Investigating Methods for the Selection and Evaluation of Everyday Sounds for Use As Auditory Icons

Eoin Joseph Brazil, M.Sc.

Department of Computer Science & Information Systems
University of Limerick
Limerick, Ireland

Supervisors: **Professor Liam Bannon** and **Mikael Fernström**

Submitted for the degree of **Doctor of Philosophy**
to the **University of Limerick**

Submitted to the University of Limerick, July, 2009.

Declaration

Author: Eoin Brazil
Title: Investigating Methods for the Selection and Evaluation of Everyday Sounds for Use As Auditory Icons

I hereby declare that this thesis is entirely my own work and does not contain material previously by any other author, except where due reference or acknowledgement has been made. Furthermore I declare that it has not been previously submitted for any other academic award or to any other university.

Signature:____________________________ **Date**:____________________

Abstract

Auditory Displays use sound rather than graphics to represent information. Such displays can in certain cases deliver more natural, creative and intuitive methods of communication between computers and people. An existing gap in this field is the issue of how to link the knowledge created in laboratory environments to the techniques developed in everyday practical world of software design.

This thesis provides a range of techniques to help designers in selecting *non speech* auditory sounds, in particular everyday sounds for use as Auditory Icons. These techniques provide a cohesive foundation that supports the systematic selection of everyday sounds. This work organises existing research and extends it to help support early stage empirically inspired design methods within the field of Auditory Display.

The research questions in this thesis explore the subjective realism responses of people to everyday sounds, how people perceive certain subjective properties of everyday sounds, the use of prior classification for improving sound identification accuracy, the subjective confusion of sounds both with and without classification, and what tacit criteria people use for attributing meaning to everyday sounds. A detailed discussion of the qualitative and quantitative results of these research questions led to the development of a practical empirically-based design framework for Auditory Displays. The methods and results used to explore these research questions form key parts of the foundations of this framework.

The presented studies found that current synthetic sound models of impact/ bouncing events are not as accurate as recorded sounds for providing quantifiable information in multiple dimensions. A new method of similarity scaling using multiple comparisons of the sound set was presented. Further studies found that presenting everyday sounds concurrently with prior classification provides approximately 7% better identification accuracy when compared to everyday sounds concurrently that are presented without any prior classification. The final studies explored the identification of everyday sounds combining causal uncertainty measures with the repertory grid method. This approach provides insight into sound identifications, providing metaphors and vocabulary in the participant's own language.

The results of the presented studies contribute to a better understanding in the field of Auditory Display. The methods and purposed framework provide designers with a foundation for an empirically-based design framework for Auditory Displays. This framework provides a contribution that will open areas for further research into Auditory Icons and in the wider field of Auditory Display.

Publications from or related to the research in this thesis

Materials, ideas, and figures from this dissertation have appeared previously in the following publications. After each reference, I note the chapters in which material is used.

Book Chapters

Brazil, E., Fernström, M. and Ottaviani, L. (2003*a*), *The Sounding Object*, Mondo Estremo, Firenze, Italy, chapter Psychoacoustic validation and cataloguing of sonic objects: 2D browsing, pp. 257–294. — Chapter 4.

Fernström, M. and Brazil, E. (in press MIT Press, due 2010), *Principles of Sonification and Auditory Display*, T. Hermann, A. Hunt, J. Neuhoff, eds., chapter Auditory Icons, pp. in press. — Chapter 2.

Journal Papers

Fernström, M., Brazil, E. and Bannon, L. (2005), 'HCI design and interactive sonification for fingers and ears', *IEEE Multimedia* **12**(2), 36–44. — Chapters 2 and 5.

Refereed Conference Publications

Brazil, E., Fernström, J. and Ottaviani, L. (2003*b*), A new experimental technique for gathering similarity ratings for sounds, *in* E. Brazil and B. Shinn-Cunningham, eds, 'International Conference on Auditory Display (ICAD-03)', pp. 238–242 — Chapter 4.

Ottaviani, L., Brazil, E. and Fernström, M. (2003), Psychoacoustic experiments for validating sound objects in a 2-d space using the sonic browser, *in* 'Proceedings of the XIV Colloquium on Musical Informatics (XIV CIM 2003)', Firenze, Italy, pp. 90–94. — Chapter 4.

Brazil, E. and Fernström, M. (2006), Investigating concurrent Auditory Icon recognition, *in* 'Proceedings of ICAD 2006 - The 12th International Conference on Auditory Display', Queen Mary, London, pp. 51–58. — Chapter 5.

Brazil, E. and Fernström, M. (2007), Investigating ambient auditory information systems, *in* G. P. Scavone, ed., 'International Conference on Auditory Display (ICAD-07)', pp. 326–333. — Chapter 6.

Hermann, T. and Willamson, J. and Murray-Smith, R. and Visell, Y. and Brazil, E. (2008), Sonification for sonic interaction design, *in* 'CHI-08 Workshop on Sonic Interaction Design: Sound, Information, and Experience', Florence, Italy, pp. 35–40. — Chapter 6.

Brazil, E. and Fernström, M. (2009), Subjective experience methods for early conceptual design of Auditory Displays, *in* 'Proceedings of ICAD 2009 - The 15th International Conference on Auditory Display', Copenhagen, Denmark, pp. 11–18. — Chapters 1 and 6.

Brazil, E. and Fernström, M. and Bowers, J. (2009), Exploring concurrent Auditory Icon recognition, *in* 'Proceedings of ICAD 2009 - The 15th International Conference on Auditory Display', Copenhagen, Denmark, pp. 56–59. — Chapter 5.

Research Acknowledgements

Portions of my research were conducted in collaboration with others; here I make clear the role of each person.

Firstly, I owe great and deep thanks to Mikael Fernström, my mentor and fellow collaborator on many research projects, for his role as advisor in my research.

Secondly, I would like to thank Laura Ottaviani, who helped in the running of the studies in Chapter 4 and who provided a different viewpoint to look at data.

Thirdly, I would like to thank all the members of the Sounding Object project as the work in Chapter 4 gained much from various conversations with these members.

Fourthly, I would like to thank John Bowers, who helped in reviewing aspects of Chapter 4 and in the design of third study in Chapter 5. He provided a different viewpoint and much sagely advice during various conversations.

Fifthly, I would like to thank all the members of the Perception and Sound Design Group at the Institut de Recherche et Coordination Acoustique/Musique (IRCAM), France for their many conversations as part of the COST SID STSM ICO061-3756 – 'Expanding and Synthesizing Sound Taxonomies Knowledge and Methods'. These conversations contributed to an understanding of their work and motivations for the CLOSED project taxonomy used as in Chapter 6.

Acknowledgments

I could not have accomplished my research without the support, guidance, and inspiration of many people.

To Mikael Fernström: Without your guidance, patience, advice, and help this simply would not have happened. Thanks for illuminating the way and giving me the confidence to follow it.

To Liam Bannon: Story and continuity, without this advice I know this thesis would be so much poorer. Thanks for all your support and guidance.

To John Bowers: The craziness and inspirations you brought to the IDC helped everyone see things differently. I am especially grateful for all your guidance, patience, advice, and help that was so freely given. Thanks also for taking the time to read the thesis and pass along those really insightful comments and suggestions.

To Anne Murphy: Without your help I do not think that I could have managed to get everything done and I am sure I would have forgotten something like registering!

To Colm McGettrick: Whatever the need was you were there to provide an open ear and when necessary point out alternatives. I really want to say thanks for sharing all those mad ideas as they inspired me in my own research.

To my colleagues and friends from the Interaction Design Centre and UL: thank you for your support, ideas, and the friendships we have shared. The Interaction Design Centre is, in my opinion, one of the very best research groups to work in and around. You made it the crazy, interesting, maelstrom of creativity that it is. Tony and Michael, thank you for helping me understand the trees and the forest.

To my colleagues and friends in ICAD, Cost SID, DAFx, ConGas, ISON and others, thanks. Without such a great academic community I would never have managed to stay the course. A final but especially big thanks to all the participants and mentors at the ICAD Doctoral Symposiums I attended which helped me focus my research.

To my family: Thanks for putting up with this for so long. Your help and support made this a reality.

To all on life's journey, a prayer passed along to me from another academic traveller with credits to Michael Leunig.

> God bless the lost, the confused, the unsure,
> the bewildered, the puzzled, the mystified,
> the baffled, and the perplexed.
> Amen.

Contents

List of Figures

List of Tables

Part I

The Introduction

This part of the thesis covers an introduction to the research and the relevant background material for the questions being posed.

Chapter 1

Introduction

> The leaves of the trees along the Mardyke were astir and whispering in the sunlight. A team of cricketers passed, agile young men in flannels and blazers, one of them carrying the long green wicket-bag. In a quiet bystreet a German band of five players in faded uniforms and with battered brass instruments was playing to an audience of street arabs and leisurely messenger boys. A maid in a white cap and apron was watering a box of plants on a sill which shone like a slab of limestone in the warm glare. From another window open to the air came the sound of a piano, scale after scale rising into the treble.
> *A Portrait of the Artist as a Young Man*, James Joyce

An increasing number of the activities that people perform are technologically-mediated, e.g. through the use of computers, PDAs, mobile phones, etc. Interfaces and interaction models for modern technologies utilise a rich set of dimensions to improve the user experience by more deeply engaging the user's attention. As a result, there has been increased interest in alternative sensory modalities other than vision, with sound and haptics being the most studied. The development of alternative sensory modality interfaces has lead to greater interest in research on how to design more useful and usable interfaces. The problems of how to design for non-visual interfaces and the problems of understanding their interactions, and how people use and appropriate them have lead to an increasing demand for research to answer these and related questions.

This thesis provides a set of methods for use in Auditory Display, that address a range of these questions for interfaces using the auditory modality. It is focused in particular on non–speech sounds, primarily the everyday sounds of the world and how to design and select sounds for use in Auditory Displays. These studies highlight a number of investigations focused on the concurrent presentation of Auditory Icons, on the identifiability of presented sounds, on the perceptual scaling of sounds, and on the associations and semantics created by listeners for everyday sounds.

Human computer interaction (HCI) and interaction design (ID) are the research areas that concerned with this type of question and many others. Both of these research areas are interdisciplinary and draw on a wide variety of sources to assist in answering research questions. These research areas have helped and influenced the development of many related fields, one such sub-field is that of Auditory Display (AD). It is in the sub-field of AD that this thesis and its work is situated. In the context of this work, speech communication is considered as part of the wider definition of Auditory Display. However for this work, a narrower view is taken, that focuses on a smaller subset of Auditory Display excluding speech. The AD field is multi-disciplinary as shown in Figure 1.1 and is primarily focused on the use of computational devices and interfaces that use sound to communicate information to a user.

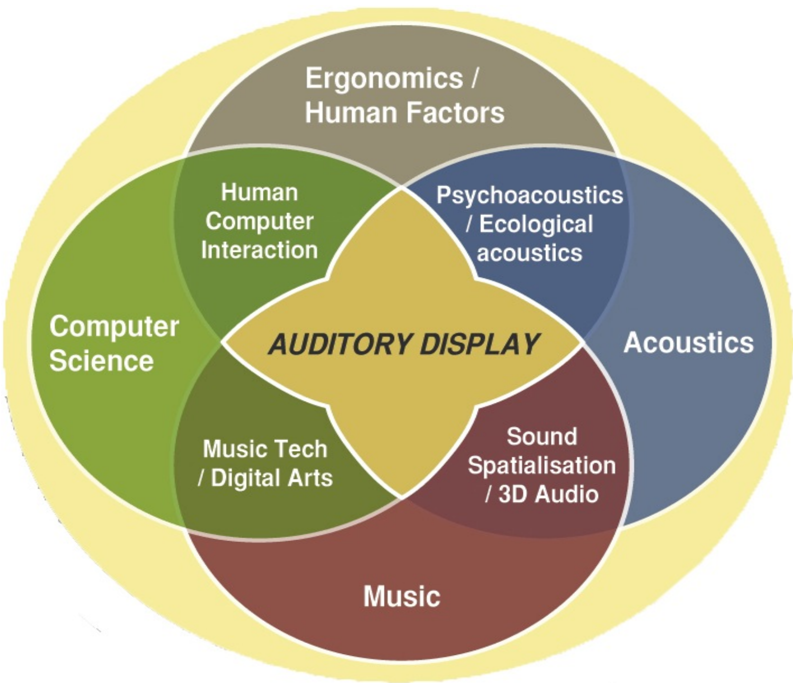


Fig. 1.1*:* Where Auditory Display fits within the wider research contexts.

It has been show that sound is an effective means of conveying information, and there is great potential to simplify user interfaces through the use of well-designed Auditory Displays. There are several situations where vision is not fully available for use and thus an Auditory Display may be suitable for presenting information. Examples of such situations include:

- Computer use in low-vision or dark environments.

- Computer use by visually impaired users.

- Users "*on the move*" where attention needs to be focused on the environment and potential dangers.
- Mobile computing or ubiquitous computing devices where a small form factor limits the visual display size.

The use of sound to convey information must be mindful of a number of issues. These issues can occur individually or simultaneously. Auditory Displays when designed poorly can be annoying and this can cause confusion or the Auditory Display being ignored. Auditory Display design must also consider the resulting æsthetic qualities of the sounds. Many sounds occurring simultaneously or close together can result in a cacophony, which increases the difficulty in hearing any one of the sounds and resulting in a difficulty in terms of interpretation. Masking is when one sound cannot be heard due to another sound and it or other similar interactions of two or more sounds can be problematic. The lack of specific guidance or methods to help a designer select the best sound for use in an Auditory Display means than the aforementioned problems will occur more often. Sound disappears once heard and as such it can be difficult to remember any distinctive part of the sound or of the wider soundscape, especially when too many different sounds or when a long sound has been played. A known problem (McGookin and Brewster, 2004) is the field of Auditory Display is the *lack of available guidance for Auditory Display designers*[1] who wish to use sound for their designs (Lumsden and Brewster, 2001, Frauenberger et al., 2007).

Sounds, in the context of Auditory Displays are often divided into three categories: speech, music and everyday sounds (Gaver, 1989). Speech and music are easily understood types but in this thesis the focus is on *everyday sounds*. An everyday sound is any non-speech or non-musical sound heard in the world. This varies from animal sounds (e.g., cows mooing, a dog barking), people (e.g., laughing, shouting, walking, opening a door), tools (e.g., hammering, sawing), transport (e.g., a car, helicopter), electromechanical signals (e.g., telephone, doorbell), liquids (e.g., water splashing, a river flowing) and so on. Auditory events encode many high-level perceptual attributes, such as the material of a door slamming, its intensity, the size of room or space, type of environment, and many others.

Everyday sounds are often heard by the action of agents (e.g., a person sawing) on objects (e.g., a block of wood) in the environment. These sounds also include the actions of objects upon other objects such as the sound of waves hitting a beach or the sound of a wave breaking. People develop a special listening skill through their lives that is attuned to everyday sounds.

[1] In the context of this thesis, designers refer to Auditory Display designers. Their experience can range across a variety of topics including interaction design, acoustics, music, software, and hardware but are typically proficient in only one or two of these topics.

This type of sound is very rich in detail and can be easily comprehended without extended training (Gaver, 1988). Everyday listening (Gaver, 1989) is a continuously occurring activity. It is a type of listening focused on the attributes of events in the world. Examples include listening to the speed of a person walking, to the number of bounces of a ball or to the type of car that is passing.

The work of Gaver (1989) was inspired by the work of Gibson (1966, 1979) on an ecological approach to vision. This perspective takes a different view than many other theories of perception. It states that perception picks up the complex events or entities in the everyday world where the world and the events or entities within it *afford* the necessary perceptual information to the viewer. In this case, perception can be seen as the pick up of information from the environment. This would mean that perception is not mediated by inference or by memory or by any other processes which uses a mental representation (Fodor and Pylyshyn, 1981). This differs from other theories, as the descriptions of the world are not limited to physical dimensions. It also means that explorations of the world can occur over time rather than as a viewing of a particular stimulus in isolation. It was the ecological views of Gaver (1989) and of Gibson (1966, 1979) that has most influenced previous research into Auditory Icons. A deeper coverage of the concepts and how they frame aspects of Auditory Icons and how this relates to the work in this thesis is discussed in Chapter 2. Bell (1999) uses the term "*ecology*" to describe in a qualitative fashion the relationships between people and their environment where the definition includes all the aspects for a specific experience. Bell (2001) further suggests that ethnographic studies into these ecologies can help in understanding the experience and makes it possible to constructively inform the design. This thesis does not extend to ethnographic studies but it does follow a similar belief to Bell's whereby improving the understanding of people's experiences of Auditory Icons will help a auditory designer ensure that their sound choices will fit with people's expectations.

The work of Gaver (1989), of Battarbee and Koskinen (2005), and of Bell (2001) show the need for the human computer interaction community to consider ecological and ethnographical approaches. In Auditory Display, user experience (UX) evaluation has not been explored in great detail. User experience is the practice of designing devices or systems, in particular the architectural and interaction model where the focus is placed on the quality of the experience and on culturally relevant solutions with less emphasis being placed on improvements in the functionality of the devices or systems. It is important to have methods that can obtain this type of experience and cultural information directly from users that are free from bias or interpretation. This type of procedure improves reliability by ensuring that interviewers do not bias participants and ensures diversity by preserving the individuality of participants. The methods that constitute the foundation of the auditory design framework are aimed at eliciting this type of information in the early stages of Auditory Display design. This foundation is based and created from the experiences derived carrying out these studies, it is outline in Section 7.3.1 in Chapter 7. This information can improve the quality of Auditory Display development and is precise in comparison to methods such as interviews or questionnaires. This approach can help to avoid any mismatch between the auditory design specifications (i.e., what the designers intended for the listeners to hear) and the actuality of what the users hear. These techniques are used to populate the design space opened by the proposed framework. In this thesis the techniques are showcased with the results being used to highlight some interesting trends, however the purpose of this work is to develop the foundations of the framework and suitable methods for inclusion within it.

The three techniques that form the foundation of the framework in this thesis are similarity scaling using multiple comparisons, listening tests for multiple sounds, and the repertory grid for obtaining tacit knowledge from participants. The approach of Bonebright (2001), Scavone et al. (2000) to scaling of sounds within a multidimensional space can help in collecting, in a structured fashion, a cross section of diversity with regard to the participant's responses. Understanding the subjective experiences of a listener is multifaceted in relation to sound. There is no single technique or approach that addresses all the issues or insights raised. The only approach to successfully gather all the necessary information is to use triangulation with complementary approaches. The listening test paradigm can complement multidimensional scaling by exploring the listeners own internal concepts by analysing their written responses. The causal uncertainty method (Ballas, 1993) can further be applied to the responses to further understand the subjective listening experiences with regard to a single sound or to combinations of sounds. The results of this method show where beyond the masking of audio, subjective interpretations of the sounds as similar sounds may result in sounds being merged to a single concept by the listener, further increasing subjective difficulty of identification. A third technique, the Repertory Grid (Kelly, 1955) can help in the triangulation of listeners subjective experiences. It provides information for design inspiration and also for evaluation of the user experience. The rich elicited attributes provide an insight into the idiosyncrasies of users and how users infer information or aspects from the particular designed attributes of the product or interface. The three techniques open new areas that were not previously explored in Auditory Display with the first study of synthetic parametrically controlled models based on physical modelling, the first exploration into techniques that can address the issues of concurrent Auditory Icons, and the first use of the repertory grid to gain a deeper insight into listeners categorisation of Auditory Icons.

This thesis investigates the everyday sounds of the world that people have adapted to understand and interpret through their daily lives and expands the existing knowledge in this area of Auditory Display. The understanding of the sounds varies from person to person and is linked to their experiences and the particular situation. Auditory Icons were defined by Gaver (1997) as "*everyday sounds mapped to computer events by analogy with everyday sound producing events*". They are similar to visual icons and replace the graphic or glyph, which represents the meaning of the sign, such as an on/off button with an auditory equivalent.

The SonicFinder Gaver (1989) used a number of Auditory Icons, such as mapping the selection of a file, folder, or application to a hitting / tapping sound where type of item (file, folder, or application) was used to determine the materiality of the struck object and size of the item was linked to the frequency. The bigger the item selected the bigger the sound of the object being struck. A number of other mappings including dragging, dropping, opening, and copying were presented using similar Auditory Icons. These mappings represent actions upon items of computer events and where each item-action combinations has a related sound producing event such as the hitting / tapping object mapping. This type of iconic mapping can help in making Auditory Icons both intuitive and coherent. Understanding the identification of sounds and of how sound producing events are interpreted by listeners helps in mapping events from the everyday world onto events in the computer world. Gaver (1989) highlighted two solutions for creating Auditory Icons to represent events, which do not have a counterpart in the everyday world, film / foley sounds and sounds that are analogies of existing sound producing events. The techniques presented by this thesis provide a number of ways to understand how people map everyday, film / foley, and sound analogies of existing sounds. These types of sound and, in particular, everyday sounds are defined and discussed in greater detail in Chapter 2.

Auditory Displays are designed to use sound to communicate information to a user but many of the designers of such displays come from computer science or other similar programming backgrounds and may not be aware of the differences between the design of a graphical interface and the design of an auditory interface. In many cases techniques adapted for spatially oriented visual displays cannot be applied to temporally oriented Auditory Displays. The findings in this thesis primarily help in casting light into new avenues in Auditory Icon research, however the methods used can help designers build similar types of Auditory Display. The methods can help designers address the issue of how to select the appropriate sounds for their interfaces. The focus of this thesis is not specific to the production or dissemination of new methods for this rich design space, but the purpose is to provide the foundations of a design framework, that can provide empirically inspired methods for the design of Auditory Displays.

There are many issues to be addressed when designing an Auditory Display and they include answering such questions as: *what do people think of when they hear a sound in terms of the identity of the sound and its source*, *what identification or understanding they have of a heard sound or source*, and *what are the sounds or sources are most likely to be confused*. An understanding of such issues is important as it helps to ensure that the correct mappings, metaphors, and sounds are chosen for use in Auditory Displays. Humans have rich discrimination ability when it comes to sound and interfaces that effectively exploit this ability will ensure better clarity and understanding of the the information being transmitted. The spatial dimensions of sound are not addressed by this research, which focuses on the macro temporal and spectral properties of sounds. Exploring issues of sound spatialisation in the context of these questions is an area for future research.

