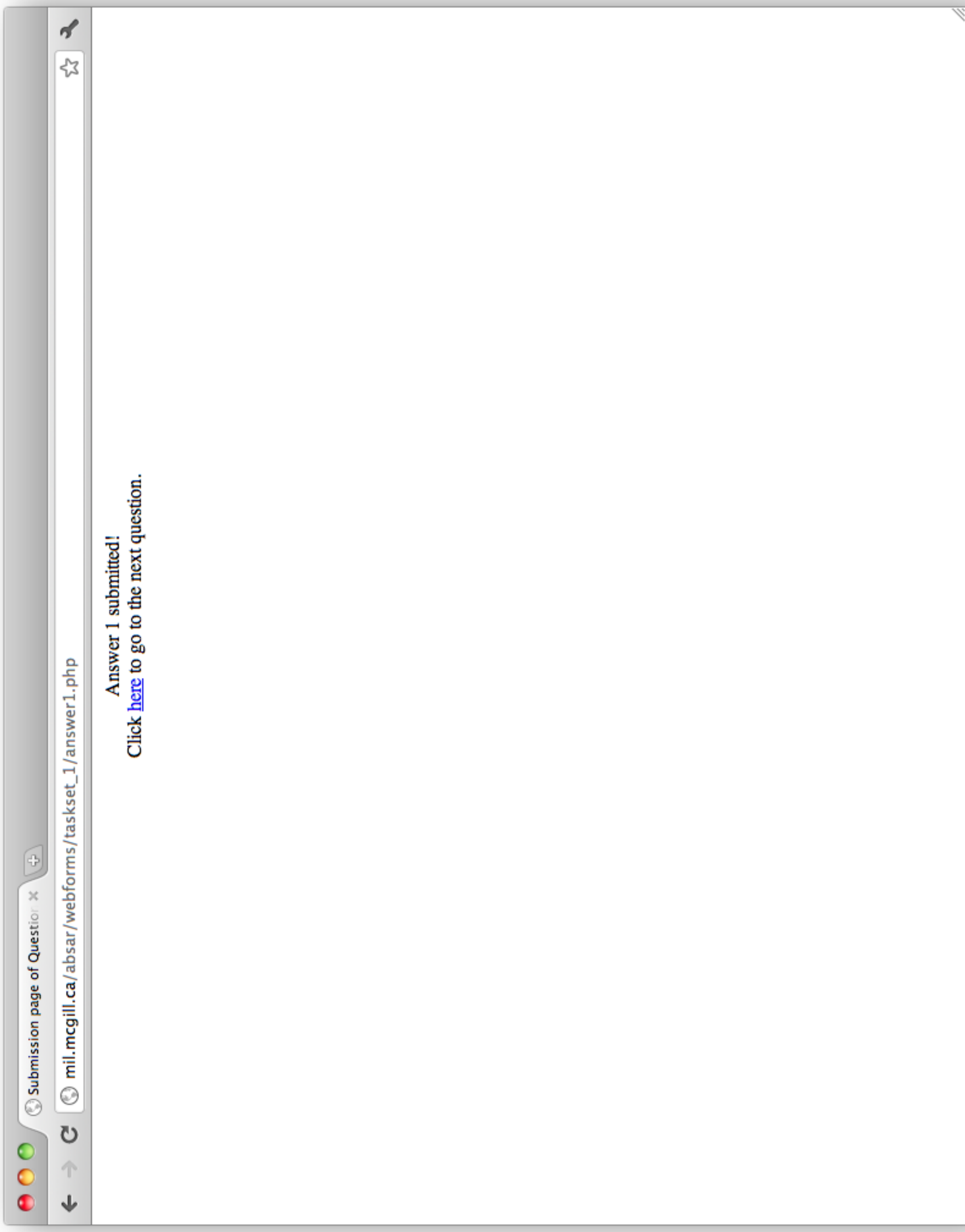
12. Do you ever play video games (those that include auditory feedback)?	Yes	No			
	○	○			
If yes, which ones do you play? If no, go on to question 13.					
How often do you play?	Once to a couple of times a year	Once to a couple of times a term	At least once a month	At least once a week	
	○	○	○	○	
13. In general, I am sensitive to noise.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	○	○	○	○	○
14. Do you have any hearing problems?	Yes	No			
	○	○			
If yes, please specify.					