The lack of guidelines or frameworks for designing with Auditory Displays using Auditory Icons has resulted in difficulties when attempting to maximise the advantages of Auditory Icons. The explorations presented in this thesis can help Auditory Display designers to make the best selections of Auditory Icons. The areas and issues explored include multiple sound presentation, the use of synthetic sounds, and providing an understanding of a listener's tacit knowledge. Presenting multiple sounds at the same time reduces the time to present information and allows for real time comparisons between the sounds and the events or the data they represent to be carried out. The issues of realism and of a sound's ability to convey information for synthetic sounds are explored. Another issue tackled is the understanding of how an everyday sound is classified or identified by listeners. A better understanding of the inner workings of how people think with regard to sound can offer more effective mappings and metaphors. These results and the application of the methods presented help to provide the information necessary for designers and for researchers. This information includes determining what people think of when they hear a sound in terms of the identity of the sound and its source, what identification or understanding they have of the sound or source, and what sounds or sources are most likely to be confused.

1.1 Thesis Statement

Evaluating and validating everyday sounds for use as Auditory Icons is difficult, in particular the issues of perceptual mapping of object properties, of identifying concurrently presented everyday sound, of the metaphors and categorisations used by people for everyday sounds. There are a number of methods that could be applied to solving or minimising these issues. The research questions in this thesis will explore and defend these methods by providing answers of use to Auditory Display designers.

A general observation from many Auditory Display designers (Fitch and Kramer, 1994, Mynatt, 1994, Frauenberger et al., 2007, Perry et al., 2007) is that Auditory Icons are not easy to design. This thesis has synthesised and organised existing research on Auditory Icons, and extended this work by exploring new methods to address this design problem using three exemplar studies. The studies provide an empirically based auditory design process suitable for use at the early stages of design of Auditory Displays. These studies and the methods provide indicative trends to assist researchers and designers in the selection, design, and use of Auditory Icons. These methods are psychoacoustically inspired and take a lightweight approach making them suitable for use outside of strict laboratory conditions. Designers that do not have access to dedicated facilities, such as listening booths or anechoic chambers can benefit from these methods. This dissertation uses an interdisciplinary approach which fits at a point between the more detailed psychoacoustical studies (Zwicker and Fastl, 1990) (i.e., signal masker approach) and the type of traditional HCI study focused on application usability heuristics or task completion measures. Psychoacoustical studies have a long history using a signal masker approach where two sound stimuli were typically presented and where pink noise or similar artificial sound are used so at times one of the sounds is masked by the other. This type of experiment uses artificial tones to study the auditory system at its lowest level, however this approach is not useful or meaningful for designers of Auditory Displays, as the results do not provide immediate solutions for their designs. This thesis does not give any prescriptive rules and guidelines for designing specific types of Auditory Displays. It does reiterating some well established psycho-acoustic guidelines as each Auditory Display application will have to operate in a particular set of requirements. These requirements are determined by the tasks and contexts, that the Auditory Display is to be used in.

Evaluating and validating Auditory Icons for application use is an important aspect for the success of an Auditory Display. Typically, Auditory Icons are chosen in an *ad hoc* manner (Lumsden and Brewster, 2001, Frauenberger et al., 2007) and are used in applications where appropriate studies on the suitability of the sounds were not conducted. The effects of *ad hoc* Auditory Icon selections can contribute to negative effects such as increased confusion. This thesis provides empirically based design methods that help designers investigate Auditory Icons. The results of these methods can explore if modifications significantly improve Auditory Icon identification or if they must be constrained due to the need to preserve the existing mapping inherent in the sound.

The methods in this thesis differ from an experimental paradigm as the focus is on a wider view. The aim is to provide methods and studies that can allow designers to replicate this approach within their own design process. Many existing approaches for quantitative analysis and validation make it difficult for designers to deal with information other than that strictly related to Auditory Display use. User experience, in particular the feelings, needs, and tacit participant knowledge are aspects of subjective experience information. This kind of information requires an analysis of the psychological relationship between the users and the Auditory Displays or sounds. This thesis uses a number of exploratory studies to illustrate, with examples, how the techniques presented in this thesis can be used as subjective user experience information gathering tools in the early stages of Auditory Display design:

- The multidimensional scaling technique gathers information about listeners' responses to a certain group of sounds and extracts the diversity of perceptual scaling used to get design relevant information about how the listeners respond to certain subjective features of the sounds.
- The listening test combined with the causal method for simultaneous sounds can be considered both an informational and inspiration gathering technique as it gathers written responses from listeners that can illustrate metaphors of inspiration and it can provide subjective informational measures about which sounds complement each other.
- The repertory grid technique can provide an experience landscape of listeners' responses to a certain group of sounds and obtains tacit attributes of the sounds along with any idiosyncrasies of how an individual listener infers information from a particular sound or aspect of the sound.

The work in this thesis sought to answer the following five research questions:

RQ1 Does the subjective realism of a sound affect the response of a listener to the sound ?

RQ2 Do listeners subjectively hear the same physical properties of objects when both synthesised and sampled versions of the same sounds are used ?

RQ3 Do listeners subjectively use the action and object categories of multiple sounds for sound identification ?

RQ4 Is it possible to reduce the subjective confusion of the sounds using action and object categories ?

RQ5 Can a listener's tacit criteria about how they attribute meaning to everyday sounds be elicited ?

These five research questions span a range of issues faced by Auditory Display designers when attempting to map or use everyday sounds in an Auditory Display. The methods presented through work seeks to answer these questions and through the use of these methods to improve the design of Auditory Displays using Auditory Icons.

The sounds used in Auditory Displays are typically not 'live' sounds as they occur in the world but are either recordings / samples or synthetic sounds. It is important to understand the different types of sound and what they are most suited for before they can be best used in an Auditory Display. There have been few studies exploring both synthetic and sampled sounds in the field of Auditory Display. There are different advantages and disadvantages to each type of sounds. Synthetic sounds can be computationally less expensive than sampled sound files both in terms of processing time and storage space, and can produce more dynamic and varied sounds avoiding repetition of the same sound. Recent developments in sound synthesis allow for real-time dynamic control of sounds where they represent physical interaction sounds such as impacts or scratching. Sampled sounds are typically more natural but are limited to what was actually recorded. These physical interaction sounds are useful due to their familiarity to listeners as everyday sounds and because they can easily be linked to physical actions in interfaces or mobile devices. This variety of sound and the additional control over the individual sound can improve interfaces, however, there are few studies (Cook, 2002a, Aramaki and Kronland-Martinet, 2006, Peltola et al., 2007) into synthesised sounds and particularly into the area of a sound's ability to convey relevant perceptual information. The previous research has concentrated on the physical modelling or artistic applications of such sounds rather than on the subjective understanding of how listeners interpret the sounds. The first study in the thesis is focused on exploring both types of sounds, the rest of the studies use a wide selection of everyday sounds. The studies in this thesis explored a wide selection of sound types within the category of everyday sounds. This prevented any misleading effects from arising due to the use of any one set of sounds.

1.2 Methodological Approach

There is no single methodological framework that can deal adequately with the complex socio-cultural context of Auditory Display design in a coherent and non-reductionist manner. This is a similar problem faced by most design oriented research, one suggestion by Melles (2008) has been to take a pragmatic stance towards methodology, where methods are selected and combined according to their usefulness for achieving specific goals. This view on design research has found support in many methodological dialogues such as those discussing multi-method research (Morse, 2003). This thesis follows this approach to provide a set of explorations and methods for Auditory Icons selection in the early stages of design.

The results of these explorations generated the foundations of a design framework, structured to support the selection of sounds while allowing the exploration of specific aspects of these sounds. The framework was designed to be accessible and open to contribution. An understanding of these methods is a prerequisite to an understanding of the framework itself. It is presented in the conclusion of this work in Section 7.3.1 as part of Chapter 7 and is an area for future studies.

The selection of methods used in the explorations in this thesis was based on a number of criteria including ease of use, prior similar use in the field or related field, ability to concisely present the results, and time required to use the method. This aim in using a selection of techniques, was similar to the discount technique approach in HCI. When the techniques are compared with traditional psychoacoustic methods, they are typically easier to learn, can be finished within a single session of less than an hour, and provided detailed data. The methods used in this thesis are multidimensional scaling (McAdams et al., 1995, Bonebright, 2001), causal uncertainty (Ballas, 1993), and the repertory grid (Fransella et al., 2004).

The analysis stage of these methods uses several other statistical techniques (Everitt and Hothorn, 2006), which are discussed with examples to illustrate how they are typically used. The explorations are supplemented in the studies within this thesis by questionnaires and participant comments. These supplementary instruments focused on items of particular note such as age or the type of childhood environment of the participant within the individual study. The methods covered in this thesis address both the sequential and the simultaneous presentation of Auditory Icons, as shown in Figure 1.2. The exploratory studies are used to illustrate the methods and highlight the different types of information or inspirations that each can contribute about the particular Auditory Icons being investigated.

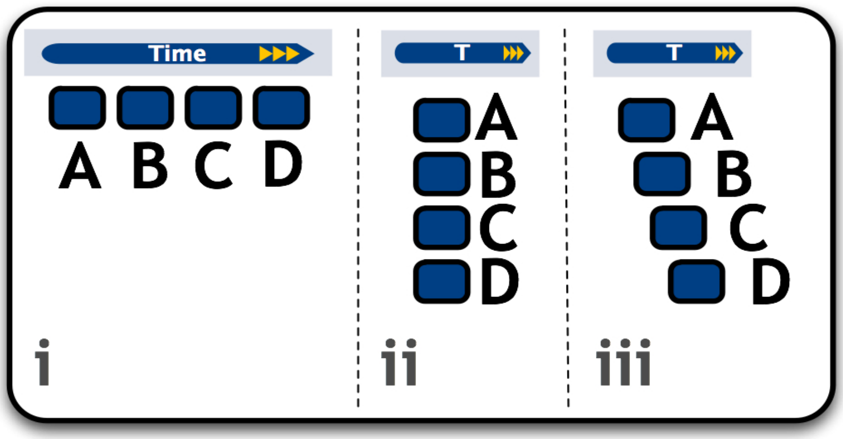


Fig. 1.2: *Sounds A,B,C and D in i) sequential presentation, ii) concurrent presentation, and iii) onset staggered concurrent presentation.*

1.2.1 Overview of Explorations

In this thesis, three broad issues and five related research questions are addressed through the use of three explorations. These studies provide results that open new avenues in the area of Auditory Icons in the wider field of Auditory Display. The methods presented are a further by-product of this research and if applied by researchers or designers can help them with the selection, design, and use of Auditory Icons for their Auditory Displays.

First exploration The first exploration in this thesis explores synthesised sounds for their realism and their ability to convey physical information to the user. A new method for exploring the perceived source identification and perceptual properties of synthetic sounds is developed. It is a multidimensional method, where sounds can be explored using direct sonification (Brazil, 2003). This allows for the mapping of relationships between the sounds to be explored in the context of the entire set of sounds, rather than in one to one comparisons for each possible sound pairing in the set of sounds being examined. This type of methodological approach has been suggested by Gygi and Shafiro (2007, p. 3160-3161) for "*developing perceptually motivated synthesis models*" and for "*designing new sounds for representing complex information structures*". They suggest that by exploring the mapping of relations between acoustic parameters and physical or perceptual aspects that the final synthesised sound can be "*substantially enhanced*". The first study echoes the thinking behind Gygi and Shafiro (2007) and explores if a particular set of synthetic sounds can convey the same type of perceptual information as recorded everyday sounds.

Second exploration Examining prior research for using simultaneous everyday sounds in the first exploration highlighted the lack of studies in this area. This motivated the second exploration, which looked particularly at the confusion and identification of sounds where several sounds happened at once. The results provided empirical details in the area where previously researchers had used educated guesswork. The methods provide results, which showed the particular sounds that were confused and the extent of the confusion. This information can help in determining what sounds are appropriate to use in a case where several different sounds need to be presented concurrently.

Third exploration The third exploration elicited and identified the key attributes used by participants when describing everyday sounds. The motivation for this study was based on both the first and second explorations. The second exploration showed that listeners use a rich vocabulary to discuss and define their meaning for everyday sounds. The first exploration looked at perceptual features and the mapping of sound properties to synthetic sound models. These discount HCI style studies highlighted new methods focused on subjective data that helps provide deeper insights into the perception of the sounds by listeners. In specific, the results provide tangible metrics or inspiration that can be applied by designers. The understanding of attributes and vocabularies relevant to listeners for a set of sounds had not been explored in the area of Auditory Icons. This understanding can point to the important attributes, scales and metaphors for the set of sounds in their tacit classification of everyday sounds. This type of tacit or implicit knowledge is difficult to elicit and there are no existing studies exploring this issue for Auditory Icons in Auditory Displays. The third exploration is concerned with the question of how it is possible to make the listeners' tacit knowledge explicit, with regard to the sounds they heard sounds during the study. This has been a known problem in the field of Auditory Display as posed by Walker and Kramer (2004, p. 168) who noted both researchers and designers need to "*know their users at a more cognitive level*" than has previously been addressed in traditional studies. They point out that these studies typically "*diminished the importance of learning and experience*" for interpreting sounds. The categorisation and similarity data from this exploration provided new insights into the listeners' auditory perceptual spaces. It does this as the first stage of gathering a listener's subjective experience information. This can be used to help determine their needs or desires whether latent or explicit in a manner that helps communication between the user and designer.

1.3 Contributions to the existing research in the field of Auditory Display

This thesis has synthesised and organised existing research on Auditory Icons, and extended the prior research in this area. The results of the explorations in this thesis have provided new insights into three different issues and their related questions. The explorations in this thesis showcase a set of techniques that can assist in understand behaviours closer to the users' needs and intentions for use with early design development. The thesis has furthermore provided a number of methods and adaptations of existing methods that can assist in the selection, design, and use of Auditory Icons.

The areas explored in this thesis provide a range of methods addressing a number of topics. These topics include how to incorporate synthesised sounds that convey accurate information about objects to listeners, how to select sounds for use where many sounds will be presented simultaneously while still remaining identifiable, and how to develop a vocabulary for a sound set which highlights relevant sound attributes, metaphors, and listeners perceptual spaces of the sounds. These methods can be applied to all types of sound but this thesis uses and only makes its claims for everyday sounds.

1.3.1 Contributions to the methods and techniques in the field of Auditory Display

A secondary benefit of the research and explorations carried out in this thesis is a concise set of methods, which address a number of areas in Auditory Display design. These methods are useful to both researchers and designers. They can help in providing details and knowledge about Auditory Icons. This can ensure that the best selection of sounds can be made whether for realism, identifiability, segregatability, or relationship to language, metaphors and labels used by the participants.

The first study provides a method to explore synthesised sounds and that can determine if these sounds can accurate convey information about the objects being represented. Experimentally validated synthesised sounds selected using this method can ensure that the mapping designed by the Auditory Display designer are those they intended. An example is where the size of an object may be used to represent a variable where this method ensures that size of the synthesised sound as perceived by the listeners or users of the Auditory Display is the same size as that intended by the designer of the Auditory Display.

The second study is the first work to examine situations with multiple simultaneously occurring Auditory Icons and provides details on one method that can be used to select appropriate sounds, which will remain identifiable and segregable. Games or complex Auditory Displays for industrial processes where many variables or events are tied to auditory alerts are examples of where this method is useful.

The third study provides a method for eliciting an individual's tacit knowledge and provides a set of words and language from the individual, which describes the entire set of sounds and their interrelationships. This method can produce labels and mappings using the participants' language, which can assist the designer in selecting appropriate terms, mappings, and metaphors for the audience of their design.

The methods explored in this thesis are not meant to be taken as strict implementations or guidelines but rather as a set of exemplars that are open to customisation. The psychoacoustically inspired methods are suitable for use outside of laboratory conditions and are a lighter weight approach than standard psychoacoustic methods making them more suitable for the needs of Auditory Display researchers and designers. These methods aim to overcome certain limitations of existing interaction design techniques when the focus is not solely on strictly functional usage and requires the evaluation of emotions, affect, or tacit knowledge. The existing HCI techniques find it difficult to measure this complex and subject dependent data that can provide useful and relevant design information whether it is solely inspirational (metaphors) or more informational (identification / confusion details). These types of participatory methods (Sanders, 2005) allow for participants needs to be expressed by themselves and gathers more subjective information allowing for their more proactive involvement in the design process.

1.4 Thesis Contents

This thesis is divided into seven chapters. The first chapter is the introduction. The rest of the thesis is structured in the following way:

Chapter 2 "*Everyday Sounds*" - This chapter introduces what everyday sounds are and how they relate to Auditory Icons. It provides a history of the development of Auditory Displays. A background on everyday sounds and issues such designing with everyday sounds are introduced. A review of Auditory Icons in Auditory Displays and where the contributions of this thesis extend the existing work in the field of Auditory Display complete this chapter.

Chapter 3 "*Research Methodologies*" - This chapter introduces the methods for understanding Auditory Icons used in this thesis. It highlights the methods developed, the rationale and the prior work that informed and contributed to the development of new methods for the comprehension and design of sound.

Chapter 4 "*Investigating Auditory Icons Using Multi-Dimensional Similarity Ratings*" - This chapter introduces a new method for understanding Auditory Icons. It highlights how the method developed, the rationale and the prior work. It also describes an exploration to highlight the method's use. This chapter describes a novel study that provides answers for RQ1, "*Does the subjective realism of a sound affect the response of a listener to the sound ?*" and for RQ2, "*Do listeners subjectively hear the same physical properties of objects when both synthesised and sampled versions of the same sounds are used ?*". This method provides a useful approach for the investigation of parametrically controlled synthesised sounds.

Chapter 5 "*Investigating Auditory Icons Identification Using Causal Uncertainty*" - This chapter explores the issues of confusion and identification of simultaneous everyday sounds. Existing work on single Auditory Icons is expanded and opens a new avenue of research in concurrent Auditory Icons. The results and methods of the three novel studies in this chapter provide answers for RQ3, "*Do listeners subjectively use the action and object categories of multiple sounds for sound identification ?*", and for RQ4, "*Is it possible to reduce the subjective confusion of the sounds using action and object categories ?*". This provides a starting point for the understanding of concurrent Auditory Icons for Auditory Displays where comprehension and identifiability are important issues.

Chapter 6 "*Investigating Auditory Icons Identification Using The Repertory Grid Technique*" - This chapter takes an existing method and applies it for use in understanding everyday sounds. It highlights the method developed, the rationale and the prior work, that informed and contributed to the development of the methods and describes an exploration to highlight the method's use. This chapter describes a novel study, which provide answers for RQ5, "*Can a listener's tacit criteria about how they attribute meaning to everyday sounds be elicited ?*". This provides a vocabulary for everyday sounds that details the individual auditory attributes of the sounds as determined by the individual. This can be used to provide appropriate language, metaphors, and mappings for Auditory Displays.

Chapter 7 "*Conclusions*" - This chapter presents a summary of the work from the previous chapters and relates it back to the three research questions, discussing to what extent the five research questions have been answered. The limitations of this thesis are discussed and suggestions for their potential resolution are proposed. The future directions for work based on this thesis are also outlined.

Chapter 2

Everyday Sounds

> "A noise like of a hidden brook. In the leafy month of June, That to the sleeping woods all night singeth a quiet tune."
> *The Ancient Mariner (pt. V, st. 18)*, Samuel Taylor Coleridge

This chapter introduces what everyday sounds are and how they relate to Auditory Icons. A background on everyday sounds and on Auditory Icons with issues relevant to their design is introduced. This background covers material from sound design, communication theory, and electroacoustic music to theories of perception. The design of an effective Auditory Icon and of an effective Auditory Display requires an interdisciplinary approach with an awareness of concepts from human perception, acoustics, design, the arts, and engineering. The development of Auditory Display systems using Auditory Icons is then explored. This establishes the current state of research in the field of Auditory Displays using Auditory Icons. The lessons learned from the past development of Auditory Display systems using Auditory Icons is useful in determining what was successful and where more research is required. In the case of most Auditory Display development, one key area that is often neglected is that of evaluations or usability studies. Many of the successful Auditory Display systems of the past where prototypes which highlighted a particular technology or novel approach in Auditory Display design. A major issue with these early Auditory Displays was a lack of detailed studies. This chapter concludes with a brief summary and an introduction to the next chapter on the issues of auditory presentation.

2.1 Defining what is an everyday sound

The sounds used in this thesis in its explorations are Auditory Icons, which are based on everyday sound, and do not typically fit into the categories of music or speech. Musique concrète is included in the category of music for this thesis. This definition can be interpreted as all naturally occurring sounds excluding speech or music sounds that occur in the real world. The definition of an everyday sound is taken from Vanderveer (1979) who defined it as:

> "Any possible audible acoustic event which is caused by motions in the ordinary human environment. ... Besides having real events as their sources ... [everyday sounds] are usually more complex than laboratory sinusoids, ... [everyday sounds] are meaningful, in the sense that they specify events in the environment. ... The sounds to be considered are not part of a communication system, or communication sounds, they are taken in their literal rather than signal or symbolic interpretation." (1979, p. 16-17)

2.1.1 Previous research investigating everyday sounds

The work in this thesis combines many concepts and techniques from a spectrum of disciplines and illustrates the range of knowledge required when researching in everyday sound. The interdisciplinary nature of sonification as discussed by Walker and Kramer (2004) can be equally applied to Auditory Icons and everyday sounds as knowledge from a range of disciplines need to be considered:

> "By its very nature, sonification is interdisciplinary, integrating concepts from human perception, acoustics, design, the arts, and engineering. Thus, development of effective auditory representations of data requires interdisciplinary collaborations using the combined knowledge and efforts of psychologists, computer scientists, engineers, physicists, composers, and musicians, along with the expertise of specialists in the application areas being addressed." (2004, p. 154)

There have been few studies solely focused on the issue of Auditory Icon presentation but there are relevant studies from the areas of everyday listening and of everyday sounds. These studies can provide a starting point for how everyday sounds have been studied which can be used to inform investigations into Auditory Icons. The seminal work of Nancy Vanderveer (1979) on everyday sounds in the field of perception research is a starting point for how these types of sounds have been studied. Vanderveer followed the Gibsonian (ecological) approach (Gibson, 1966, 1979) to perception, where sound affords information to a listener within an environment. According to Vanderveer (1979) the three types of information sound affords are orientation information for within an environment, orientation information related to a specific event, and specific information about an event. In particular interest to this thesis, were the experiments on identification by Vanderveer (1979), where participants provided a free text response or description to the tape recordings of everyday sounds played to them.

These responses were made in a free text format by the participants and as such were highly individualised. These results were judged by Vanderveer (1979) using a heuristic which defined the "*correct*" response or responses to a sound. The different responses from the listeners can be very difficult to analyse using the heuristic and the listeners' responses themselves were very informative. Vanderveer noted that the participants in the studies did not tend to refer to the sound itself but rather referred to the events that caused the sound. The participants focused on three particular points with regard to the sounds and these were the action, the object/s involved in the action, and the place where the action occurred. The studies by Vanderveer (1979) found that the agent was rarely described even if it was the person or the object that had created the sound. These studies provide an important reference point for this thesis as many of the experiment techniques presented in this thesis were inspired or directly taken from the work of Vanderveer (1979). The methodological approach of using listening tests with free sorting was used in this thesis as the exploratory approach for researching large heterogeneous sets of auditory stimuli.

2.1.2 Taxonomies & Categorisation schemes

In order to answer the question about what organisational criteria and features are used by a listener, it is important to explore the prior work into these issues which is concerned with Auditory Icons and everyday sounds. Gaver (1988) work followed an ecological approach that to the development of the concept of everyday listening. This methodology is founded on the concept that sound provides information about the interaction of materials within an environment. Gaver's viewpoint, of a hierarchical structure of sound-producing events, is shown in Figure 2.1. In this view, the world of sound occurs between interacting materials of the type aerodynamic, liquid, or vibrating objects.

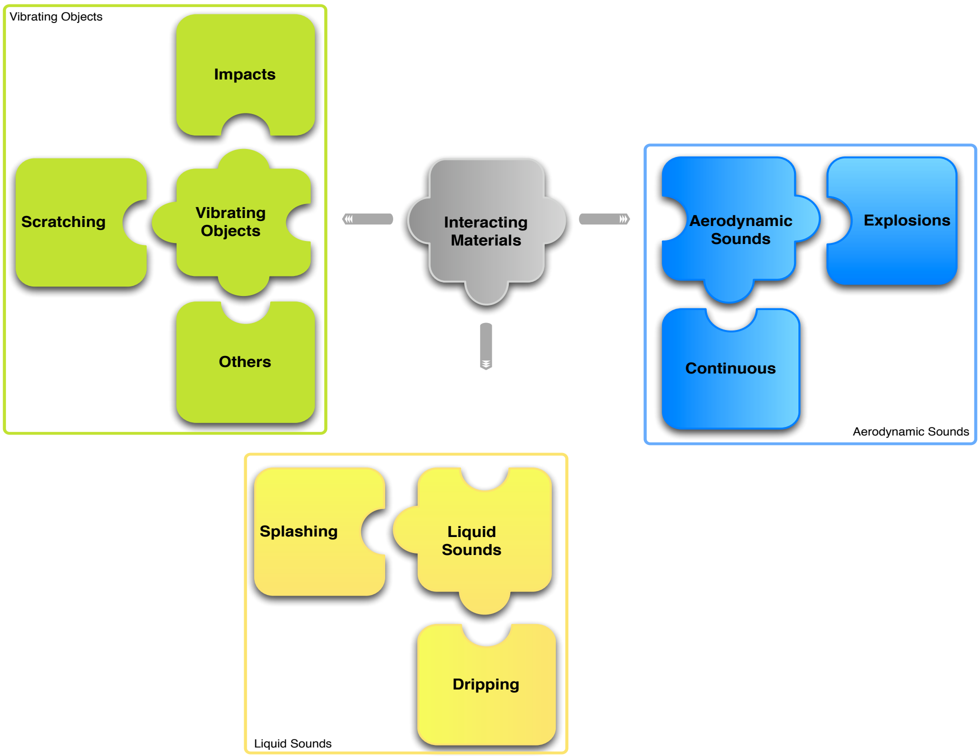


Fig. 2.1: *Hierarchical description for simple sound events as discussed in Gaver (1988).*

A sound event, according to Gaver, occurs when there is an interaction between two materials. The sound depends on the shape, size, material, and textures of the objects involved in the interaction. These sound events were broken down into one of three initial categories of interacting materials (liquid, gas, or solid), as shown in Figure 2.1. These initial categories break down further into particular sub categories such as impacts or splashes depending on the parent category. In the solid sub-category, Gaver proposed four distinct event types of deformation, impacts, scraping, and rolling. These distinct events have a related set of perceptually relevant attributes for these physical properties. In the instance of the impact event type, this would consist of the vibrating object and parameters related to this such as its materiality, its size, its shape, its surface hardness, and the force of the impact upon it. These event types and material categories form part of more complex events. These were classified by Gaver as temporal, compound, or hybrid. Temporal events consisted of a sequence of distinct event types of a particular sub category. Compound events consisted of more than one of the distinct event types. Hybrid events consisted of more than one of the three initial interacting materials categories (liquid, gas, or solid). Gaver (1988) saw this framework as a beginning and was open to it being organised in alternative ways. It was deemed to be sufficient for describing a wide selection of sound events by providing a framework for understanding complex sound events with physical attributes that have a relation with the human perception of everyday sounds. This framework provided a second important reference point for this thesis, as it built upon the work of Vanderveer (1979) and expanded the research in the field.

The work of Howard and Ballas (1980) built upon the work of Vanderveer (1979) but had a different account of perception as they explored similarities between the perception of everyday sounds and the perception of speech. This approach differed from Gaver's (1988) where the identification of sounds is seen as solely a bottom-up process where the information is taken from the context and the sound. Ballas and Howard (1980) presented a dual approach for the identification of sounds consisting of a bottom-up process and a top-down process. The top-down process used prior knowledge and expectations for the identification process. This approach stated that in the interpretation of sound events that human listeners rely on both the perceptual information from the sound and from the environment but in addition they rely on their own implicit knowledge. The first study that lead Howard and Ballas (1980) to form this approach was focused on the categorisation of sequences of brief sounds and explored the semantic and the temporal organisation of the sounds and the effect of these on identification performance. The result of this study found that the sequences of sound events that followed a grammatical structure rather than those randomly selected were more easily learnt.

These studies were further developed by Ballas et al. (1986), Ballas and Howard (1987), Ballas et al. (1987), Ballas and Mullins (1991), Ballas (1993) where the parallels between everyday sound perception and speech perception were explored. One finding from these studies was the importance of context for perceptually distinguishable sounds, which had confusing identification. This is similar to the homonym concept in language. The context of the sound event helped listeners to select from the potential alternative identifications for the sound. An example of this confusing identification would be where a sound such as frying bacon could be easily mistaken for the sound of rain. The homonym-type sounds were further investigated by Ballas and Mullins (1991) where quasi-homonymous type sounds were presented in the actual case of the particular sound or in an alternative case for the particular sound. The incoherent sequences were found to direct listeners towards the alternative causes and their identification for the sounds, while coherent sequences did not lead to better identification than in a neutral case. Ballas (1993) further explored factors such as ecological, cognitive, and acoustical variables and how these effected the identification of everyday sounds. The results of the study by (Ballas, 1993) showed that performance was related to many different variables. These variables included the typicalness of the sound, its ecological frequency (how often a listener heard that particular type of sound), its causal uncertainty (a measure for the number of reported alternative causes for the sound) and acoustical variables such as pitch.

The acoustic variables accounted for approximately half of the variance in the accuracy and in the identification times of listeners. This study suggested that the identifiability of a sound was related to several factors including the strength of the mental image associated with the sound, the context independence of the sound, the listener's familiarity with the sound, the typicalness of the sound to the stereotype for that particular sound, the ease of description by the listener of the sound, and the clarity of the sound. The studies by Ballas and Howard were an important point in establishing a dual approach for classification using bottom-up and top-down processes. In particular, this thesis uses their method of causal uncertainty as a measure in several of the explorations presented. The analysis of descriptions, the causal uncertainty measures, and the analysis of "correct" responses combined together can be useful in determining the sounds more suitable for use by virtue of them being less confusing. These results can also provide metaphors and descriptions from the listener's descriptions with regard to the sounds.

Another influence in this area of perception studies was the work of Guyot (1996) where 25 sounds were presented to listeners and they were asked to classify the sounds into distinct classes based on their perceptual similarity. The result of this work showed two strategies being used by listeners, the first used psychoacoustical criteria such as the type of excitation of the source, pitch, and temporal progression. The second strategy was related to how the sound was produced where it related to the source or to the action. These strategies were statistically analysed by Guyot and showed two different cognitive processes where one level of identification dealt with sources and a second, more abstract level dealing with actions. Guyot (1996) proposed a classification for everyday sounds based on Rosch's (1977, 1978) three levels of abstraction. The classification's basic level dealt with psychoacoustical, the supraordinate level dealt with identification by source, and the superordinate level dealt with abstract sound identifications like electronic sounds. This hierarchical classification of domestic sounds by Guyot is shown in Figure 2.2. Guyot's use of an additive tree to obtain categories from sorting data was an approach that introduced the use of statistical classification for determining hierarchical categories of sounds in perception studies of everyday sounds. This research was important in highlighting one approach using statistical classification in this context and its potential application to the work in this thesis. It reiterated the importance of sound source, event, action, and sound identification as criteria for listener classifications.

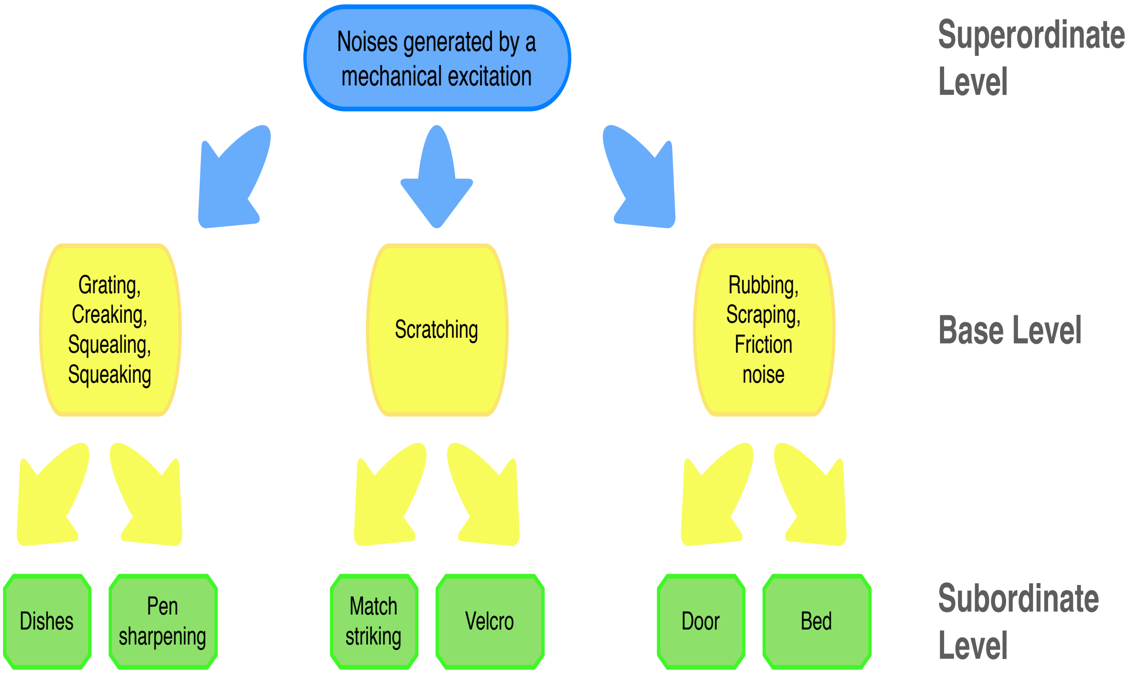


Fig. 2.2: *Guyot's (1996) hierarchical description for domestic noise.*

The work by Marcell et al. (2000) aimed to create a corpus of unambiguously identifiable everyday sounds and used free text classification to elicit the categories that listeners felt the sounds belonged to where the other sounds in the same category had similar characteristics as judged by the listener. The results after judging for equivalence were 23 distinct categories. The resulting categories ranged from machines to animals. The categories were very broad and varied from sound sources to physical locations to abstract ideas such as sleep or sickness. The judging for equivalence was an idea that echoed Vanderveer's (1979) heuristic for correct response and it is the method used in this thesis for the classifications of free text descriptions from listeners. Marcell et al. (2007) further explored these types of sounds with a focus on longer sounds. These sounds were non-musical and non-linguistic yet each sound formed a short sonic narrative or established a sense of place. This reiterates the importance of using complex and meaningful everyday sounds. The categories in this taxonomy were quite general and broke down into three type categories. The first was sources such as 'air transport' or 'tool', the second was locations such as 'nature' or 'kitchen', and the third covered abstract ideas such as 'sleep' or 'hygiene'.

The work of Gérard (2004) used classification experiments to study the perception of everyday sounds. Listeners were asked to classify the sounds together where they would be heard. The second experiment asked listeners to group the sounds based only on their acoustical properties. These experiments were analysed using cluster analysis and the first experiment resulted in two types of sounds, those from inanimate objects and those from live people or animals. These types were further broken into particular thematic subcategories such as sounds from within a house, transportational noises or the sounds from farm animals. The second experiments resulted in the sounds being sorted by acoustical similarities such as the same pitch or rhythmic structure. The use of cluster analysis and the representation of the groupings by dendrogram by Gérard (2004) influenced the selection of these techniques for use in this thesis.

Work from the Sounding Object project can be seen in Chapter 4 as this project's physically based sound models produced the synthetic sounds used in the study. This project also discussed a tentative outline for an incomplete taxonomy of everyday sounds. It was highly influenced by Gaver's (1988) taxonomy and built upon this taxonomy with minor differences with respect to the description of sound events. The taxonomy was further refined as part of the CLOSED project (Houix et al., 2007b) and can be seen in Figure 2.3. The development of this taxonomy, as shown in Figure 2.3, was led by researchers from the University of Verona and their focus on synthesised sound models is reflected in the taxonomy. It uses a bottom-up approach starting with low level sound models such as friction or bubble models. The next level contains basic events or sound textures formed from the low-level sound models such as rubbing or dripping. The third level contains process or temporal patterns formed from one of the basic textures or events such as sliding or splashing. The fourth and last level contains several implementation scenarios. These scenarios used the results of the lower levels. Examples include rubbing glass or footsteps on gravel. The taxonomy differs from Gaver's (1988) as instead of decomposing specific sound events into simpler events in order to understand the basic perceptual attributes, it focuses on using physically based sound models as building blocks. These are used to create compound models to represent higher-level events. The hierarchy in this taxonomy starts with the low level sound events and moves progressively to the most complex events. The differences between naming, positioning, and the absence of some events from one taxonomy, when compared to the another are due to the fact that both are incomplete representations or taxonomies. The low level models correspond to Gaver's basic events, the basic events and textures correspond to the concept of simple events and some aspects of temporal patterning, the derived processes roughly correspond to Gaver's compound events, and the simulation examples are equivalent to Gaver's hybrid events. This taxonomy influenced the direction of this research as it was conceptualised in part during the Sounding Object project and contributed to several studies including the work presented in Chapter 4.

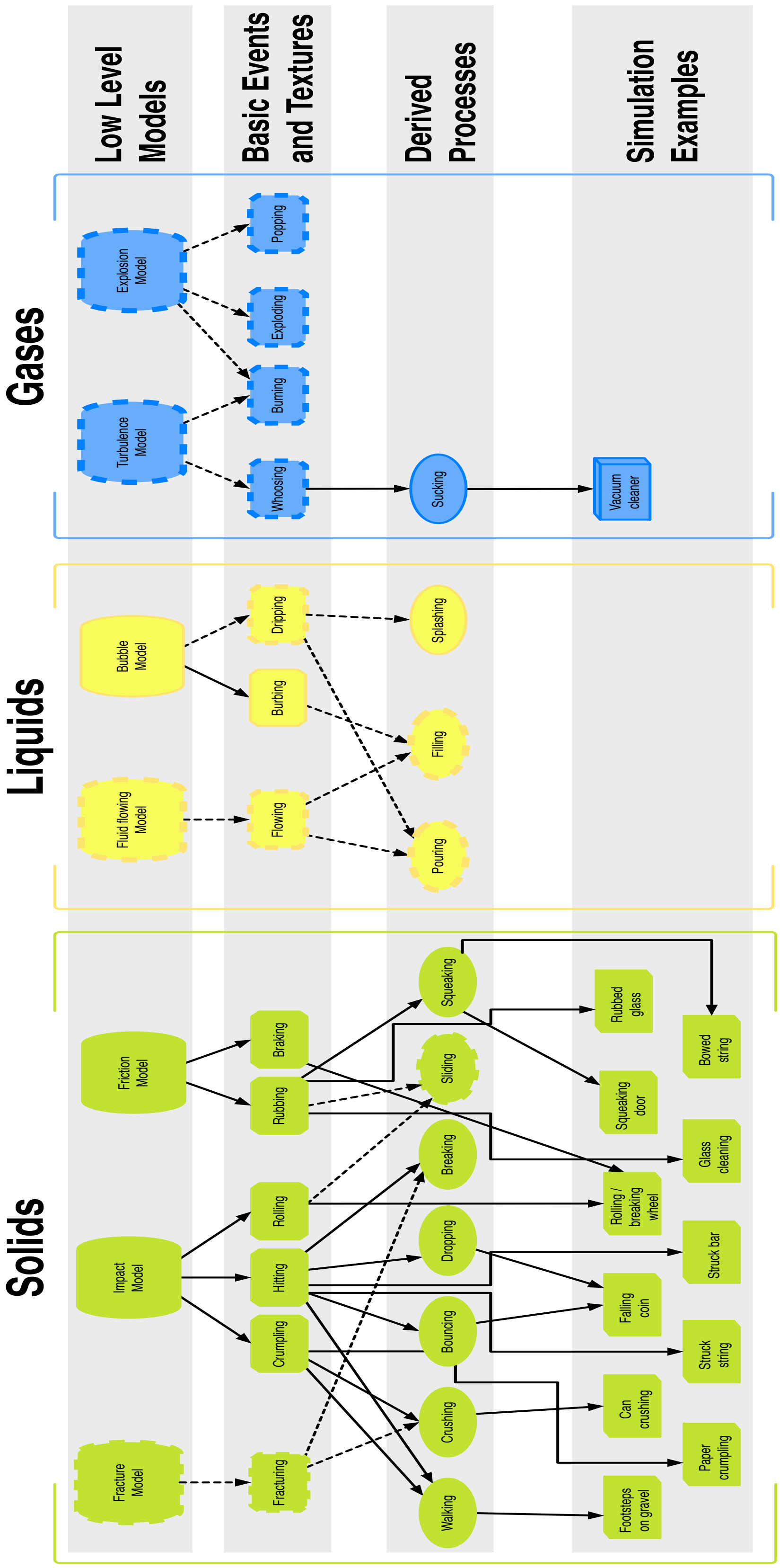


Fig. 2.3: *Sound Object outline taxonomy as discussed in (Houix et al., 2007b).*

The previous taxonomies have shown that listeners group sounds according to various properties or similarities such as event, acoustic, or semantic similarity. The previous studies highlight various types of grouping where the listener used the various properties or similarities:

- Acoustical properties or similarities such as the same rhythm, timbre, or duration.
- Interaction with the same kind of action or movement.
- The same object or sound source is used to group the sounds.
- The sounds fit into the same abstract category or function.
- The sounds happen in the same location or for the same type of event.

The previous studies show that sound identification and source identification have strong relations to the classification of a sound. The only case where event similarity is not taken into account for grouping is where the grouping occurs at the acoustic property level between sounds. This indicates that for many of the grouping classifications there is an implicit knowledge of an action and of an object creating the event in the sound.

The studies presented and discussed in this section have found that when a listener perceives a sound event that the perception of the sound event uses both bottom-up and top-down processing. The auditory attributes of the sound and its context as well as the listener's own knowledge and expectations are at the core of this perception. These studies have produced different theories for the categorisation and for the classification of sounds. These theories are relevant to this thesis as it is critical to have an understanding of how people categorise and classify everyday sounds before attempting to design interfaces using this type of sound. The studies in this section presented several classifications and from these theories the work of the CLOSED project and its hierarchical classification (Houix et al., 2007a), shown in Figure 2.4, was selected as the most appropriate theory for this thesis. This provides a well-defined core based on empirical studies and forms a general framework for the classification of everyday sounds.

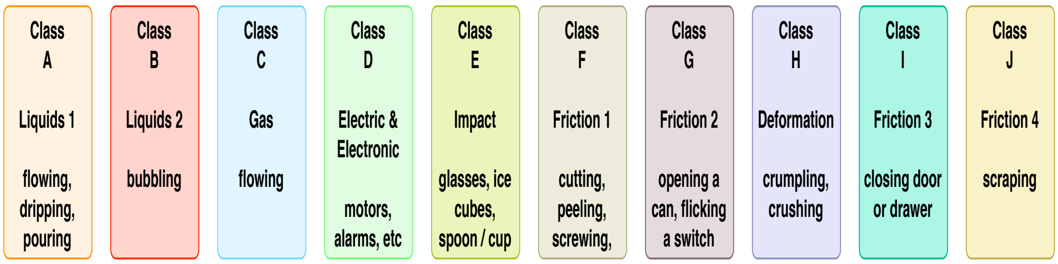


Fig. 2.4: *CLOSED project's hierarchical taxonomy as discussed in (Houix et al., 2007a).*

2.2 Definition of an Auditory Icon

Auditory Icons were defined by Gaver (1997) as "*everyday sounds mapped to computer events by analogy with everyday sound producing events*". These are representations of everyday sounds designed to convey information from computer events (Gaver, 1993c,a, 1994). An example of an Auditory Icon is where the sound of coughing has been used to represent carbon monoxide levels in an airplane (Perry et al., 2007). This approach exploits the fact that people often hear events, rather than sounds. Traditional psychoacoustics (Zwicker and Fastl, 1990) is concerned with the ability to perceive attributes of sounds such as frequencies, spectral contents, amplitudes, etc. Everyday listening contrasts to this approach as it is concerned with listening to the attributes of events in the world, examples would be the speed of a person walking, the number of bounces of a ball, the type of car that is passing, etc. Gaver (1989) has discussed this type of listening and suggest that it used as a basis for Auditory Icons

> "Objects in the computer world should be represented by the objects involved in sound producing events; actions by the interactions that cause sounds and attributes of the system environment ... by attributes of the sonic environment." (1989, p. 71)

In this thesis, the definition for Auditory Icons is that Auditory Icons sound like *everyday sounds* and have an iconic and metaphorical mapping to the system they represent. As Auditory Icons are caricatures of naturally occurring sounds where the source of sound is used as the source of information, Gaver (1988) suggested that where these are used to represent conceptual objects in an Auditory Display they may be more clearly perceived than other sounds. This suggestion has lead to many studies to determine if Auditory Icons are easier to use, more direct, more tangible, and more engaging than symbolic sounds such as Earcons. Earcons (Blattner et al., 1989) are short structured audio messages formed from a "grammar", typically, but not always these are musical motifs. The focus of this thesis is not to explore the comparison between these types of sound; rather the work is aimed at opening new avenue in Auditory Display. Future studies should explore these types of sounds in sole and in hybrid conditions to determine what type of sound (Auditory Icon, Earcon, or Speech) is most suitable for the particular situation. The new avenues from the research in this thesis will help future studies to select the most suitable Auditory Icons for mapping and this may then be compared with other forms of Auditory Display. The design of these mappings for Auditory Icons or for everyday sounds is not a simple process and researcher and designers in this process face many problems. The next section discusses points that need consideration when designing Auditory Icons.