15. We would be grateful if you would add any comments or observations you may have had during the experiment. Your insight can be very useful to us!

Submit


Start form
mil.mcgill.ca/absar/webforms/taskset_1/start_form.html
Participant ID entry
We are about to start the main experiment. Please enter the participant ID below.
ID:
Submit


Question 1
mil.mcgill.ca/absar/webforms/taskset_1/question_1.html
Question 1
Find a subject which directly belongs to "Operations Research":
Submit


Submission page of Questio
mil.mcgill.ca/absar/webforms/taskset_1/answer1.php
Answer 1 submitted!
Click here to go to the next question.

4. Appendix D

4.1. Measurement Scales for User Experience. The following pages contain the measurement scales for user experience referred to in Chapter 3:

(i) Davis' items on Perceived usefulness and ease of use (Davis, 1989)
(ii) Doll and Torkzadeh's EUCS instrument (Doll & Torkzadeh, 1988)
(iii) Wixom and Todd's survey items (Wixom & Todd, 2005)
(iv) Lund's USE questionnaire (Lund, 2001)
(v) User Engagement Scale (UES) (O'Brien & Toms, 2010a)

Scale items adopted into the final user experience measurement scales are denoted with a *.

4.1.1. *Davis' measurement scales.* Each scale was rated using a 7-point Likert scale ranging from "Extremely likely" to "Extremely unlikely". "X" signifies the system being evaluated.

Perceived Usefulness:

- Using X in my job would enable me to accomplish tasks more quickly.
- Using X would improve my job performance.
- Using X in my job would increase my productivity.
- Using X would enhance my effectiveness on the job.
- Using X would make it easier to do my job.
- I would find X useful in my job.*

Perceived Ease of Use:

- Learning to operate X would be easy for me.
- I would find it easy to get X to do what I want it to do.*
- My interaction with X would be clear and understandable.*
- I would find X to be flexible to interact with.
- It would be easy for me to become skillful at using X.*

- I would find X easy to use.*

4.1.2. *Doll and Torkzadeh's EUCS instrument.* The "G" items correspond to Global scales, "C" to Content, "A" to Accuracy, "F" to Format, "E" to Ease of Use, and "T" to Timeliness.

- G1. Is the system successful?
- G2. Are you satisfied with the system?*
- C1. Does the system provide the precise information you need?
- C2. Does the information content meet your needs?
- C3. Does the system provide reports that seem to be just about exactly what you need?
- C4. Does the system provide sufficient information?
- A1. Is the system accurate?
- A2. Are you satisfied with the accuracy of the system?
- F1. Do you think the output is presented in a useful format?
- F2. Is the information clear?
- E1. Is the system user friendly?
- E2. Is the system easy to use?*
- T1. Do you get the information you need in time?
- T2. Does the system provide up-to-date information?

4.1.3. *Wixom and Todd's integrated model survey items.* The survey items were rated using a 7-point Likert scale.

- Completeness
  - X provides me with a complete set of information.
  - X produces comprehensive information.
  - X provides me with all the information I need.

- Format
  - The information provided by X is well-formatted.
  - The information provided by X is well laid out.
  - The information provided by X is clearly presented on the screen.*
- Accuracy
  - X produces correct information.
  - There are few errors in the information I obtain from X.
  - The information provided by X is accurate.
- Currency
  - X provides me with the most recent information.
  - X produces the most current information.
  - The information from X is always up to date.
- Information quality
  - Overall, I would give the information from X high marks.
  - Overall, I would give the information provided by X a high rating in terms of quality.
- Reliability
  - X operates reliably.
  - X performs reliably.
- Accessibility
  - X allows information to be readily accessible to me.
  - X makes information very accessible.
  - X makes information easy to access.
- Flexibility
  - X can be adapted to meet a variety of needs.
  - X can flexibly adjust to new demands or conditions.
  - is versatile in addressing needs as they arise.