2.3 Viewpoints on the creation of Auditory Icons from everyday sounds

Sound design has long been addressing the issue of how to select and use the appropriate everyday sounds as sound designers are not recreating “real” sounds, rather they are attempting to create the impression of the real event in the mind of a listener. The cultural and physical experience form part of the user’s expectations and contribute to the listener’s mental model. Using both a cultural understanding and natural cognitive mappings learnt from everyday life, a sound designer creates metaphors, balancing the particular sounds within a changing cultural context. Sound designers require methods to these areas, such as the fact that some sounds are less definite than others, or, that they can have a multiplicity of meanings dependent upon the context of the sound. The results from this thesis help in providing details on a number of these problems. An effective design recognises these problems and combines them in a manner that conveys the desired message. An essential part in creating an effective design is easing the identification of a given sound or combination of sounds. People often create stories to explain a sound or set of sounds but these stories may not always be the story/stories intended by the sound designer (Cohen, 1994a). A sound in isolation can be very ambiguous in its meaning but by putting it in context with other sounds, its meaning can be made clearer. In designing with everyday sounds it is very important to establish if sounds are ambiguous and either place them in a contextual relationship with other sounds to clarify their meaning or remove them and select a sound with a clear meaning that still suits the particular context where the sound is being used. In this thesis, sounds are seen as a particular union of events in a soundscape where the acoustic sound, the current state of mind of the listener, the environmental factors of a sound’s recording (if a sound sample) and its playback are all factors. The details of how, where, and when a sound is recorded must also be carefully considered by sound designers. The perspective provided by a recorded “*scene*” offers a window on a real event and documents the event in an unreproduceable sense as it is impossible to recapture that instance or exact sound again. The details of equipment, setup, distance from the event of the microphone, room type, etc. are important to note as these will affect this perspective (Shafiro and Gygi, 2004). The first experiment of this thesis presented in Chapter 4 explores the use of synthetic sounds to convey the same information in a convincing manner where synthesised sounds were designed to maintain the perceptual invariants of the original sounds that they were modelled on. The second experiment explores the questions of identification and confusion while the third experiment explores listener’s classifications of sounds. These studies provide results and methods that could potentially be of use to sound designers. Likewise the factors in sound design need to be considered by those developing new Auditory Displays.

2.3.1 Creating Auditory Icons

Auditory Icons can be created and generated in several different ways. Historically, Auditory Icons have used recordings (samples) of real-world sounds. This allows for high-fidelity exact reproduction of the sound. This may becoming annoying to listeners after a time as it is always the exact same sound they heard with no variety. Capturing many different situations and instances of an event can improve the expressiveness of an interface by allowing multiple different versions or interpretations of an event to be used by the selection of a different recording of the event. With increasing computational power, it has become possible to use synthesised sounds as Auditory Icons. The most recent versions of these synthesised sound can be parametrically control to offer fine grain control over the sound produced. This makes it easy to reproduce a similar but not exact reproduction of a sound, which helps to provide variety in the sound. The parametrical control of the sound also allows for easier mapping of computer data or values to the different perceptual parameters of the sound (e.g. the force of impact or height dropped). The first experiment in this thesis explores a particular kind of synthetic sound, a physically informed sound object, and these introduced in more detail later in this Chapter 2.4.3. This thesis provides methods that caters to both types of Auditory Icons, those based on recordings of real-world sounds and those based upon synthetic sounds such as those created using physically informed sound object models. The type of Auditory Icon, real or synthetic, is one factor in creating an Auditory Icon. There are a number of other factors and lessons from many fields such as sound design, cinematography, electroacoustic music, and ecological thinking that are important in creating a successful Auditory Icon.

The only previous guidelines for researchers or designers wishing to work with Auditory Icons are from Mynatt (1995) who suggested four factors that affected the usability of Auditory Icons. These factors were identifiability, conceptual mapping, physical parameters, and user preference. The specific guidelines from Mynatt (1995, p. 71) are as follows:

1. Choose short sounds that have a wide bandwidth, and where length, intensity, and sampling quality are controlled. The set of sounds should represent the variety and meaning needed for the anticipated design space.

2. Evaluate the identifiability of the auditory cues using free-form answers.

3. Evaluate the learnability of the auditory cues that are not readily identified.

4. Test possible conceptual mappings for the auditory cues.

5. Evaluate possible sets of Auditory Icons for potential problems with masking, discriminability and conflicting mappings

6. Conduct usability studies with interfaces using the derived Auditory Icons.

In addition to these guidelines there was more generalised design advice given by Mynatt (1995, p. 87) such as evaluating the Auditory Icons outside of the interface they were to be used in. Mynatt (1995) stated that Auditory Icons must be designed as "*a cohesive set*" where their semantic mappings are coherent and complementary, otherwise there is a slim chance that the icons will complement each other. The studies in this thesis highlight a number of method, which help designers select a cohesive set of sounds. In the conclusions of this work, the foundations of a new framework are highlighted as part of Section 7.3.1 in Chapter 7. This approach was implicitly used in this thesis to provide a framework for the selection of cohesive sound sets. The work in this thesis uses Mynatt's (1995) second guideline for two of the experiments and elicits text descriptions of Auditory Icons written by participants to evaluate the identifiability of the sounds. These guidelines can help in avoiding many problems such as sound confusion or masking but more work is need to extend the existing guidelines to address the problems and issues faced by concurrent Auditory Icons.

Designing Auditory Icon Mappings

The guidelines from Mynatt (1995) talk about testing the conceptual mappings. However, prior to testing such conceptual mappings an understanding of how the particular sounds used in the mappings are identified and interpreted by listeners is required. Gaver (1989) highlighted the need for understanding how everyday world events are mapped onto events in the computer world. In many cases, it is possible to have direct iconic mappings between real and computer events. These iconic mappings are literal metaphors of real events and are constrained as they must represent a real event, sight, or sound. Gaver (1989) realised this limitation and discussed the use of foley / sound effects to create sound for an event, which has no real world counterpart such as the formatting of a hard drive on a computer. He further discussed the use of source metaphors, which use analogues of a real sound producing event to convey the mapping of the auditory icon. In the SonicFinder (Gaver, 1989), an example of a sound analogue is where the sound of pouring a liquid into a container is used to represent the progress of an operation. The three types of mapping suggested by Gaver (1989) for auditory icon all require information about the mappings and metaphors the listeners attribute to the sounds and the methods in this thesis help provide approaches to gather this type of information.

Film sound

Adding more and more sounds to an auditory design is not always the best approach. Dense, layered soundtracks are often used in movies but these are accompanied by linked visuals that help in the interpretation of the sounds. Chion discusses the theory of this layering with respect to film where as many layers as desired can be added to a soundtrack but the layers are situated at different hierarchical narrative levels such as background music or synch dialogue (Chion, 1994). In an Auditory Display careful attention must be given to the layering and sequencing of elements in the soundscape. This layering of elements can be informed by knowledge from various fields in science and engineering such as psychoacoustics, acoustical engineering and audio engineering to supply methods, techniques and concepts to assist in the layering process. The methods from these disciplines often focus on the sound as signals and the transfer of energy. The work from these fields does not answer all the questions for sound designers so other approaches need to be considered.

Electroacoustic music

One such complementary approach comes from the field of electroacoustic music. It provides additional insights for Auditory Display design by asking questions about the listener's interpretation of everyday sounds. Windsor (1997) highlights the fact that "... musical and everyday sounds are merely labels for how we use or hear sounds rather than epistemological categories"(Windsor, 1997, p. 77). In later work, Windsor (2000) talks about the "*mutual relationship*" between listeners and the environment and how the listener explores the acoustic structures to build a meaningful interpretation of these structures. Exploring the meaningful interpretation of these structures is difficult. This thesis proposes the repertory grid as detailed in Chapter 6 as one method for addressing this. These interpretations are a "best guess" by the listener and are "part-based upon the predictability of the natural environment, part-based upon the predictability of our cultural environment" (Windsor, 1997, p. 80). This thesis explores methods to investigate a listener's interpretation of everyday sound categories. The work of Windsor indicates a note of caution, as cultural factors are dependent upon listener's prior experiences.

Gibson's theory of perception and related concepts

The view of sound in this thesis, and in particular on the importance of source identification and of sound identification has been influenced by many theories. Gibson's theory of perception (1979) formed a core part of Gaver's reasoning when creating the first Auditory Icons. An important aspect of this theory is its consideration of how information is picked up from the world. It proposes two ways of listening to the information provided by the world. The first way looks at information as having a mutual relationship with action, the heard sound is perceived in terms of how a person can act towards it. This relationship between the sound and the world is its affordance. The sound itself allows for the perception of events and of objects, that afford particular courses of action in the world. These affordances are not fixed but change with context (Gaver, 1993c). An example would be where a car mechanic hearing the sounds of a faulty engine would perceive a set of actions (e.g. replace the faulty spark plug that is causing the misfiring of the engine) that may not be the same set as those perceived by a lay person in the same situation. The second way considers information where it is perceived for its own sake as a 'sensation'. In particular, the first way of listening, where the relationship with action is particularly noted, and influenced the research in this thesis. The close ties between objects and actions in sounds are explored by the methods in this thesis. Several researchers have noted the dominance (Gaver, 1993a,c, Vanderveer, 1979) of 'causal listening' or listening for the source of the event. Kendall noted that

> "In everyday life, sound events arise from action, in fact, from the transfer of energy to a sounding object. The auditory system provides us with perceptual characterizations of the energy transfer and of the internal structure of the objects involved. Early in childhood one learns to recognize the occurrence of sound events and to relate them to physical events" (Kendall, 1991, p. 71)

Visual perception of actions

The core essence of both Gibson and Windsor is that "sounds are intimately tied to action, whether natural, human or artefactual" (Windsor, 1997, p. 81). Studies in the visual perception of actions by Runeson and Frykholm (1983), Runeson (1989) suggest that there may be action-related representational structures serving the production and the perception of both actions and action effects. An important concept from this visual perception research by Runeson (1989) is that of the "*incomplete invariant*" occurs when observers in a situation only possess a limited or subset of the critical information necessary to solve the problem or to decide the particular meaning for the given situation. It is important to note that the events themselves afford little information unless they are related to a particular context or contexts provided by the listener and the environment as in situations where there is "*incomplete invariants*", the organism's perceptual system will "hunt" for meaning (Gibson, 1966, p. 303-304), and it is from a social or cultural meaning that additional information can be used to pick up or more of the available affordances from the environment. The third experiment in this thesis in Chapter 6 elicits attributes for the description of everyday sounds to help in determining the subset of information the listener uses to decide on the particular meaning for the sound. Windsor (2004) suggests a broader definition of affordances, where objects and event may afford different things depending on the needs and capability for actions of the perceiving organism (Heft, 1989), where the affordances are both culturally relevant as well as open to social mediation (Noble, 1991, Costall and Still, 1989, Heft, 1989) and where the affordances result from a mutual relationship between the object/event and the organism (Heft, 1989). Windsor (2004) suggests that this broader definition should span all the available contexts of human environments building upon earlier work by Sanders (1997) who states that

> "... affordances are opportunities for action in the environment of an organism, the opportunities in question include everything the organism can do, and the environment includes the entire realm of potential activity for that organism ..." (1997, p. 14)

This wider definition allows for associations to be made between events and cultural or social affordances. Once made, these affordances are available for discovery in the environment. As the natural or cultural environments change, they will require adaption by a process of learning and Gibson (1966, p. 285) has previously stated that "*learning is vital to the perception of affordances*". These ideas influenced the lens through which the experimental results from the thesis were viewed. Addressing these open questions in the field of ecological psychology is not the goal of this thesis. Ecological psychology did provide many ideas that were useful in providing a wider view of the topics and research questions addressed by this thesis.

Representative design or ecological validity

Another important concept that helped in framing this research was that of "*representative design*" as proposed by Brunswik (1956). "*Representative design*" as a theory seeks to describe psychological processes. These processes are seen as being adapted to the properties of an environment. Representative design uses either the random sampling of stimuli from the environment being explored or the creation of stimuli in which the properties of the environment being explored have been preserved. This is often incorrectly defined in the existing literature as the less defined concept of "*ecological validity*". Brunswik (1956) defined "*ecological validity*" as referring to the validity of the cue (i.e., the perceptual variable) in predicting the outcome state of the particular environment. It was Egon Brunswik's work on this concept that influenced Gibson (2001). The work on "*ecological validity*" has clearly pointed to the distinction between the real world and a controlled laboratory setting. This is an important reminder to designers that while experimental laboratory based results may provide one result, the real result for designers is found in the real world where the Auditory Display is to be used. Brunswik stated that

> "There is little technical basis for telling whether a given experiment is an ecological normal, located in the midst of a crowd of natural instances, or whether

> it is more like a bearded lady at the fringes of reality, or perhaps like a mere homunculus of the laboratory out in the blank." (1955, p. 204)

The key idea is that the an organism has evolved to respond to meaningful stimuli in a meaningful way. As part of this thesis, everyday sounds were used to satisfy this concept of "representative design" and to avoid Brunswik's "bearded lady at the fringes of reality". Brunswik was referring to the signal - masker approach where the use of a background of stationary random noise and one relevant sound are used as the stimuli of an experiment. According to Brunswick concept, this is not a "representative design" for the stimuli, or in the case of this thesis, the everyday sounds that people have adapted and evolved to. It is important to point out that the signal - masker approach in psychophysical research is useful for exploring the fundamental abilities of the auditory system. This approach is not suitable for providing an answer to everyday sound perception. A more useful approach to answer this question is the "ecological" approach of Gibson. Gygi and Shafiro (2007) has described this approach of auditory perception as a process of "tuning in" to the perceptually relevant properties of the stimulus. Brunswik (1952) used an approach that explored the variables at the analysis stage rather than the design stage of research using statistical methods such as partial correlation. The experiments in this thesis use various methods including several similar statistical methods to provide additional insight on the perceptual processing of meaningful sounds. A key difference between Gibsonian and Brunswikian thinking is that Brunswik saw limitations to the information available via environmental cues, while Gibson believed the environment afforded complete information to a person. Taking a Gibsonian approach to the analysis of perception requires that one conduct a simultaneous appraisal of the environment and of the information that it affords. This thesis takes a lighter weight approach than either of the experimental approaches available to these perceptual theories. This work attempts to capture as much information about the sounds, environment, and the listener as possible while minimising the overhead necessary in using these methods. The methods in this thesis aim to provide meaningful and useful information without the overheads found in a traditional strict laboratory study such as those used in psychoacoustics or in the cognitive psychology.

Definitions of embodied

Ideas from *embodied cognition* such as Lakoff's and Johnson's (1999) idea of *embodied realism*, the work by Varela et al. (1991) and their definitions of *embodied* and *action*, the research on a two level category-specific organisation of concepts (Rosch, 1977, 1975, 1978) helped to frame a number of relevant concepts for this thesis. These ideas influenced the structuring and reasoning behind the research questions and how they explored in this thesis. The support of ideas from *embodied cognition* has been strengthened by research in neurophysiology which focused on the role of action in cognitive processes. It was specifically focused on the perceptual recognition of objects, actions and their conceptual categorisation (Garbarini and Adenzato, 2004). Garbarini discusses a viewpoint combining these concepts where the mind is seen

> "Acting in the world, interacting with objects and individuals in it, representing the world, perceiving it, categorizing it, and understanding its significance are perhaps simply different levels of the same relational link that exists between organisms and the local environments in which they operate, think, and live" (Garbarini and Adenzato, 2004, p. P105)

Garbarini further refines the idea of *mental representation* as being

> "intrinsically linked to the sphere of action and is expressible in the same terms that control it ...(the) mental representation in which the experience is "constructed" on the base of categories, which are no longer theoretical, but pragmatic, deriving from the dynamic interaction of the organism with its adaptive environment" (Garbarini and Adenzato, 2004, p. P106)

The studies in this thesis used two categories, objects and actions, as the basis for the classification of the everyday sounds. The reasoning behind investigating these classification categories was informed by research in soundscape classification, interactive sonification and embodied cognition. McGregor et al. (2006) have found that in their soundscape classification studies participants included both source (object) and actions in 100% of their descriptions of a soundscape. The spatial dimensions of the sources were the next most common item included in their descriptions with 88% of descriptions including this aspect. These studies point that sources (objects) and actions are the most salient items in a soundscape for listeners.

The idea of action and source was further inspired by research (Fernström, Brazil and Bannon, 2005), that helped users to experience the *flow* (Csikszentmihalyi, 2000) of a complex yet continuous interaction in an Auditory Display to create a pseudohaptic experience for a gesture-based device. This work relates to the definition of interactive sonification from Hermann and Hunt (2005, p. 20) where sound is the integral part of "*a tightly closed*" loop where the "*auditory signal*" provides information on either the particular data being analysed or on the actual interaction itself. The sound can even be used for "*refining the activity*" or interaction in a dynamic fashion. This type of interaction has highlighted the potential for Auditory Displays, which concentrate on mapping human activity to actions. In many Auditory Displays prior to this, activity or events had been mapped to objects or sources rather than actions. This work when, considered with the results of McGregor et al. (2006) from soundscape classification studies highlight the possibility for using action mappings. This thesis expands upon the limited research on mapping of activity to actions in Auditory Display by exploring the more fundamental question of how action classification of a sound can affect its identification.

The choice of action as a category was further influenced by the action-related representational structures for perception and of the "*incomplete invariant*" as suggested by Runeson (1983, 1989). Both actionhood and objecthood are important concepts in many previous theories varying from Gibson on *affordances* (1979), Lakoff and Johnson (1999) on *embodied realism*, Varela *et al.* on their definitions of *embodiment* and *action* (1991), to the work of Garbarini and Adenzato (2004).

The identification of the change and persistent features of events that are ecologically significant as the produced structural information can afford further perception or action. Event perception is the pick up of the invariant properties of events that detail the static and dynamic features of a environment to a organism, as events contain an affordance structure containing two types of invariants. Work by Ballas and Howard (1980, 1982, 1983) has investigated various abilities of people including the identification of sources in complex and noisy signals, the influence of source classification, and the effect of semantic interpretation of sounds on the syntactical parsing of soundscapes. Their findings would indicate that the ability to identify a sound is related to the ability to classify acoustic features. Their methods are discussed in Chapter 5 where everyday sounds, and in particular for situations or interfaces using many simultaneous sounds, are explored.

Temporal and Spectral concepts

It is useful to review the prior research on sound structure as an understanding of the importance of temporal and of spectral content helps in understanding these points and their place within the research field of Auditory Display. Temporal structure is also important as shown by the work of Warren and Verbrugge (1984) whose breaking and bouncing experiments on glass found that subjects made the breaking or bouncing distinction based upon on temporal information rather than on spectral information. Macrotemporal properties of sounds are also important for identification, e.g. the sound of a single isolated footstep can be heard as a book being dropped on a table, while if several footsteps are heard there is seldom any doubt about the source. Gygi (2001) found that when all spectral (timbral) content was removed, the macro-temporal structures of the sounds still afforded identification (22-46%). This suggests that the spectral information in a sound specifies the interactions of the materials as they relate to a real physical event. Gaver (1993a, 1993c) proposed that both real and synthetic events can be perceived in terms of their probable causation, regardless of whether this causation is real or modelled. Work in synthesised sounds (Rocchesso et al., 2003, Rocchesso, 2004) has shown that a recording of a real event is not required for the perception of a physical cause once a particular event's invariant structures can be modelled and controlled. This was one of the motivations for the first experiment in Chapter 4 where synthetic event sounds were used to determine if proper modelling of an event's invariant structures could still convey the same information with regard to the perception of the physical cause. In the case of the synthesised sounds used in the experiment in Chapter 4, it was found that synthesised sounds could still convey this type of information. Earlier research (Ballas, 1993, Vanderveer, 1979, Gaver, 1988) provided guidance and outlined how sound and event invariance was previously explored. This informed the choice of methods used. This question is explored in greater detail by the experiment in Chapter 5.

Source and sound identification are two areas that this thesis provides additional depth to the existing work in the field, in particular with the work in Chapter 5. Identification is a key element within the methods provided by this thesis however it should be noted that sometimes it may be useful to have ambiguous sounds. Work by Gaver et al. (2003) has highlighted the role of ambiguity in developing "*engaging and thought provoking*" artefacts. This idea can equally be applied to sound in Auditory Displays. This work (Gaver et al., 2003) has provided a taxonomy of methods and relates these to the design of interactive systems. In these cases, the methods from Chapter 5 are useful for determining if there is a masking or identification problem where multiple ambiguous sounds are used. The points raised in this section are important in establishing in the mind of a designer what issues they may be faced with in the process of designing an Auditory Display. Creating an Auditory Icon as a representation of an everyday sound is relatively easy, however as the previous paragraphs have shown, the design or composition of an Auditory Icon such that it creates the desired impression in the mind of a listener is a more complex issue.

The next section gives a history of Auditory Icons in interfaces where a number of Auditory Displays using Auditory Icons are discussed. These systems show how previous designers created systems using Auditory Icons and provides a foundation for understanding how Auditory Icons have been used in the past. These applications all suffer from a lack of empirical studies and most had only a brief evaluation, if any, conducted. These studies have been typically conducted in the late or post stages of development. The experiments in this thesis provide a set of methods aimed at researchers and designers of Auditory Displays. A by-product of this research is that if researchers or designers applied these methods then better Auditory Displays would be created. This would be due to the fact that they are less dependent on exceptional luck or exceptional design skill due to methods providing more detailed information on the Auditory Icons being used. This will facilitate researchers or designers in making more informed design choices about the suitability and use of the Auditory Icons within their Auditory Display. The review of Auditory Icons in interfaces highlights the major design choices taken in prior applications. These choices were due to a variety of factors from the Auditory Display technology to domain specific factors, however an awareness of the rationale behind these choices can help in future designs.

2.4 A History of Auditory Icons in User Interfaces

Auditory Icons have been used across many domains and in different interfaces, from the computer desktop to the factory floor. The previous systems using Auditory Icons have been reviewed in this thesis and divided into three generations. Each generation was informed by the earlier generations. The first generation of work is typified by systems such as SonicFinder (Gaver, 1989) and SoundShark (Gaver and Smith, 1990). They added Auditory Icons to existing desktop-based applications. The second generation of work are exemplified by ARKola (Gaver et al., 1991) or Varèse (Albers, 1994). These systems build on the first generation Auditory Displays by providing more complex soundscapes and addressing topics such as awareness and collaboration. These displays were still typically limited to a single desktop machine. It is in these second generation of systems where multiple parameterised Auditory Icons were first used. The third generation of work is best seen in the approach of Cook (2002b) and of the Sounding Object project (Rocchesso and Fontana, 2003) to investigate physically based sound synthesis for Auditory Icons within a ecological paradigm. These Auditory Displays used parametrically controlled sounds and moved to more technically complex platforms such as multiple machines with web server monitoring (Malandrino et al., 2003, Gilfix and Crouch, 2000) or to ubiquitous and wearable computing (Rocchesso et al., 2003).

Understanding how Auditory Icons were used in these systems can help in determining what problems must be addressed when designing these types of Auditory Displays. The results presented in this thesis address several of these questions by providing a deeper insight into specific areas such as sound confusion and sound identification. A problem of all the previous Auditory Icon based Auditory Displays is that they focus solely on the areas of the design of the system architecture and of its sound design. Many other topics were not explored and areas such as the identification of the Auditory Icons used or the confusion of Auditory Icons have not been previously studied. Evaluations when carried out, were often of a short exploratory nature and did not conduct systematic or in-depth studies. The results of these evaluations were brief and specific to the particular system such that these results were not applicable to the design of other Auditory Displays. The systems presented in the following paragraphs did not carry out any evaluations that provided results, which could be more broadly generalised. The lack of documented reflections on these systems is due both to the prototype nature of the systems and the lack of deep evaluations. This type of approach is problematic; as even today as many of the Auditory Displays developed reach the functional prototype stage and only then undergo a preliminary evaluation. This lack of further work is problematic for other researchers and designers who wish to understand the reasons and the design choices that lead to the success or the failure of the particular system.

2.4.1 The first generation

This section discusses SonicFinder (Gaver, 1989) and SoundShark (Gaver and Smith, 1990). These are typical examples of the first generation of Auditory Icon Auditory Displays. These Auditory Displays shared a number of common aspects with regard to the use of Auditory Icons. Direct iconic relationships reflected user interactions and process activity. In these early systems, hardware was more expensive, disk space was at a premium and the modern sound card had not yet been integrated as a standard computer peripheral device. A successful system required either using a large amount of disk and of memory space, or, external hardware such as samplers controlled via MIDI. These options limited the widespread commercial deployment of these Auditory Displays but this work did succeed in highlighting the possibilities for Auditory Icons in the user interface. These systems listed in the next paragraphs can be seen as some of the first prototypes in the field of Auditory Display.

The SonicFinder

The SonicFinder was a novel Auditory Display that mapped qualities and quantities of events occurring within a computer to perceptible attributes of sounds. This allowed users to intuitively map the everyday sounds to the computer events as the mapping exploited the causal structure around which everyday listening is based. This interface was the first specifically designed system to incorporated Auditory Icons; it was designed as an extension to the existing Finder application in Apple's Macintosh operating system. The Finder application is the file manager in the Macintosh platform and is used to organise, manipulate, create and delete files. Extending Finder to become SonicFinder (Gaver, 1989) was done by adding sampled sounds at appropriate points to play the sounds according to the attributes of the relevant events. The actions which had a related Auditory Icon in SonicFinder included: selecting, dragging, copying files; opening/closing folders; selecting, scrolling, and resizing windows; and the dropping of files into and the emptying of the trashcan (this holds the delete files on the Macintosh platform). The complete list of mappings for the SonicFinder is shown in Table 2.1. Gaver (1989) claimed that the intuitive mappings between the events and the Auditory Icons resulted in a much higher sense of direct engagement from those who used the interface. The Auditory Icons used were based on literal mappings between the computer's events and the Auditory Icons, which reflected sounds in the real world. It aimed at building upon listener's existing abilities such as the skill of everyday listening while providing information that could be understood with minimal training.

Event to Sound Mappings for the SonicFinder	
Computer Finder Event	***Auditory Icon***
Objects	
Selection Type (file, application, folder, disk, trash) Size	*Hitting Sound* Sound Source (wood, metal, etc.) Frequency
Opening Size of opened object	*Whooshing Sound* Frequency
Dragging Size Location (window or desk) Possible Drop-In ?	*Scraping Sound* Frequency Sound type (bandwidth) Disk, folder, or trashcan selection sound
Drop-In Amount in destination	*Noise of object landing* Frequency
Copying Amount completed	*Pouring sound* Frequency
Windows	
Selection	*Clink*
Dragging	*Scraping*
Growing Window size	*Clink* Frequency
Scrolling Underlying surface size	*Tick sound* Frequency
Trashcan	
Drop-in	*Crash*
Empty	*Crunch*

Table 2.1: The mappings from computer events to Auditory Icons in the SonicFinder (1989).

There were two types of problem encountered when creating this type of mapping, the first was developing Auditory Icons for events with counterparts in the everyday world and the second was providing a sound for a computer event that does not have any informational sound associated with it. The first problem is illustrated by windows in the user interface of the computer, in particular what is the sound of selecting or growing a virtual window. Real windows open slowly whilst computer based windows typically zoom in or pop out very quickly. This meant that the real sounds of a window would be an inappropriate mapping. The SonicFinder (Gaver, 1989) used a '*whooshing*' sound and highlighted the potential for sound effects or Foley sounds as alternatives when mappings based on real sounds are inappropriate. The difficulty in finding a sound to relate to an event, which has no informative sound can be seen in the ''*copying*' Auditory Icon, where a pouring sound was used. The rationale for using this type of sound was that it provide an analogical mapping. In the SonicFinder (Gaver, 1989), the link was between pouring in the real world and copying on the computer. The '*pouring*' fits a metaphorical mapping between events and helps provide useful information about the closeness to the end of the task, which is highly salient to users.

The mapping of events to sounds and the related problems highlight the difficulties in creating and developing Auditory Icons that accurately provide salient information to users of Auditory Displays. In this thesis, three broad issues and related methods are discussed. The mapping of perceptual properties to sound is explored in Chapter 4. This work provides an approach to explore if sounds are conveying the desired mapping to listeners. The issue of identification and the confusion of everyday sounds is investigated in Chapter 5. Identification and confusion are important factors in determining the success of an Auditory Icon's mapping. An understanding of listeners' perceptual spaces, how they relate attributes to one another, and the vocabulary or metaphors they use to describe them is explored in Chapter 6. This information can help designers in understanding existing mappings and in creating new mappings for Auditory Icons.

The Auditory Icons in the SonicFinder (Gaver, 1989) were used to provide feedback and information about the interactions in an intuitive and informative way, in support of the visual GUI. Whilst this application was found to be useful and intuitive by many users, it was intended as a prototype to outline a potential future version of the Finder application. A number of constraints mitigated against it success at the time including the size constraints on hard disk sizes of the machines of the time. It did succeed in highlighting the potential of auditory interfaces and the real-world challenges that face those building and designing them.

SoundShark

Following on from the design of the SonicFinder, Gaver worked with a colleague to create a new system, SoundShark. Gaver and Smith (1990) expanded upon a multiprocessing, collaborative environment, SharedArk (Smith, 1988) to add Auditory Icons to create this new system. SharedArk was a collaborative Auditory Display designed as a virtual physics laboratory for distance education (Smith, 1988). Auditory Icons were added to this system to create SoundShark. It used Auditory Icons to indicate user interactions, ongoing processes and modes, to help with navigation, and to provide information about other users. User actions were parameterised to indicate attributes such as object size and this was incorporated into the Auditory Icons. Sounds were used to indicate the nature and activity of ongoing processes even when not within the window or view on the screen. This aided in co-ordination as even when collaborators could not see each other they could still hear each other. Modes within the system were indicated by low volume background sounds. The distance between the user's cursor or hand in the system and the source of a sound was indicated by the relative loudness combined with low-pass filtering of the sound. This was one of the earliest uses of "auditory land-marking" where relative loudness was used to indicate the distance to object being represented by the Auditory Icon. These landmarks functioned as repetitive sounds and were used as orientation aids. This Auditory Display used simultaneous sounds via an external MIDI controller and demonstrated ideas that would require the advent of cheap audio cards to see widespread availability.

2.4.2 The second generation

The second generation of Auditory Icon systems used either parameterised Auditory Icons (Gaver, 1993b) or state based Auditory Icons (Albers, 1994) to convey information about a larger number of variables in a dynamic and simultaneous manner. Fitch and Kramer's Auditory Display (1994), is a typical example of this kind of system. The display was designed as an aid for anaesthesiologists; two parameterised Auditory Icons were used to convey information to the anaesthesiologist about a patient's vital data. The patient's heart rate controlled the rate of a heart-like sound with the patient's systolic blood pressure indicated by the pitch of this heart sound. It used a set of two parameterised Auditory Icons mapped to seven physiological variables and one traditional high-pitched alarm sound for one physiological variable. This system and others in the second generation expanded upon the work of previous Auditory Icon systems and applied them in more complex situations using parameterised Auditory Icons.

ARKola

After the work on SoundShark, Gaver (1991) explored the use of Auditory Icons in conveying information about user-initiated events, processes and modes, and about location within a complex environment. Building upon SoundShark, a model of a soft drink plant called the ARKola bottling factory (Gaver et al., 1991) was created to explore these ideas. The bottling plant simulation consisted of a single assembly line consisted of 9 machines. These machines represented the various processes in the system and included cooking, bottling, the provision of supplies, and financial tracking of the various processes. During the design of the simulation, a deliberate decision was made to ensure that the simulation was larger than a single screen so that participants could only see approximately half of the icons or represented machines on the computer screen at any given time. The scenario was to observe and repair the factory as it ran. The simulation incorporated breakdown of the machines within the simulation and users were expected to "repair" these malfunctioning machines in the assembly line. Each machine had a unique Auditory Icon to indicate its function; in addition the rate of each machine was indicated by the repetition rate of the sounds it made, while problems with machines were indicated by a variety of alarm sounds such as breaking glass and overflowing liquid, amongst others. Auditory Icons were designed for each machine to indicate its status over time and also to reflect the semantics of the particular machine. Examples include a '*whooshing*' sound that represented the heating machine or the sound of '*clanking bottles*', which represented the bottling machine. The interface for the ARKola system is shown in Figure 2.5.

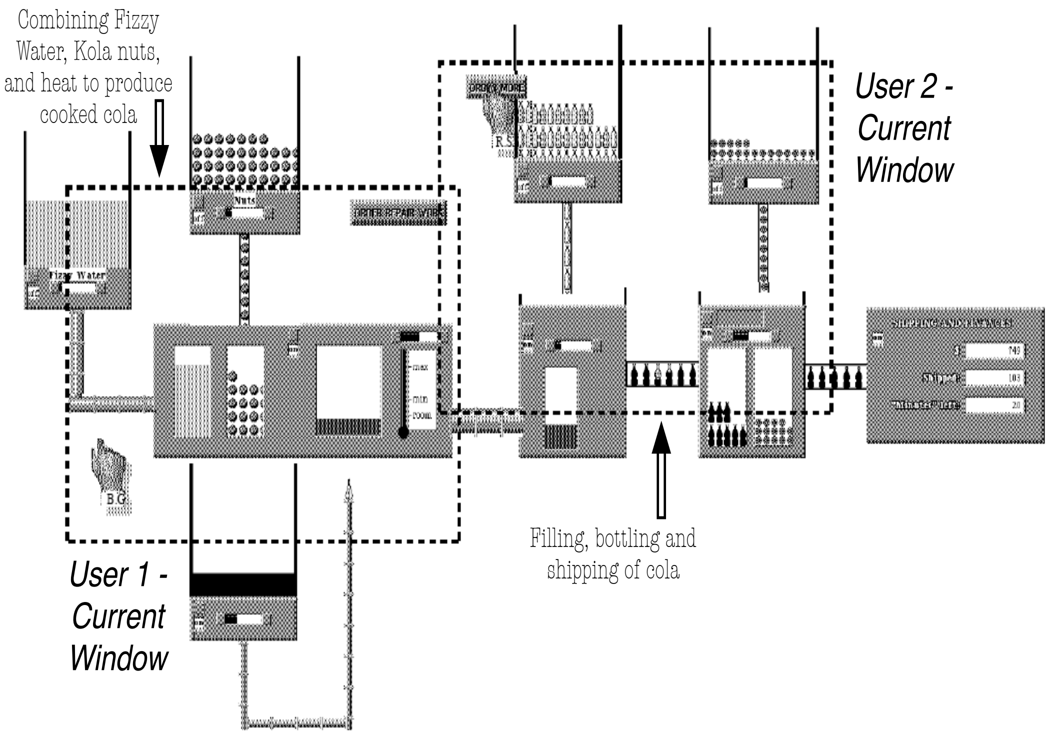


Fig. 2.5: *The ARKola (Gaver et al., 1991) bottling plant simulation. The rectangles indicate the extent of the plant that each user sees at a given moment.*

This Auditory Display used up to 14 simultaneous sounds to reflect the machines and processes occurring within the bottling plant. Firstly, temporally complex sounds were used to maximise discriminability and the sounds where designed to be semantically related to the events they represented. Masking was avoided by selecting sounds spread fairly evenly in frequency and the sounds were not played continuously but rather played in a repetitive stream with gaps between the repetitions to improve the chance of hearing the other sounds. This approach created a number of repetitive streams of Auditory Icons, which reflected the parameters and state of the processes and machines. This produced a complex but dynamic and changing sound scene allowing users to perceive the plant as an integrated complex process. The gaps used were an *ad hoc* solution to prevent sounds overlapping and did not actively consider these sounds in relation to one another. The work in Chapter 5 presents a method that solves this type of problem and helps designers ensure that the sounds they select will complement each other in a complex soundscape.

Sound was found to help users track the activity, rate, and functioning of machines within the plant. Studies were carried out to explore the use of this simulation as a collaboration process between the two participants, each focused on half of the processes. The results found that sound lead to a greater collaboration between partners as they could directly hear the status of their partner's half of the plant. It helped in increasing both the realism and engagement of users with the simulation while providing foreground information visually and background information using sound.

Varèse

The Varèse system (Albers, 1994) was a typical second generation implementation and used a parameterised Auditory Icon approach. The Varèse system was designed for satellite-ground control where operators were only able to communication to a satellite for 10 minutes, 4 times a day. In these designated time slots or windows, operators had to determine the state of 6 sub-systems of the satellite and the overall operational state of the satellite. Each sub-system had three operational states, normal, warning or critical. The data from the satellite's sub-systems consisted of time-series data and this was represented by 6 continuous sounds conveying both the current state and proximity to the next operational state for each subsystem. For example, the communications system was represented by a stream of Morse code and the power system being represented by a car engine. The parameterisation of the Auditory Icons was used to increase the rate of the Auditory Icon so that changes from normal to warning states were presented in a increasing sense of urgency to the operators. In the case of the power system, the normal operational state was represented by the sound of a smoothly running engine, a spluttering engine indicated the warning operational state, and the sound of a dying or seizing engine represented the critical operational state. The published work on this Auditory Display is a typical example of many Auditory Display papers where the mappings were discussed but there was no detailed commentary or reflection on the design of the Auditory Display. In the case of the Varèse system (Albers, 1994), it is simply reported that the system assisted operators in fault detection by isolating the information from the various satellite sub-systems. These comments were not substantiated by any studies or evaluation data. This lack of data or reflection increases the difficult for those who wish to learn from the successes and from the failures of past Auditory Displays. The lack of evaluations has left many open questions in the field of Auditory Displays and with the growing acknowledgement about the important of interface design, it can be hoped that this type of evaluation will be conducted in current and future Auditory Displays.

Audio Aura

The Audio Aura (Mynatt et al., 1998) was designed as a soundscape that used background auditory cues to provide peripheral awareness. These cues were tied to three scenarios. These were email quantity, colleague presence and a background continuous "group pulse" representing the group's activity level. A distinct set of auditory cues where created for each scenario. The mappings for email and for "group pulse" are shown in Table 2.2. The Audio Aura system was not a strict Auditory Icon only display as it offered four possibilities for Auditory Display with speech, Auditory Icon, Earcon or a hybrid model of all three. The Auditory Icon soundscape was modelled on a beach metaphor where group activity was mapped to wave activity, email amount mapped to amount of animal calls, with particular email senders being mapped to different birds or seals and the colleague presence was mapped to buoy bells. The mapping in this Auditory Display further highlights a number of areas that were designed in an *ad hoc* fashion were perceptual properties such as the mapping of quantity to sound, the use of concurrent everyday sounds, and the creation of new mappings for Auditory Icons. Methods for addressing these particular issues are presented in Chapters 4, 5, and 6. This implementation is similar to many of the second generation systems but it was one of the first Auditory Displays that took the Auditory Display away from the computer interface and made it part of the environment.

Peep Network Auraliser

Peep (Gilfix and Crouch, 2000) was designed as a network monitoring system using Auditory Icons that represented network events. It differs from ARKola (Gaver et al., 1991) in that it uses a continuous sonic environment to allow for a real time approximation of network status. It used three types of sound representation to achieve this, the first represent single occurrence events, the second represented states of the network and the last represented regular but changeable states (such as the presence of particular machines). Single occurrence events were represented by single peeps or chirps, states were represented by changing the type, volume or stereo position of an ongoing background and heartbeats were represented

Email Quality				
	Sound Effects	Music	Voice	Rich
Nothing new	A single gull cry	high short bell melody rising pitch at end	'You have no email'	Same as SFX: single gull cry
1-5 new messages	A couple of gull cries	similar bell melody but longer and falling at the end	'You have **N** new messages'	Same as SFX: A couple of gull cries
5-15 new messages	A few different gulls cries	similar bell melody but lower and longer	'You have **N** new messages'	Same as SFX: A few different gulls
15+ new messages	Gulls squabbling, making a racket	similar bell melody but longest and falling at end	'You have **N** new messages'	Same as SFX: Gulls squabbling
Group Pulse				
	Sound Effects	Music	Voice	Rich
Low Activity	Distant Surf	Vibe	None	Mix of Vibe and Surf
Medium Activity	Closer Waves	Same Vibe with added sample at lower pitch	None	Mix of Vibe and Closer Waves
High Activity	Closer more active Waves	As above, three Vibes at three pitches and rhythms	None	Mix of Vibe and Closer Waves, but more active

Table 2.2*:* The mappings from computer events to Auditory Icons in the Audio Aura (1998).

by playing a sound at varying intervals, such as by altering the frequency of cricket chirps. Events were single Auditory Icons and selected to be short and staccato in nature. In the case of this thesis, animal and bird vocalisations are counted as everyday sounds and are treated as Auditory Icons, however there may be cause to classify them as in a hybrid category of both Earcons and Auditory Icons due to the musical aspects to bird song, etc. An example was how email events were divided into incoming and outgoing emails, and these were represented by the sounds of two conversing birds to create a sequence of call and response. State sounds corresponded to measures or the magnitude of something such as the load average or number of users on a machine. These were represented by continuous stream of background sounds, like a waterfall or wind and were scaled to an ordinal measure, rated from quiet to loud. Regular changeable states (Heartbeats) were represented by regularly occurring sounds such as crickets chirping at night. Network load was represented in this category with intermittent chirps meaning low load to a chorus meaning high load. Heartbeats could also be used to report server checks such as ping to see if a machine is present on the network. Peep offered a variety of sound themes to allow a user to select a particular soundscape or even to create their own personalised themes to reflect their tastes and particular network situations. Peep was designed by and for system administrators with no attention being given to evaluations or design reflections.

2.4.3 The third generation

At this point in the development of Auditory Icons, there was a growing awareness of their possibilities with interfaces using them but they were not longer the single key feature of the interface. It is also at this point that the first synthesised Auditory Icons functioning in real time are more widely available. Real time systems are defined as those where the total correctness of an operation such as the synthesis of an Auditory Icon depends not only on its logical correctness but also upon the time in which it is performed. Embedded systems such as a car engine control systems, avionics, heart pacemakers, factory automation systems, and portable computing devices often have aspects which require that an event is reacted to within a specific and strict timeframe. The requirements for the synthesis of Auditory Icons in real time has resulted in several synthesis techniques each designed with the aim of creating realistic, high quality realtime sounds. One of the most promising approaches is that of physically based sound objects (Rocchesso, 2004). This approach used physical descriptions and listening tests to develop physically based sound object models that can be manipulated according to everyday experience and embodied into artefacts that support continuous interaction. The descriptions and listening tests were used to empirically determine the important acoustic features of the sound being modelled. These acoustic features were used to develop the sound object models and the parameters that controlled these models. An example of this kind of Auditory Icon system is the Ballancer developed by Rath and Rocchesso (2005) where real time parametric control of the Auditory Icons was used in a simulated task of balancing of a virtual ball on a real stick. The Ballancer is shown in Figure 2.6. Several people who evaluated the Ballancer did not feel the easiness or solvability of the task improved when using the Auditory Display with larger visual displays even though the results showed an improvement when using the combined audio-visual display. The equilibrium task explored by Rath and Rocchesso (2005) did highlight that a well designed sound model can improve performance and the illusion of substance in continuous interaction tasks.

The next paragraph describes in more detail physically based models. These models were informed by earlier work in ecological acoustics and by Gaver's earlier works. These types of Auditory Display demonstrated that it was not just in notification or awareness systems were Auditory Icons could be successfully used. They introduced the possibility of linking real time aspects of systems or processes to Auditory Icons, whose parameters could be manipulated in real time. Many of the older systems lacked the fine detail of control of the Auditory Icons offered by the parametric controls and those that could offer fine control of the Auditory Icons were unable to do so in real time.