- Integration
  - X effectively integrates data from different areas of the company.
  - X pulls together information that used to come from different places in the company.
  - X effectively combines data from different areas of the company.
- Timeliness
  - It takes too long for X to respond to my requests.
  - X provides information in a timely fashion.
  - X returns answers to my requests quickly.
- System quality
  - In terms of system quality, I would rate X highly.
  - Overall, X is of high quality.
  - Overall, I would give the quality of X a high rating.
- Information satisfaction
  - Overall, the information I get from X is very satisfying.
  - I am very satisfied with the information I receive from X.
- System satisfaction
  - All things considered, I am very satisfied with X.*
  - Overall, my interaction with X is very satisfying.
- Attitude
  - Using X is (not enjoyable/very enjoyable).
  - Overall, using X is a (unpleasant/pleasant) experience.
  - My attitude toward using X is (very unfavorable/favorable).
- Intention
  - I intend to use X as a routine part of my job over the next year.
  - I intend to use X at every opportunity over the next year.
  - I plan to increase my use of X over the next year.

- Ease of use
  - X is easy to use.*
  - It is easy to get X to do what I want it to do.*
  - X is easy to operate.
- Usefulness
  - Using X improves my ability to make good decisions.
  - X allows me to get my work done more quickly.
  - Using X enhances my effectiveness on the job.

4.1.4. *Lund's usability questionnaire.* Each scale is rated using a 7-point Likert scale ranging from "Strongly disagree" to "Strongly agree".

- Usefulness:
  - It helps me be more effective.
  - It helps me be more productive.
  - It is useful.*
  - It gives me more control over the activities in my life.
  - It makes the things I want to accomplish easier to get done.
  - It saves me time when I use it.
  - It meets my needs.
  - It does everything I would expect it to do.
- Ease of Use:
  - It is easy to use.*
  - It is simple to use.
  - It is user friendly.
  - It requires the fewest steps possible to accomplish what I want to do with it.
  - It is flexible.

  - Using it is effortless.
  - I can use it without written instructions.
  - I don't notice any inconsistencies as I use it.
  - Both occasional and regular users would like it.
  - I can recover from mistakes quickly and easily.
  - I can use it successfully every time.
- Ease of Learning:
  - I learned to use it quickly.*
  - I easily remember how to use it.*
  - It is easy to learn to use it.*
  - I quickly became skillful with it.*
- Satisfaction:
  - I am satisfied with it.*
  - I would recommend it to a friend.*
  - It is fun to use.*
  - It works the way I want it to work.
  - It is wonderful.
  - I feel I need to have it.
  - It is pleasant to use.

4.1.5. *User Engagement Scale (UES).* Each scale was rated using a 5-point Likert scale ranging from "Strongly agree" to "Strongly disagree", with the wording of the scale items specific to the online shopping experience which was evaluated (O'Brien & Toms, 2010a).

1. I lost myself in this shopping experience.
2. I was so involved in my shopping task that I lost track of time.

3. I blocked out things around me when I was shopping on this website.
4. When I was shopping, I lost track of the world around me.
5. The time I spent shopping just slipped away.
6. I was absorbed in my shopping task.
7. During this shopping experience I let myself go.
8. I was really drawn into my shopping task.
9. I felt involved in this shopping task.
10. This shopping experience was fun.*
11. I continued to shop on this website out of curiosity.*
12. The content of the shopping website incited my curiosity.*
13. I felt interested in my shopping task.
14. Shopping on this website was worthwhile.*
15. I consider my shopping experience a success.
16. This shopping experience did not work out the way I had planned.
17. My shopping experience was rewarding.
18. I would recommend shopping on this website to my friends and family.*
19. This shopping website is attractive.*
20. This shopping website was aesthetically appealing.*
21. I liked the graphics and images used on this shopping website.
22. This shopping website appealed to my visual senses.*
23. The screen layout of this shopping website was visually pleasing.
24. I felt frustrated while visiting this shopping website.*

25. I found this shopping website confusing to use.*
26. I felt annoyed while visiting this shopping website.*
27. I felt discouraged while shopping on this website.*
28. Using this shopping website was mentally taxing.*
29. This shopping experience was demanding.
30. I felt in control of my shopping experience.
31. I could not do some of the things I needed to do on this shopping website.

5. Appendix E

5.1. Controlled evaluation data. The tables below display the descriptive statistics for the two dependent variables (Time and Accuracy) for each of the six tasks and for each of the two systems (audio-visual and visual-only) respectively.