Physically informed object modelling systems

Work by Cook (2002b) and by projects such as the *Sounding Object* (Rocchesso et al., 2003) [1] created systems based on new methods of physically inspired modelling of sounds for real

[1] see http://www.soundobject.org

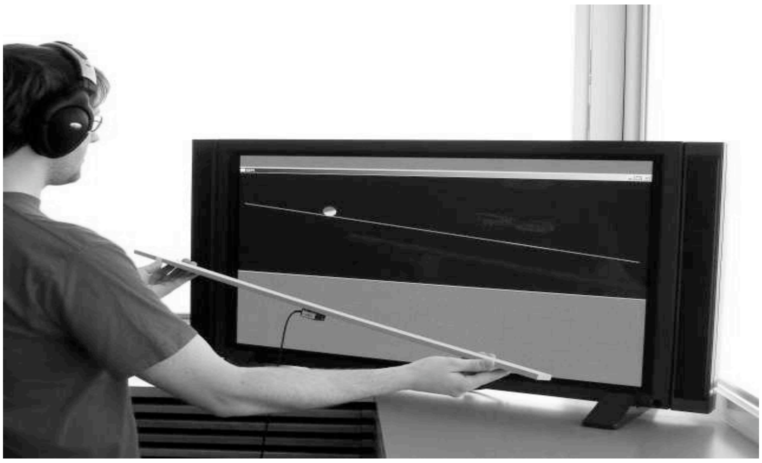


Fig. 2.6: *The Ballancer (Rath and Rocchesso, 2005) auditory equilibrium task. The screen displays the ball whilst auditory feedback helps a user position the ball in the specified spot on the stick displayed on-screen. The physical stick is used to control the movement of the ball.*

time sound synthesis. An example of the difference available with realtime parametric control can be seen in musical interfaces such as the Vodhran (Rocchesso and Fontana, 2003). This was based upon a traditional Irish instrument, the Bodhran. This instrument used a physical based sound model for the sound creation and created a embodied interface where the music emerged as a by-product of the interaction (Fishkin et al., 2000). This type of system is part of the growing area of interactive sonification (Hermann and Hunt, 2005) where the data a user is interacting with, is directly mapped to Auditory Icons and generated in realtime. These types of models are parametrically controlled and responsive, allowing for the control of the sound to be linked to user interactions. This opens a wide range of new interaction possibilities, which were unavailable to previous systems which relied on the playback of sampled sounds. The experiment in Chapter 4 deals with this type of Auditory Icon and was one of the first studies to explore in detail these types of sounds. The work of the *Sounding Object* (Rocchesso et al., 2003) project has lead to many reflections on this type of Auditory Icons such as those by Rocchesso (2004). These systems have also influenced several European COST Actions such as Cost287-ConGAS [2] on *Gesture Controlled Audio Systems* and COST IC0601 Action on Sonic Interaction Design [3] where this type of system has been further developed in a number of prototype interfaces. COST Actions are pan-European research networks with a focus on a particular research area or topic.

[2] see http://www.cost287.org/
[3] see http://www.cost-sid.org/

2.4.4 Lessons and problems of the previous Auditory Icon systems

The previous Auditory Displays, which have used Auditory Icons, have shown that there is a wide range of possible domains where Auditory Icons can be used. Interactions can be dynamically mapped to parametrically controllable sound models, which can provide richer Auditory Icons than previously available. However, technological improvements in Auditory Displays and in Auditory Icons have not yet been supplemented by improvements in evaluations or in reflections on the success or failure of these types of systems. The previous systems have highlighted that technological development without reflection or evaluations on the design will certainly produce more advanced Auditory Displays but that these systems are dependant on the individual skills of their designers.

Auditory Icons and Mapping

A core problem of creating an intuitive Auditory Icon is to link the desired action or concept with an appropriate everyday sound. Mapping between the sound and information or event being represented can vary from arbitrary, to metaphorical, to iconic. Arbitrary mappings are a straight relationship between data and sound, often coming under the category of auralisations. This type of mapping can be problematic especially where many mappings are required or where the Auditory Display is used in a sensitive environment such as a hospital or airplane cockpit. The main problem is that it can be difficult to design clear an unambiguous arbitrary mappings. This type of mapping can be explored using the approach discussed in Chapter 4.

The next type of mapping is a metaphorical mapping, which are more systematic than arbitrary mappings. An example is where high pitches represent high values and low pitches representing low values for a particular data range. The similarities between the item being represented and the representing system, where these are not arbitrary mapping but do not have to depend on physical causation. This type of mapping includes metonymic mappings where a feature is used to indicate the whole, an example being a bird's call being used to represent a bird.

The last type of mapping is iconic, where the characteristics are causally related to the item or items that it is representing. An Auditory Icon can be a recorded sound, or can be synthesised to capture the important features of an everyday sound. The main characteristic is that the attributes of the sound convey information using the causal relations to the attributes they are representing. Iconic mappings are closely related to events they are meant to represent, however this is also a constraint as the representation is limited by the same laws and physics that relate to real world event it is representing. This can raise a practical issue as computer based events do not always map neatly to events in the everyday world.

The repertory grid approach presented in Chapter 6 is one approach to determining what mappings listener's perceive for a given set of sounds. The methods presented in this thesis and the framework presented in Section 7.3.1 in Chapter 7 provide a means of addressing in a structured fashion the issue of what mappings work, what are the actual mappings listener's perceive for a set of sounds, and if the mapping, in particular the auditory aspect conveying the information is likely to be confused. The methods presented in this thesis can be used outside of the suggested framework but to ensure the best design, designers must carry out an evaluation of their displays and of the proposed mappings. The next section presents three hypothetical applications, these are used in the thesis to give an example of how the methods can be used to address particular design challenges.

2.5 The design challenges

The problems of previous Auditory Icon systems have shown the need for better methods for the selection of Auditory Icons. Section 2.1.2 on categorisation and taxonomies highlighted that a listener's perception of a sound event involves both bottom-up and top-down processing. In order to understand how people categorise and classify sounds it is important to understand these types of processing. The work in this thesis uses the CLOSED project's hierarchical classification (Houix et al., 2007a) to provide a well defined core based on empirical studies. This is used as a general framework for the classification of everyday sounds in this thesis. The methods used in this thesis are discussed in the next chapter with an emphasis on how they deal with the challenges presented in this chapter. In order to help elaborate upon these methods, three domains or hypothetical Auditory Display applications will be used and how each method addresses some of the design challenges for these applications will be covered as part of the relevant chapter which discusses the particular method. The three hypothetical scenarios are shown in Figure 2.7, the first scenario is related to mobile devices and messaging, the second to large scale monitoring environments such as a control centre, and the third scenario is concerned with the Auditory Display aspects for an interactive surface or table. The scenarios demonstrate the application of the empirically based auditory design methods to addressing real design problems in Auditory Displays.

Fig. 2.7: *The three hypothetical domains / Auditory Display applications which will be used to highlight where the methods presented in this thesis can address certain design challenges.*

2.5.1 Mobile Devices and Messaging

Mobile technology is a growing area, which is an ideal area for auditory feedback. Current mobile phone technologies deal with single isolated events such as the arrival of a text message and queue events dealing with each sequentially. Spearcons (Walker et al., 2006) are short speeded up spoken phrases or words, these have been shown as one mechanism for more rapidly browsing telephone directories and contact lists on mobile phones. Shoogle (Williamson et al., 2007) proposed the use of sensing technologies and Auditory Icons to convey information about the number of SMS messages or the current battery level. The idea would be to use a shaking metaphor to excite a number of virtual sound objects representing the SMS messages where materiality could be tied to work or friend groups and size of the message could be linked to its weight. A number of other potential scenarios were suggested including notification of forthcoming appointments to the shaking of an iPod to get a rapid overview of the genres in a playlist. This type of interface presents a multimodal display, which does not require any visual attention and leverages user's existing familiarity with real world dynamics and physics.

The potential for non-visual feedback using auditory or a combination of auditory and haptic is a growing area of research. Gestures and auditory feedback can be combined in novel ways as illustrated by Shoogle (Williamson et al., 2007). Designing the mappings for these complex interactions, in particular, the possible auditory mappings and metaphors can be explored in a structured manner using the framework and methods presented in this thesis. The potential for new auditory feedback using gestural mechanisms such as pseudo haptic auditory buttons (Fernström et al., 2005) are opening new avenues for research with mobile technologies. The key features of this domain are:

- Activities and notifications can occur whilst the user's vision is engaged in another task, e.g. driving, cycling, walking, etc. Auditory feedback is one alternative communication mechanism to provide information without distracting from the primary focus of attention.

- Input opportunities are more limited than using traditional keyboard and mouse, gestures and rapid feedback is necessary to help avoid incorrect data or command entry.

2.5.2 Network and/or Process Monitoring

The area of networking monitoring has been briefly discussed with the review of the work on the Peek system (Gilfix and Crouch, 2000) in Section 2.4.2. It used Auditory Icons to represent network events in a continuous sonic environment and was able to provide a real time soundscape. This type of domain was also discussed in the ARKola (Gaver et al., 1991), which provided process monitoring of a factory environment. These previous applications show the definite possibilities for using complex auditory soundscapes to provide temporal information about ongoing processes or events to listeners. This type of domain or environment involves divided attention, multiple tasks, and a variety of information sources providing the information. The use of an Auditory Display is one approach to collating and providing this information in an effective and efficient manner. Unlike visual scanning, aural scanning is better suited to detecting transient or subtle changes in the data as it can process all spatial directions simultaneously whilst visual scanning is concentrated on focal points.

The complex nature of this type of system means that there are many events occurring over a period of time. These events can occur in patterns or overlap. Designing complex auditory representations for this type of situation requires that designers use clearly identifiable mappings and provide a means to minimise confusion where overlapping may occur. The work in this thesis on concurrent Auditory Icons and on the repertory grid provide a new approach for Auditory Designers dealing with these types of problems. The key features of this type of data or process are:

- Complex data with many events, which can occurring simultaneously.
- Multiple processes or data sources representing a wide variety of information at different priorities.

2.5.3 Interactive Surfaces or Tables

The growing ubiquity of technology has lead to a growth in interactive tables or surfaces. This type of application uses various sensor technologies to provide new gestural based interfaces. An example of this type of table is DiamondTouch (Dietz and Leigh, 2001), which allows multiple users through gestural recognition. In specific, it is a type of Single Display Groupware (SDG) (Stewart et al., 1999). A SDG allows a number of co-located people to work together around a single, shared display that functions as the input device. A problem with many of these multiple user touch technologies is how to convey the correct response to the user who initiated the interaction. The current solutions typically use individual headphones to provide individual channels of communication as well as a global group communication channel. An example of this approach is SoundTracker (Ringel Morris et al., 2004).

An Auditory Display could provide similar information without the need for headphones where the information is customised to each user. An example is where gestures on the surface are tracked to each user and where these are mapped to different types of material so one user hears scratching of wood, whilst the next user hears their gestures as water based splashes. Auditory Icons could be used to not only differentiate between users but to highlight particular global events in a similar fashion to SonicFinder (Gaver, 1989) or Audio Aura (Mynatt et al., 1998).

A related area is where auditory feedback is used to provide information of menu, button, and function layout with auditory software buttons (Fernström et al., 2005) providing an auditory only feedback mechanism of non-visual user interfaces. This type of approach is suitable for wearable gestural interfaces such as the WUW – Wear Ur World (Mistry et al., 2009). This technology augments the environment with visual projections onto real objects, which are controlled via hand gestures. This type of interface could be combined with the previously discussed work on physically informed object modelling systems to create gestural interfaces with auditory feedback.

The complex nature of the types of events and interactions that can occur simultaneously in this type of interface require designers who wish to use sonic feedback to ensure that it is identifiable. The method presented in Chapter 5 for concurrent Auditory Icons can help in ensuring the success of such feedback. The key features of this type of interface are:

- Complex gestural interactions with different functions, and potentially several hierarchies of functionality.
- Multiple people working collaborative together. This requires mechanisms to distinguish gestures and provide feedback on the success or failure of those gestures. Auditory Icons are one possible mechanism to provide the feedback.

2.6 Conclusions

This chapter presented an introduction to everyday sounds, Auditory Icons, the various categorisation and taxonomies used for everyday sounds, and to the questions surrounding the creation of these kinds of sounds was discussed. This review highlighted the interdisciplinary nature of the relevant concepts and methods. The difficulties from the various taxonomies highlighted the potential for classification and categorisation error of sound events. Determining the problems and possible methods for addressing everyday sounds provides a useful foundation prior to the creation of a sound design or a sound mapping in a interface. Three major issues are often faced by Auditory Display Designers and will be addressed by the work in this thesis.

- The mapping of perceptual qualities of sounds to concepts or values in an Auditory Display.
- The need for mechanisms to help provide simultaneous feedback of multiple events. One solution is presented in this thesis using concurrent Auditory Icons.
- The categorisation and metaphors used by people to describe the domain or problem. Gathering deep insights using the repertory grid can help find the best metaphors to solve this issue.

The next chapter focuses on topics associated with the methods for the identification of everyday sounds and Auditory Icons. It highlights the methods developed, the rationale and the prior work that informed and contributed to the development of new methods for understanding Auditory Icons.

Part II

Research Methods for Empirically Based Design of Auditory Icons

The part of the thesis provides an introduction to the methods used in this research. It then presents a number of explorations using these methods in a range of issues exploring the identification of Auditory Icons, the meaning of Auditory Icons as construed by listeners, the effect of the realism of Auditory Icons, the scaling of synthetic Auditory Icons. These methods can provide new insights for researchers in the Auditory Icon field and in the wider field of Auditory Display.

Chapter 3

Research Methods

> "There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after."
> *The Lord of the Rings*, JRR Tolkien

This chapter introduces the research methods used in this thesis. The methods provide the foundation for an empirically based auditory design process as part of early stage conceptual Auditory Display design. The details and procedures for using the methods are introduced to enable designers to apply these techniques to their own Auditory Display design process. This can help by providing a firmer ground through explorations and probes than the existing *ad hoc* approaches used by many designers (Frauenberger et al., 2007). The methods provide new details on the identification of salient features of a sound, the determination of confusion of sounds and of the organisation criteria for sounds used by listeners. The determination of confusion can highlight the degree of confusion or ambiguity associated with a sound by a listener. The salient features of a sound can provide details such as realism or on the perceptual scaling of the sound. Organisation criteria provide category or taxonomy information which can be used to provide an insight into the listeners' tacit knowledge and provides information on how listeners organise sounds. In the following chapters the methods and approaches introduced here will be discussed in greater detail, this chapter provides an overview of these methods and approaches. The next section provides an introduction into the particular methods used in this thesis.

3.1 An overview of the methods used in this research

In this thesis several existing methods, guidelines, approaches and taxonomies were used. This section provides an overview of the methods and approaches; each of the methods is discussed in greater detail in the relevant chapter where it is first used. The methods presented are useful for either researchers or designers. The focus of these methods as presented in this thesis is aimed at the early conceptual design stage of Auditory Displays with everyday sounds. The methods can be used at later points of the design process, however as with software development the largest gains can be seen by applying the techniques early. This thesis helps to provide the foundations and move towards an empirically based auditory design process that is accessible to designers or researchers in Auditory Display. A number of methods have been adapted from the fields of Auditory Display or of Human Computer Interaction. A number have been taken from other fields and applied within the domain of Auditory Display to provide this foundation. Newer interaction design frameworks have inspired the methods where the focus is on the interaction between the person and interface or sound in context. Experience focus techniques and similar techniques which focus on perception and meaning (Battarbee and Koskinen, 2005) have highlighted the need for subjective experience methods (Sanders, 2005) that focus on designing with people rather than for them. The first stage in this type of technique is to determine the users' needs and desires as well answering questions like what a person thinks of when they hear a sound or what sounds do they find confusing. The type of information is useful to designers who can use it to ensure the functions / product / interface / sound characteristics are translated into an Auditory Display in a way that produces a meaningful interface for the intended user. This thesis provides the first stage of this type of two-stage process, the methods of gathering this subjective data but acknowledges the need for methods to map this information into actual auditory designs. A interaction design framework with this two stage view and which used many similar methods to this thesis has been proposed by Tomico (2007) as the Subjective Experience Gathering and Inspiring Techniques (SEGIT) approach as shown in Figure 3.1. The work by Tomico (2007) provides a process based on the repertory grid technique for eliciting and analysing high level abstractions used by participants. The process then translates the basic constructs derived using the repertory grid into increasingly more complex ideas and metaphors using analogies and scenarios, respectively.

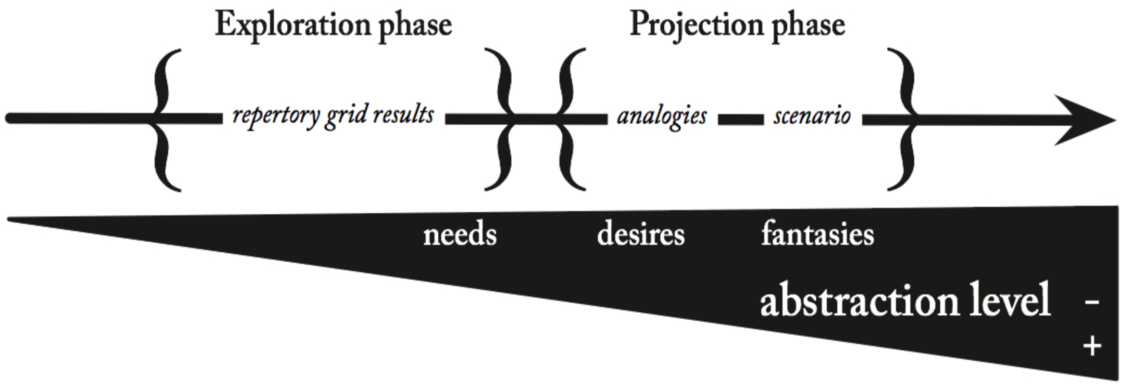


Fig. 3.1: *An illustration of Tomico's (2007) Subjective Experience Gathering and Inspiring Techniques (SEGIT), a two stage interaction design framework for products.*

The work in this thesis provides a set of techniques for Auditory Display designers which sits between psychoacoustics and design as shown in Figure 3.2. The discount HCI style techniques (Nielsen, 1989) were informed by psychoacoustics studies, however such studies are too specific and do not answer general design questions. The thinking is similar to that of Millen (2000) and rapid ethnography, which uses a collection of field methods to gain a limited understanding of users and their activities within a time limited constraints. This style of rapid HCI evaluation can also be seen in work of Pawson and Greenberg (2009) on Extremely Rapid Usability Testing, which uses questionnaires, co-discovery, storyboarding, and observational think-aloud tests in the context of a trade show or conference. These techniques are formative, light weight, and aimed at the rapid collection of good quality feedback without a large overhead. The aim was to provide accessible techniques that could provide empirical design methods at a reasonable cost in terms of time and effort for designers. The techniques aim to provide two viewpoints for Auditory Display designers. The first is an inspirational view point and the second is an informational view point. An inspirational view point is one that relates to exploratory ideas and projection techniques that display concepts at a high level of abstraction. An informational view point relates to lower levels of abstraction and provides specific measures such as identifiability or confusion. Both of these views can use statistical techniques, however it is the final level of presentation of the results of these techniques that classifies the view point. Inspirational uses of statistical techniques provide overview visualisation while informational uses provide measures such as causal uncertainty. These complementary views help in addressing issues at different levels of abstraction as part of an empirically based auditory design process.

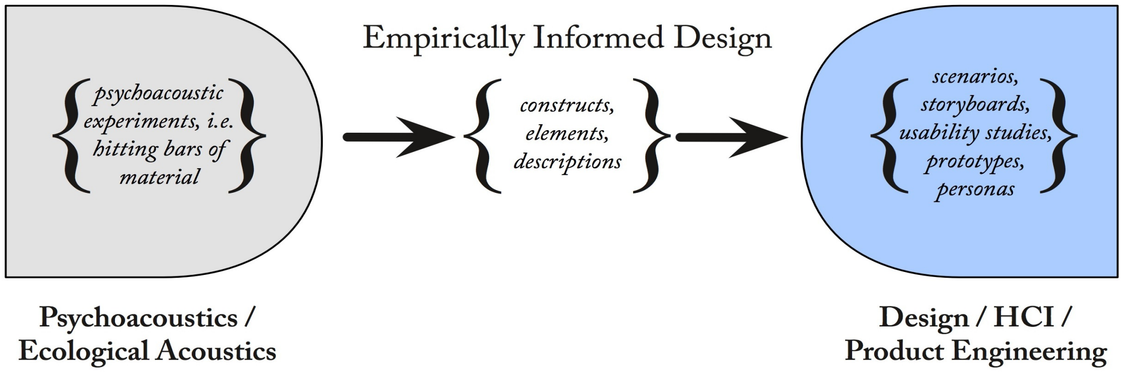


Fig. 3.2: *An illustration of where the methods in this thesis bridge two disciplines, psychoacoustics and design.*

The description of each method includes a short discussion on where and on what it was used for in the context of this thesis and its research. The first type of method focused on the study of listener's subjective perceptions of everyday sounds. A number of statistical techniques were used for this type of study and they included multidimensional scaling (MDS) which can be used for visualisation to help explore similarities or dissimilarities in data. This is mainly used for exploratory data analysis as opposed to hypothesis testing. It can show which are the similar sounds and the relationships between sounds. The goal of the approach and methods used in this thesis are to manage subjective and specific experiences of users from their subjective point of view while ensuring the information gather does not lose its design-engineering focus. This will allow users to tell designers something about something of the way in which they hear and order their worldview. The results of this approach are detailed, reliable and can be unknown to the actual user, particularly where their tacit knowledge was made explicit.

The studies and the analysis techniques used in this thesis are shown in Figure 3.3. This shows the type of analysis but to understand how this is of value to designers, it is better to break down the methods into their informational and inspirational design aspects, these are shown in Figure 3.4. This can help to illustrate how subjective experience can be explored using both types of design aspects. The mix of both informational and inspirational methods used in this thesis provides an approach that facilitates triangulation through multiple methods to more deeply explore the experiences and results of the explorations in this thesis. The mix of measures, direct or calculated are taken and analysed by the informational methods to provide information like how confused a sound was or how a listener clustered a set of sounds. The values and mappings provide information such as a dendogram (tree diagram) visualisations showing the groupings listeners' used for their sounds or textual descriptors that highlighted potential mappings or metaphors. An awareness of the strengths and weakness of each particular method and where each can be used in the design process as either an inspirational tool to guide the creative process or as an information tool to guide the actual choice of sounds within an Auditory Display can help novice designers or those unfamiliar with these methods.

	Principal Component Analysis	Multidimensional Scaling	Cluster Analysis	Causal Uncertainty	Heuristic Identification	Linear Regression
Scaling and Realism of Sounds		✓				✓
Listening Tests				✓	✓	
Repertory Grid	✓	✓	✓	✓		

Fig. 3.3*:* The studies and related analysis techniques used in this thesis.