TABLE E.1. Task Means, Standard deviations, lower and upper bounds (95% Confidence Interval)

Dependent Variable	**Task**	**Mean**	**Std. Dev**	**Lower Bound**	**Upper Bound**
Time	1	24.396	7.968	8.707	40.084
	2	119.253	7.968	103.565	134.941
	3	42.088	7.968	26.400	57.777
	4	40.765	7.968	25.077	56.453
	5	52.342	7.968	36.653	68.030
	6	135.725	8.057	119.861	151.588
Accuracy	1	97.917	3.897	90.243	105.590
	2	52.083	3.897	44.410	59.757
	3	97.917	3.897	90.243	105.590
	4	97.917	3.897	90.243	105.590
	5	100.000	3.897	92.326	107.674
	6	81.156	3.941	73.397	88.916

TABLE E.2. System Means, Standard deviations, lower and upper bounds (95% Confidence Interval)

Dependent Variable	**System**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
Time	Audio-visual	69.854	4.600	60.797	78.912
	Visual-only	68.335	4.617	59.244	77.427
Accuracy	Audio-visual	87.847	2.250	83.417	92.278
	Visual-only	87.816	2.259	83.369	92.263

5.2. Qualitative analysis coding scheme. The coding scheme for the qualitative data is given in the table below, with all the phrases, codes and comments grouped for each emerging concept described in Table 6.7 and 6.8.

Table E.3: Coding scheme for free-format comments

Concept	Phrases or partial comments
Enhanced experience	added to the experience significantly; appreciated their existence; enhanced the experience of browsing; a great experience to use the two together; enjoyed using the system with sounds; did contribute to my preference; addition of sound to this system enhances this system; overall experience rating excellent with sounds; a worthwhile experience with audio; definitely added to the experience; added something to the experience of the system;
Quicker with sounds	quicker answers; did make the tasks go more quickly; helped to orient me within the tree more quickly; helped me in answering faster; helpful when trying to quickly determine the depth of a category; auditory cues were helpful and efficient; info seemed more quickly available;
Engagement	engaging; fun; appealing; interesting; stimulating; fascinating; cool; new; neat; fresh; more animation; interactive; impressive;
Reduced visual workload	reinforced the perception; provided additional feedback; easier to get this estimate from the sounds rather than visuals; provided more information; need to read less; useful in providing feedback when visuals did not; nice to get feedback; made it clear; didn't need to check visually;
Pleasantness	nice addition; pleasant to listen to; pleasantness of the interface;

Table E.3: Coding scheme for free-format comments

Concept	**Phrases or partial comments**
Usefulness	helpful; useful; convenient; helped me to understand; got an idea of; facilitated tasks; informative; helped orient;
Ease of use	easy to use; easier to get; user-friendly;
Immersiveness	more immersive; nicer sense of immersion; added to the immersiveness;
Prefer continuous sound	could get more out of the sounds possibly if the sounds were more continuous; I'm used to continuous sound or none at all;
Not critical to use	not critical to its usefulness; other ways to see the information; not necessary; irrelevant for search;
Trust visuals more	trust my visual system a lot more than my audio system; I more trust what I see then what I hear; I prefer and am used to visual-only with computers; how we grow up or accustomed to do; more tuned for it;
Limits usage	sounds limit its usage (user needs special equipment to use it, like speakers)."
Would like better view of previous hierarchical levels at a time	like to have a better view of where I am in the tree at any given level; wish it were easier to read the parents of the node I'm currently sitting on; will be much easier to search through if hierarchy also appears in the in screen window of books, like the directory used to be in windows explorer;
Need more camera control	the zoom out function zooms too much; would be nice for the zoom out option to be more gradual;

Table E.3: Coding scheme for free-format comments

Concept	**Phrases or partial comments**
System latency	wait time can get old real fast; system latency was distracting; make it run faster; latency was a problem;
Visuals confusing or obscured	sometimes the interface was too cluttered; confusing; too many branches on the screen; preventing a clear view; yellow icons for going to the parent were a bit misleading; wish there were a clearer rule for what words pop out;

Document Log:

Manuscript Version 0.2

Typeset by $\mathcal{A}\!\mathcal{M}\!\mathcal{S}$-LaTeX — 11 April 2012

RAFA ABSAR

McGILL UNIVERSITY, 3661 PEEL ST., MONTRÉAL (QUÉBEC) H3A 1S1, CANADA, *Tel.* : (514) 398-4204

E-mail address: rafa.absar@mail.mcgill.ca

Typeset by $\mathcal{A}\!\mathcal{M}\!\mathcal{S}$-LaTeX