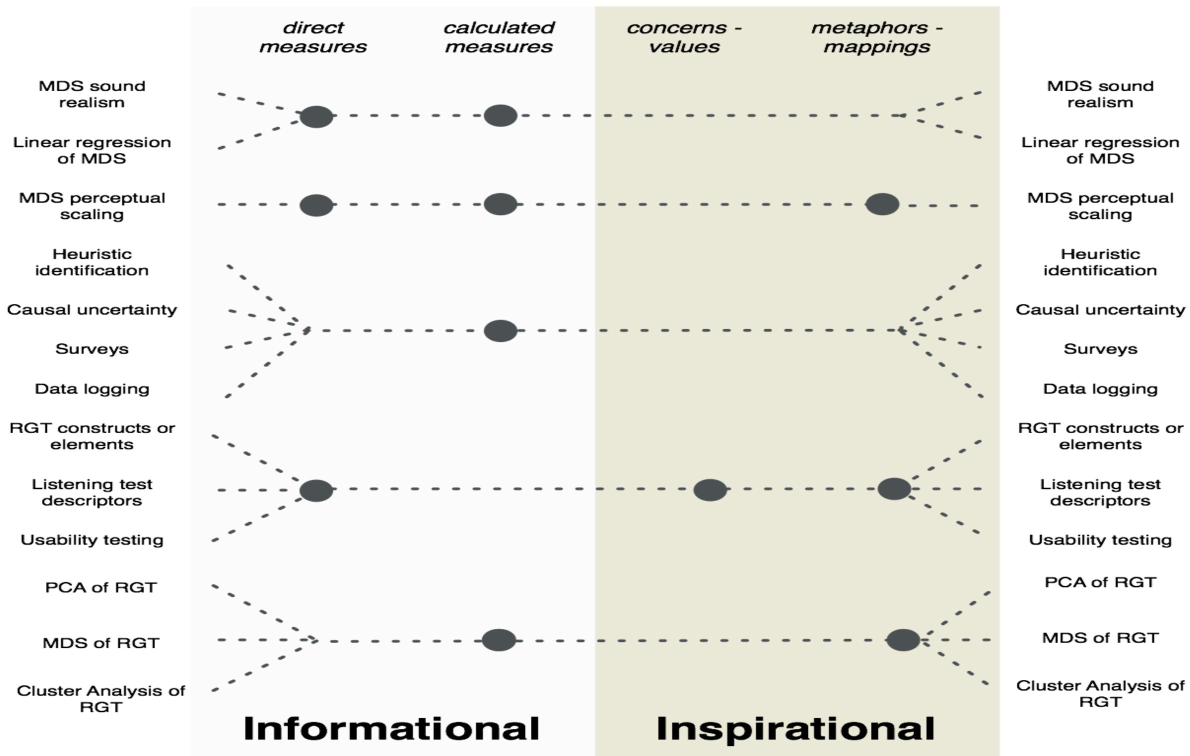


Fig. 3.4*:* The informational and inspirational aspects of the methods and techniques used in this thesis.

The methods used in this thesis to study the perception of everyday sounds are as follows:

Multidimensional scaling for the estimation of sound realism This method is used in Chapter 4 for comparing the realism of "real-world" sounds versus the realism of parameter-based synthesis models. The work of Bonebright (2001) highlighted multidimensional sorting tasks as a method for the collection of similarity data of the perceptual relations among stimuli. This technique is useful for exploring the relations of stimuli particular larger collections (more than twenty five) as suggested by Bonebright et al. (2005). It is useful for situations where the experimenter wishes to use only a single sessions to collect the responses from a participant to all the stimuli. This is a much faster method for stimuli response collection when compared to pairwise comparison (Bonebright et al., 2005).

Multidimensional scaling for the estimation of perceptual scaling of sounds The method is used in Chapter 4 to investigate the perceptual scaling of the synthesised sounds produced by parameter-based synthesis models, in particular, how the synthesised sounds were scaled with regard to their physical dimensions. Two physical dimensions, the perceived height of an object drop and the perceived size of a dropped object were investigated. This method is similar to the multidimensional scaling for the estimation of sound realism and offers the same advantages. Both of these methods allow participants to listen to the stimuli as many times as they wish while producing the groups or scalings that best represent the relations among the stimuli.

Linear regression analysis of the magnitude estimation of perceptual scaling of sounds The linear regression technique is used in Chapter 4 to investigate the perceptual scaling of the sounds, in particular, if any power function existed for the scalings made by participants. Stevens (1974) had suggested that in perceptual magnitude estimation tasks that participants used a ratio scale where x and y are values on the scale where k, a constant exists such that x = ky. This idea of a power law form, which can be used as a method for the stimulus magnitude estimation, was investigated using linear regression. In Chapter 4 exploring the findings from the perceptual scaling of the sounds supported the findings of Luce (1990) and the approach of signal detection theory.

Heuristic for number of correctly identified Auditory Icons In Chapter 5 the identification of concurrent Auditory Icons is explored, however due to the differing numbers of Auditory Icons used in the conditions of the exploration it was not possible to provide a direct numerical comparison between the number of Auditory Icons presented and those correctly identified Auditory Icons. This is a similar problem to the "*correct response*" of Vanderveer (1979). In this thesis, a comparison based on the average number of Auditory Icons identified per participant has been used. This comparison was converted to an average of “correct responses” into a percentage for the number of Auditory Icons that were concurrently presented. The advantages of this technique is that it provides a comparison figure, however the disadvantage of this heuristic is that it is determined by the experimenter and is a subjective measure.

Causal Uncertainty The method of causal uncertainty is built upon the research of Ballas et al. (1986) and is used to determine the number of potential alternative identifications for a set of Auditory Icons. This method is used in Chapter 5 and Chapter 6. A full description and background to this method is provided in Chapter 5. This thesis extends Ballas’s research by applying the method to conditions, which used the presentation of simultaneous or concurrent everyday sound. The prior work with this method focused on sequential or single presentation of everyday sounds. This thesis shows how this method can provide new insights in conditions with simultaneous sounds. This method provides a single measure for a sound or an average for a set of sounds, which easily describes the confusion of their identification. This measure summarises skewed responses, unity, or degree of split in a single figure.

Repertory Grid Technique The repertory grid technique was extended from George Kelly’s work (1955) in this thesis to address Auditory Icons. The use in this thesis of the repertory grid as a subjective experience information gathering technique places the focus not on the subject (as would occur in psychology) but rather on the information they generate, the elements and constructs. Design relevant information can be obtained from participants by analysing their constructs and elements (Hassenzahl and Wessler, 2000). The method is discussed in detail in Chapter 6. The data obtained from the repertory grid technique were analysed using three statistical techniques. These techniques were principal-component analysis, multidimensional scaling analysis, and cluster analysis. These techniques and their application in the analysis of repertory grid technique data was described by Fransella et al. (2004). This method extends the work in the field, as it is the first time it has been used to explore everyday sounds to elicit individual vocabularies to support the design of Auditory Icons. The method was used as one approach for eliciting and interpreting the descriptors provided using the listening test approach. It has been used in acoustics for loudspeaker and musical timbre vocabulary studies and these are discussed in Chapter 6. The advantages are that it provides a mathematical construction of an individual’s psychological space to derive the personal constructs and the approach used encourages personal reflection upon the qualities of the stimuli under examination. This is a better means for exploring a participant’s underlying *semantic constructs* than other methods which use ‘pre-set’ lists of adjectives from which the participant must select. The participant’s responses are therefore more natural semantic responses than elicited by some of the other force choice approaches.

Principal-Component Analysis of Participant’s Grid Data The principal component analysis technique (Jolliffe, 2002) is a statistical technique used in Chapter 6 to analyse a participant’s grid data. A full description and background to this method is provided in Chapter 6. It is used to systematically represent the relationships between the constructs and their correlations while taking into account the negative correlations that occur in the use of a bipolar scale. The advantage of this method is that is can help in determining new variables or trends of interest in the data. This is useful as the number of descriptors produced from a participant’s grid data can be difficult to interpret without statistical techniques to provide pointers to items or trends of interest. This approach is similar to the idea of experience landscapes (Tomico, 2006) where constructs and elements from a grid are used to generate a spatial analysis visualisation as one way of representing each participant’s grid. This visualisation approach is also taken for the multidimensional scaling and the cluster analysis techniques.

Multidimensional Scaling Analysis of Participant’s Grid Data The multidimensional scaling (MDS) analysis technique (Bonebright et al., 2005) is a statistical technique used in Chapter 6 to analyse participant’s grid data. This method produces a two dimensional spatial representation of the constructs and elements of a grid and is a standard method for providing a repertory grid visualisation. Using MDS to visualise the pairwise distinction of a participant’s similarity structure can help in the identification of regions or *facets* (Borg and Shye, 1995) where stimuli posses similar values. Examining the MDS analysis and supplementing it with a minimum spanning tree (MST) can help to identify any inaccuracies present in the MDS representation (Jolliffe, 2002). This method’s advantage is that it can easily visualise the underlying structure of relations between elements. There are a number of common patterns which can easily be seen in this type of plot and which can point to the number of dimensions of the phenomena being investigated.

Cluster Analysis of Participant’s Grid Data The cluster analysis (Agresti, 1996) technique is used in Chapter 6 to analyse participant’s grid data. A full description and background to the cluster analysis (CA) technique is provided in Chapter 6. This is a similar method to the principal component analysis and provides another way of illustrating the relationships among constructs and elements. CA is used to group the objects that are represented by participants' similarity structures. This highlights potential groupings of objects using a tree structure. In a similar fashion to MDS, this method offers the ability to represent deep structures from within the data in a visual and understandable manner. The difference between these methods is that MDS results in a non-hierarchical view of the underlying structure of relations between elements in the data set whereas CA presents results in a hierarchical structure of the elements in the data set.

The second type of method used related to the study of the usability and design of Auditory Icons for use in Auditory Displays. The results from these methods are only of secondary concern to this thesis but it is mentioned here for completeness. The methods from this area that this thesis used are as follows:

Usability Testing The software or scripts presented in the explorations (Chapters 4, 5, and 6) in this thesis used task lists, scripted instructions, and post-test questionnaires for usability testing (Kuniacsky, 2003). The advantage of this approach is that it can easily gauge participants responses to the tasks.

Surveys The explorations presented in Chapter 4 used a survey (Courage and Baxter, 2005) to elicit feedback from participants over the course of their use of the synthesised sounds. The rationale and reasoning behind this approach is discussed in greater detail in Chapter 4. The advantage of using a survey is that it can help in directly questioning aspects of the phenomena or system under investigation.

Data Logging The software presented in Chapter 4 used application data logging to capture usage and request data from participants. This type of data can help in determining a participant's actions and path through the session.

3.2 Rationale for selection for the research approach

Investigating Auditory Icons at the early design stage can be undertaken using a number of different research approaches. It is useful from a theoretical viewpoint to study how listeners classify sound events as it helps in understanding their ability to identify the causes of sounds as well as how they organise their knowledge about the particular sounds. The results from the explorations in this thesis can assist in gaining a perspective on various issues including what are the basic categories of sounds or can these sounds be used together. The focus of the human computer interaction approach in this research is on understanding of individuals, their concerns, desires, aspirations, values, and experiences. A sound and its classification can provide categories linked to the interactions users may experience through an Auditory Display. The range of techniques discussed in this chapter provide reliable mechanisms which can be used in the early stages of design to assure the correction interpretation of users subjective experience of non – speech auditory sounds.

3.3 Conclusions

This chapter presented an introduction to the research methodologies used in this research. The collection of a broad range of methods used in thesis helped in addressing a number of different issues for Auditory Icon research by providing different viewpoints and approaches. The chapter provides a summary of methodologies, prior research, and an overview of the methods used in this thesis.

In the following chapters the methods and approaches introduced here will be discussed in greater detail, this chapter provided an overview of these methods and approaches.

The next chapter focuses on an exploration using multidimensional scaling for estimation of sound realism and for the perceptual scaling of sounds. It highlights the methods, the rationale and how the prior work informed this research. In particular, the next chapter describes a novel study providing answers for RQ1, "*Does the subjective realism of a sound affect the response of a listener to the sound ?*" and for RQ2, "*Do listeners subjectively hear the same physical properties of objects when both synthesised and sampled versions of the same sounds are used ?*". This study is a good starting point for people who wish to use parametrically synthesised sounds in their applications. It highlights the issues of realism and of scaling which can affect the use of this type of sound. These sounds can be dynamically controlled and offers a greater range of potential than existing pre-recorded sound files.

Chapter 4

Investigating Auditory Icons Using Scaling Methods For The Estimation Of Psychological Scales

> "If you develop an ear for sounds that are musical it is like developing an ego. You begin to refuse sounds that are not musical and that way cut yourself off from a good deal of experience." John Cage

This chapter explores the perceptual scaling and the realism of everyday sounds and of Auditory Icons. It helps answers the question of mapping between the dimensions of sound and the dimensions of perception by providing an approach to explore how the perceptual attributes of sounds are scaled by listeners. A new technique was developed in this thesis to explore subjective issues and was inspired by the existing techniques in traditional psychoacoustics and in ecological psychoacoustics. The technique allowed for the collection of scaling data and realism rating of sounds in a single session. The problem of ensuring that the designer's interpretation of a mapping and that of the listener's are the same can be solved by following this approach. Auditory Icons and their mappings are often not tested to ensure these two interpretations are the same or similar. Ensuring the correctness of a mapping can prevent any difference perceptual interpretations of the Auditory Icon. A summary of the benefits to designers and of the results of this chapter are shown in Table 4.1. There are two aims in the work presented in this chapter. The first is to investigate the perceptual scaling of synthesised sounds produced by the parameter-based synthesis models, in particular, how synthesised sounds are scaled with regard to their physical dimensions and the suitability of the chosen physical dimensions for use. The second aim is to compare the realism of recorded sounds versus the realism of parameter-based synthesis models. Exploring the perceptual scaling allows us to see if parameter-based synthesis models are capable of communicating quantifiable information to users.

Dropping sounds are used in this exploration as this type of impact sound covers a range of common everyday sounds including keys jangling, sweets rattling in a jar, ping-pong balls bouncing, and water sloshing in containers. Humans are well versed in everyday listening and at inferring the physical properties of materials from the sound of their physical interaction. This inspired the choice of the two physical dimensions for examination, the dimensions were the perceived height of the object drop and the perceived size of a dropped object. This work can be seen as examining the dropping aspect from the derived processes layer in the Sounding Object project's taxonomy as discussed previously in Section 2.1.1. The results of this work highlight the effect of surface material perception with regard to dropping sounds, the effect of room acoustics on the realism of synthetic sounds, and it provided a new method that explored the perceptual scaling of parameter-based synthesis models. In particular, two pilot studies were conducted exploring the new technique, the first concentrated on determining the suitability of the approach and helped refine the sound selections for the second pilot. The second pilot used more complex two mode synthesised sounds to see if these conveyed information about their physical dimensions better than one mode synthesised sounds.

Informational	Scaling data of real and synthesised sounds. Realism ratings of real and synthesised sounds. Suitability of physical parameters in synthesised sounds.
Inspirational	Potential for use of synthetic sounds as Auditory Icons. Favouring of one mode over two mode synthetic sounds by listeners, possibly suggesting that simple synthetic sounds can be as effective as more complex sounds (cartoonification).
Difficulties	Confusion due to "Buzz tail" in two mode synthetic sounds. Effect of surface material on perception of the drop. Lack of room acoustics in synthetic sounds.
Contributions	New 2D method for scaling and comparing sounds (real or synthetic). First exploration into perceptual scaling of parameter-based synthesis models.

Table 4.1*:* Summary of informational and inspirational aspects of the methods and techniques from this chapter and also the difficulties and contributions from this chapter.

There are a number of reasons why this type of dropped sound was chosen. Firstly, it is likely that information about the dropped item's height and size is available when it is dropped and that this information is relevant for listeners. Second, the choice of synthesised and real sounds provides two different mechanisms for presenting the same or very similar information to listeners. Thirdly, the changes in material and height can be understood in terms of the physical model of the event. This type of understanding can help in determining the ways in which variations to the sounds provide information about the changes to the sound source. The fourth reason is that this is a simple event that is not tied to any specific place or environment and as such a listener should not need to rely heavily on prior knowledge of the sounds used in the study. This type of sound fits Gaver's (1988) idea of a *basic sonic event*, as it can be generalised to describe a large number of specific instances (in this case, the class of solid objects being dropped). It is not generic enough to group with other events, however many features are shared by basic sonic events within the same category. In this study, dropping might contain features or invariant information involving the sequence of onsets of each subsequent and usually less powerful drop or impact. This type of sound is common to almost any environment, which makes it likely that no listener will be more expert in their perception of this sound than any other listener.

Do the height and the size attributes of dropping sounds convey the essential attributes and information of the event? One aspect of this research examines whether height and size can convey the essential attributes of dropping events with a wider view towards using this category of sounds within Auditory Displays. This question takes Gaver's (1993b) idea of parameterised Auditory Icons and explores if the parameters of height and size of the Auditory Icons can be used to convey information to listeners. A number of studies (Giordano, 2003, Giordano and McAdams, 2006) have shown that these attributes characterise a basic sonic event. This examination will provide more details on the scaling of these attributes of dropping sounds and leads to the next part of the study, which specifically narrows this questions to the two types of sound (real and synthetic) and asks if they convey the same attribute information.

Can synthetic sounds convey the same characteristic attribute information about dropping events as real sounds? The work of the Sounding Object project built upon many of Gaver's (1988) ideas including the possibilities of using synthetic sounds. The synthetic sounds from this project were used in this study to explore if they convey the same information as real dropping sounds. The earlier work by Gaver (1988) had explored struck metal and wood bars to judge their length and material using synthesised sounds and while the results were poorer for synthetic sounds, it was judged that the model did provide a good match to the description of the event being modelled. The sounds were presented binaurally using headphones to participants who categorised them using a one of three choices. The choices were wood, metal, or 'do not know'.

These types of models can provide a vast range of similar but unique sounds at low computational costs with full parametric control of the various attributes such as height, size, or material. There is a lack of understanding about the use synthetic sounds as Auditory Icons.

The work in this chapter helps provide a method, which can judge the ability of synthesised sounds for conveying information. This will allow Auditory Display designers to explore this type of sound and determine its usefulness for their designs. In a similar fashion to the real sounds, the height and size of the synthetic sounds should characterise a basic sonic event whose perception is relatively stable across participants. The results of the study help in understanding the suitability of the studied attributes and whether they can convey the same information as real sounds.

There are several experimental methods that can be used to explore the questions raised. This includes magnitude estimation (Stevens, 1974) and multidimensional scaling (Borg and Groenen, 1996), two methods from a much broader list (Bonebright et al., 2005). Newer methods from ecological psychoacoustics, inspired by Gibsonian thinking (Gibson, 1966, 1979), offer the potential to conduct studies where more than one variable may change and that can involve ill-defined tasks. These newer methods and the studies using them explore more realistic situations of use for Auditory Displays. These types of studies are also making use of more realistic sounds, an example is Bonebright's (2001) study of perception and categorisation of environmental sounds through the use of multidimensional scaling techniques. Researchers such as Ballas (1993), Ballas (1994) and Gaver (1994) have used "*real-world*" sounds in their studies. These types of studies provide a better understanding of the sounds and their meanings.

Bonebright's (2001) approach allows for the analysis of real sounds to determine their salient features. These salient features can be extracted to create parameter-based synthesis models. This type of Auditory Icon will have certain acoustic similarities with the objects they are representing. These salient features can be further "*cartoonified*" (Fernström, Brazil and Bannon, 2005) to exaggerate the most salient features of the sound. These parameter-based synthesis models are artificial sounds and can be parametrically controlled in real time. These sounds can be as compelling as the original real sound while being more computationally efficient and dynamically controllable within the Auditory Display. A review of physically based parametrically control synthesis models by Rocchesso (2004), gives a good review of this approach and discusses how it is now possible to design, develop, and embed the models into interactive artifacts supporting continuous interaction. This chapter presents a study using both real and synthesised sounds using a new approach that differs from previous psychoacoustic methods.

4.1 Prior psychoacoustic methods and approaches

Classical pairwise comparison or similarity rating studies using a multidimensional scaling technique requires that a listener rate all the possible stimulus pairings. The number of comparisions required for similarity ratings are shown in equation 4.1.1 where N is the number of stimuli being examined and C is the number of comparisons required.

$$C = N(N-1)/2 \tag{4.1.1}$$

In this chapter, 18 different sounds are examined and using a traditional approach this would take 153 comparisons. It is open to debate whether participants can keep a coherent set of comparison ratings even for such a small set of stimulus. The technique presented in this chapter is an alternative approach to similarity rating comparisons.

4.1.1 Computer based sorting compared to pairwise comparison

Work by Bonebright (1996) found that computer based sorting was better at reflecting the emotional cues in the stimuli when compared to a paired comparison approach on the same sound stimuli. The reasoning behind interactive computer based sorting is that it does not constrain participants from applying more complex comparison strategies and it allows for participant's to change their sorting criteria based on a better understanding of the stimuli at any point. This inspired the work of Scavone et al. (2000), who developed an interactive application, the Sonic Mapper to provide an interactive 2-dimensional space were the sounds were visually and aurally displayed. This helped participants hear how the sounds related to one another simultaneously. This type of approach allows listeners to arrange and group sound stimuli according to the similarities found between them by the individual listener, while providing a full overview of the stimuli set helping listeners gain a better appreciation of the entire set when compared to a pairwise comparison task. The motivation is that this process decreases user fatigue, as it requires fewer comparisons than classical psychoacoustical methods, it is hoped that this approach also increases the listener's attention and their decision consistency. In the case of large stimulus sets, it is difficult, if not impossible, for listeners to retain the same criteria with such large number of sounds. Scavone et al. (2000) post experiment interviews with participants found that pairwise comparisons were artificial and unintuitive when compared to the stimuli sorting offered by the Sonic Mapper.

The goal of Sonic Mapper was similar to that of the Sonic Browser (Brazil, 2003). The interface provided by the Sonic Browser used an aura or cursor to provide a stereo-panned sound field with drop and drag, and tagging functionality. The sounds in this study, represented by linked visual and Auditory Icons, were played simultaneously while panned to their correct location under the aura. The panning and related multiple concurrent sound playback functionality of the Sonic Browser was the major difference between it and the Sonic Mapper. This interface provided an experimental platform for exploring the psychophysical experiments presented in this chapter. An introduction to the Sonic Browser (Brazil, 2003) is given in Appendix A.1. It provides an interactive simultaneous presentation with direct manipulation (Shneiderman, 1983) and results in a 2-dimensional rating scale.

4.2 Parameter-based synthesis models

The modal synthesis models used in this chapter provide a computer model of a sound, which reduces the interacting physical parts into mathematical formulas. These formulas are used to create the algorithm that drive the models. Driving the parameters of these models will output the appropriate sound. The number of modes used in the model increasing the computational complexity and processing required. The models used for synthesis in this chapter were created as part of the EU IST Disappearing Computer initiative and, in particular, the project "the Sounding Object" (SOb). The model's development was influenced by the earlier work of Gaver (1993a, p. 292) who highlighted that "many of the sounds we hear in the everyday world involve one solid impacting against another". This lead to the creation of two types of modal synthesised models, one for impact sounds and one for friction sounds. The models were written in *PureData* and created by the University of Verona, Italy as part of this project. A modal synthesis model is a bank of damped harmonic oscillators. These oscillators are controlled and excited by external stimulus, the settings of the oscillators such as their frequencies and dampings are based on the geometry and material properties of the real object being modelled. The sound model parameters for a given real object are used as the basis for the synthesis model. These parameters are obtained experimentally using recordings of the object's impulse responses by fitting the synthesised model's parameters to the recordings of the object. The design, development and a detailed exploration of the modal synthesis models used in this chapter can be found in Rocchesso and Fontana (2003). The models used in this research were modal synthesis impact models with a higher-level control model for bouncing with the additional parameters that included the materiality of the dropped object. A technical reference (Rath and Fontana, 2003) for the models covers their main features and the association between parameters.

4.3 The pilot studies

The aim of the studies was to understand how listeners scaled the synthesised sounds produced by the models, in particular by their perception of their physical dimensions. The method used allowed for any two dimensions to be explored using a 2-dimensional plot. The two dimensions of interest selected from the various physical dimensions were the perceived height of the object drop and the perceived size of the dropped object; additionally the perceived realism of the event was explored.

The first study explored a limited subset of the stimuli with the second study exploring the full set of synthetic stimuli. The first pilot allowed for an initial observation of the procedures and the results as well as determining the most suitable sounds that were to be the stimuli for the second pilot. The second pilot concentrated on the synthetic stimuli and looked at one and two mode synthesised sounds to see if the addition of a mode added more realism to the synthetic sounds. In the next subsections, the pilots are covered with the procedures and results being introduced and discussed.

4.3.1 The first pilot

In this study the stimuli had a fixed height to theoretically limit the variance to the perceived size of the dropped object. The procedures and initial sounds were observed and analysed using the first pilot. The following hypothesis were made with regard to the expect results of this study.

Hypothesis 1. *The synthesised sounds would convey perceptually distinct sizes and heights.*

Hypothesis 2. *The height scaling of the sounds would be consistent across all the sounds.*

Hypothesis 3. *The recording distance of the real sounds will be heard as smaller when recorded further away from the microphone.*

Hypothesis 4. *The scaling of sounds would be consistent within the same type of sound.*

Hypothesis 5. *The synthesised sounds convey height and size information as well as recordings of real ball dropping sound events.*

Participants

There were 4 participants (1 male, 3 females) in the first pilot; all were either postgraduate students or researchers. In pre-screening for the study, all reported to having no hearing or sight problems and all had prior musical training (5, 4, 2, and 10 years respectively).

Stimuli

The set of stimuli chosen for the first pilot consisted of 9 recording of live sounds and 9 recordings of synthesised sounds taken from the previously mentioned Sound Object project's dynamic impact model. All the sounds consisted of two or more “bouncing” events. The stimuli in the pilot were limited in the height scale so that all of the stimuli were synthesised or recorded at a height of 20 cm. This limitation was chosen to allow for the pilot to see if a base line existed for the height scaling of the stimuli. It was hypothesised that with a fixed height that the major source of variance would be limited to the size of the dropped object.

The recordings of the live sounds consisted of three steel balls (weighing 6, 12, and 24 grams respectively) dropped on a wood surface of 1500 x 500 x 20 mm from a single height (20 cm) where the microphone was positioned at three different distances from the surface (20, 40, and 80 cm respectively). These recordings were made using a MKH20 Sennheiser microphone at a sampling rate of 44.1 kHz and at 16-bit resolution.

The recordings of the synthesised sounds consisted of sounds taken from *PureData* programs (patches) and modelled the interactions of two modal resonators (Rath and Fontana, 2003). These models were simplified to a single mode and with the material type or property of the synthesised model set to either glass or wood. In the first pilot, the height of the dropped balls in the synthesised sound recordings was kept constant at a height of 20 cm. The synthesised sounds represented either glass or wood objects being dropped. The choice of two different materials was included to help determine if one material was more suitability or had better scaling results.

The stimuli used in the first pilot are presented in Table 4.2, the particular settings and values used to generate the particular synthesised sounds from the impact model are shown in Appendix C.1 in Table C-1. The abbreviations used in for the sounds are as follows *d* represents the microphone recording distance, *w* represents the weight of the ball and *h* represents the distance dropped. The synthetic sounds use the additional abbreviations of *gl* to represent glass and *wd* to represent wood. The impact model, its main features and the meanings of its parameters are discussed in further detail in Rath and Fontana (2003).

Procedure

The study was conducted in an isolation room with the stimuli presented over stereo headphones using the Sonic Browser (Brazil, 2003) application. The pilot consisted of three stages, the first was the perceptual estimation task, the second was the “tagging” of unrealistic sounds by the participants, and the final stage was the use of a questionnaire on the participant’s perceptions of the performed tasks and of the Sonic Browser application. An explanation of how the participant would scale and tag the sounds in the study is shown in Figure 4.1. Participants arranged the sounds within the 2-dimensional space using dragging and dropping to reflect their perceptual estimation of each sound’s height and size. These sounds were then tagged by participants to highlight those felt to be unrealistic. The participants had no time limit or playback limit with the study being completed when they felt satisfied with the scaling and “tagging” of the sounds.

Sound file	**Sound type**
d20-w6-h20.wav	Real
d20-w12-h20.wav	Real
d20-w24-h20.wav	Real
d40-w6-h20.wav	Real
d40-w12-h20.wav	Real
d40-w24-h20.wav	Real
d80-w6-h20.wav	Real
d80-w12-h20.wav	Real
d80-w24-h20.wav	Real
w-s-h-s-gl-pd-1.wav	Synthesised- Fixed at 1 Mode
w-m-h-l-gl-pd-2.wav	Synthesised- Fixed at 1 Mode
w-m-h-s-gl-pd-3.wav	Synthesised- Fixed at 1 Mode
w-m-h-s-gl-pd-4.wav	Synthesised- Fixed at 1 Mode
w-s-h-m-gl-pd-5.wav	Synthesised- Fixed at 1 Mode
w-s-h-m-wd-pd-1.wav	Synthesised- Fixed at 1 Mode
w-m-h-m-wd-pd-2.wav	Synthesised- Fixed at 1 Mode
w-s-h-m-wd-pd-3.wav	Synthesised- Fixed at 1 Mode
w-s-h-s-wd-pd-4.wav	Synthesised- Fixed at 1 Mode

d - 20/40/80 - microphone recording distance in cm
w- 6/12/24 - size (diameter) of object in cm
h - 20 - height of dropped object in cm
gl - synthesised glass wd - synthesised wood
w - s/m - size of object h - s/m/l - height of drop of object
gl - synthesised glass wd - synthesised wood
pd - 1/2/3/4/5 - PureData synthesised sound settings

Table 4.2: List of the real stimuli and of the synthesis stimuli used in the first pilot.

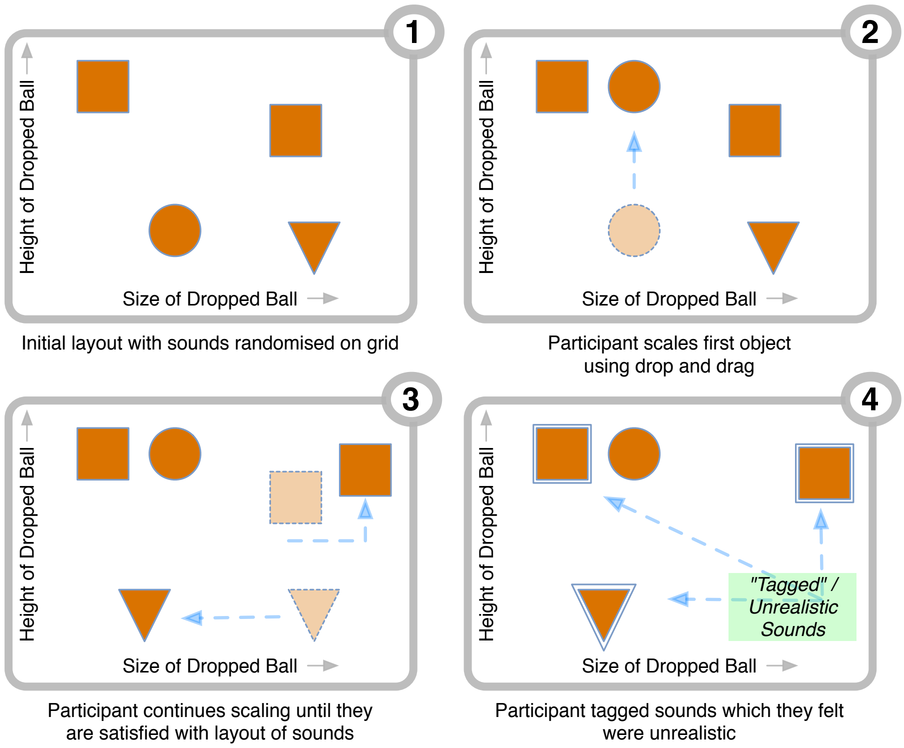


Fig. 4.1: *An overview of how the process of scaling and of "tagging" of the stimuli in this study.*

The stimuli were represented in the Sonic Browser as various coloured geometric shapes randomly displayed on its 2-dimensional plot. An introduction to the Sonic Browser (Brazil, 2003) is given in Appendix A.1. The Sonic Browser application allowed users to listen to the stimuli and using its drag and drop interface to scale the stimuli within the specified 2-dimensional scale. The plot axes were labeled according to the dimensions under investigation, i.e. perceived height of the object drop on the y-axis and perceived size of the dropped object on the x-axis. This scale used had the perceived size of the dropped object on the x-axis while the perceived height of the object dropped was on the y-axis, this is shown in Figure 4.2. The cursor in the Sonic Browser was surrounded by the aura. The aura could be resized or turned off and on by the subjects. The subjects listened to the sounds by moving the cursor over them and, if the aura was large enough, they would hear the surrounding sounds simultaneously. The participants scaled the stimuli in the perceptual space by dragging-and-dropping the shapes that represented the stimuli. The Sonic Browser allowed for the collection of object placement points within a 2-dimensional space, in addition "tagging" of sounds was used to highlight if the participant deemed the particular sound unrealistic. The "tagging" used a combination mouse and keyboard shortcut command to designate the sound as unrealistic, this could be undone using the same command. This allowed the users the opportunity to change their opinions prior to the end of the study.

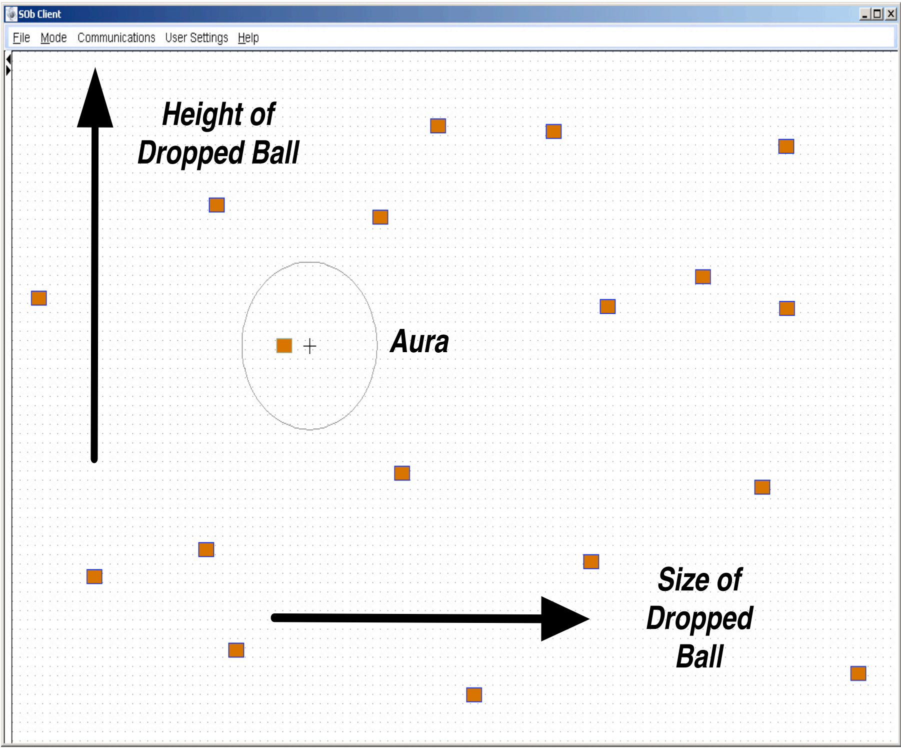


Fig. 4.2: *The Sonic Browser interface (Brazil, 2003), used for the scaling and "tagging" of the stimuli in this study.*

A post-experimental questionnaire and participant debriefing supplemented the application data logging from the Sonic Browser. This was used to give an insight into the perceptions of the participants with regard to the tasks and to the Sonic Browser. All the tests were captured by video logging of the participants and all participants signed the necessary consent forms with regard to the study.

The participants used the 2-dimensional space to estimate the positions of the stimuli without a comparison stimulus or a reference scale. There was a pre-defined range for the study, that being the screen size, however it was found that the range of perceptual evaluations was relative to each individual participant. These perceptual space boundaries were considered by all the participants and were relative to their particular individual maximum values. An observation that occurred during the first pilot was that participants had a difficulty in referring to the screen space as their boundary range, they showed a preference for setting their own boundaries. This observation meant that for a meaningful comparison of the results from each participant to be compared, the data co-ordinates of their scalings had to be normalised. This normalisation was used to map the locations from the 2-dimensional space to a range from 0 to 1. The final part of each session required the participants to fill out a 7-point Likert scale questionnaire with semantic labels for the end points. The questionnaire is shown in Appendix D.1.

Results and Observations

Figure 4.3 and in Figure 4.4 shows the perceptual scaling (similarity ratings) of the participants in the study. The exploratory nature of this pilot study means that the results are indicative rather than firm findings and where hypotheses are supported; these are merely qualitative findings. Both of the figures are broken into two sub-images as the data is drilled down to focus on particular aspects. In Figure 4.3a, an overview of the scalings are shown by sound type. This indicates support for Hypothesis 1 as the synthesised sounds can be seen to convey distinct heights and sizes to the listeners. The overview does not support Hypothesis 2 as it is evident that the scalings did not have a consistent height scaling. Figure 4.3b focuses on the real sounds and the distance they were recorded from a microphone. The trend, 1, in Figure 4.3b indicates support for Hypothesis 3 where the furthest away recorded sounds are taken as smaller than those which were recorded closer to the microphone even when the furthest away sounds were physically larger. Figure 4.4c consolidates the synthesises sounds together to focus on the perceived size and height. It shows a relatively good match between the height and the size scalings from the participants with regard to the synthetic sounds.

The trends, 1 & 2, in Figure 4.4c show that a trend for the height of the dropped ball and how this may also affect the perceived size of the ball, were larger drops make the ball size seem smaller. This indicates support for Hypothesis 4 that the particular type of sound, real or synthetic, would scale within its own type of sound. The 95% confidence intervals for the consolidated sounds are shown in Figure 4.4d. Analysis showed a high variability in the scaling of the synthesised glass sounds. This contrasted with the synthetic wood sounds, which barring a single outlier had a much lower variance. This can be seen in Figure 4.5 which unpacked the synthetic sounds and plotted only the wood sounds to graph this finding. This would indicate that the wooden synthetic sounds are better scaled than the glass synthesised sounds. Hypothesis 5 was not supported and the results show that the synthetic sounds do not convey information about size and height of drop as well as real recording of such events. This finding for the one mode synthetic sounds inspired the second pilot, which looked at using more complex two mode synthetic sounds, in addition, to the already used one mode synthetic sounds to see if these more complex sounds would better convey information to listeners with regard to the size and the height of drop.

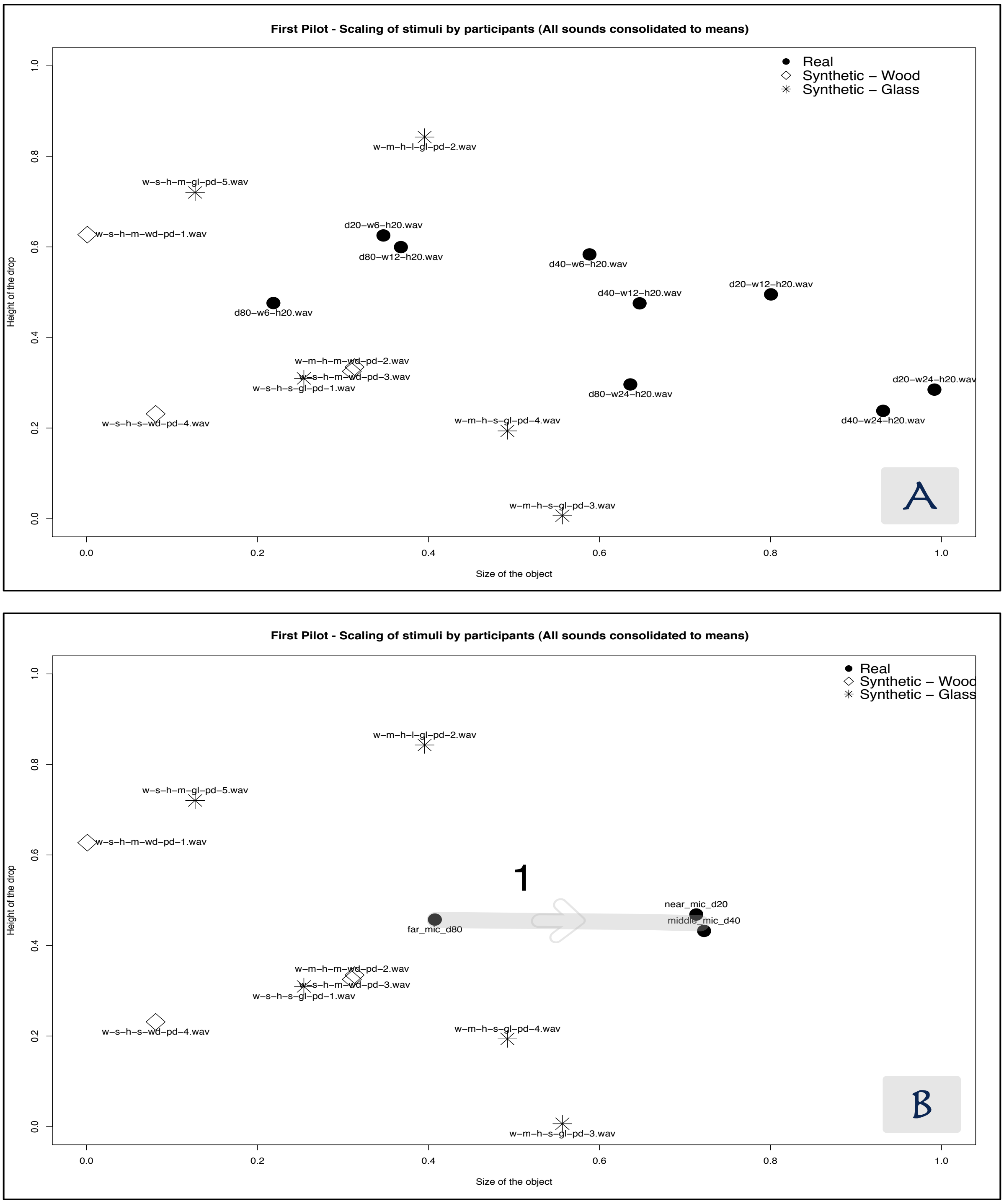
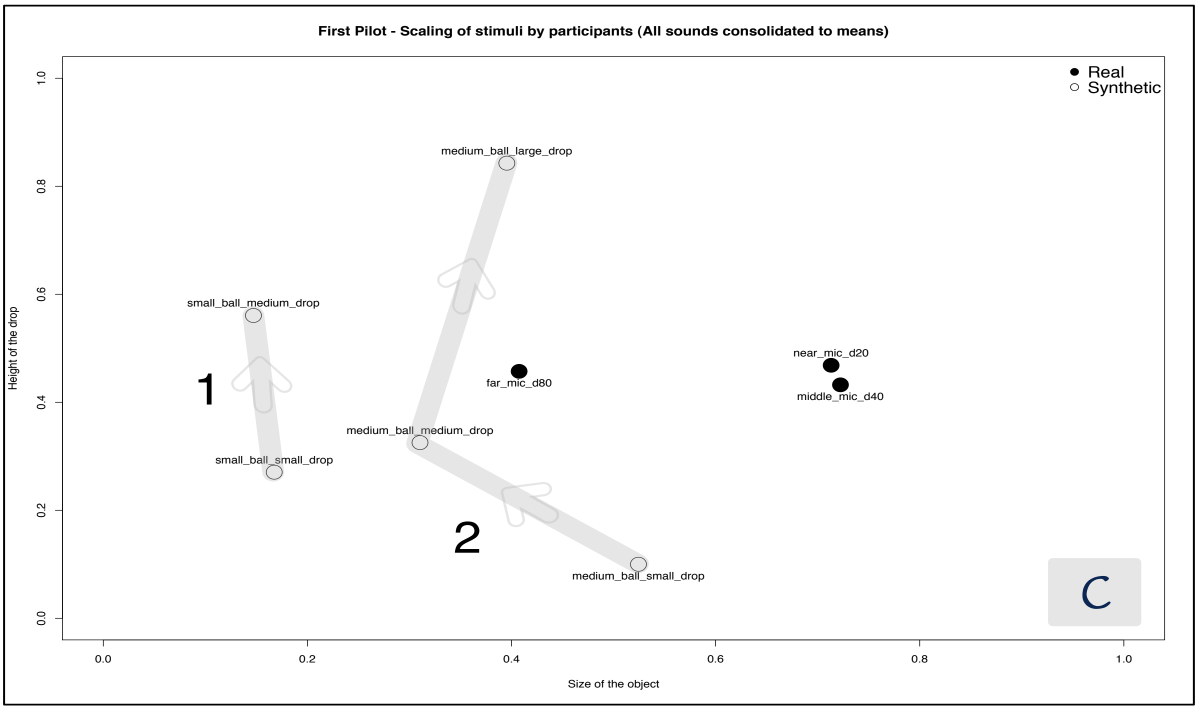


Fig. 4.3: *The perceptual scaling averages of the participants in the first pilot. The top figure, 'A', represents the complete scaling data of the sounds presented. The bottom figure, 'B', drills down and collapses the real sounds to the distances they were recorded at with a microphone. The trends, 1, in 'B' shows the scaling trend for the distance of the microphone recording of the sound. This affected the perceived size of the real sounds in the study with closer recordings distances being perceived as larger.*



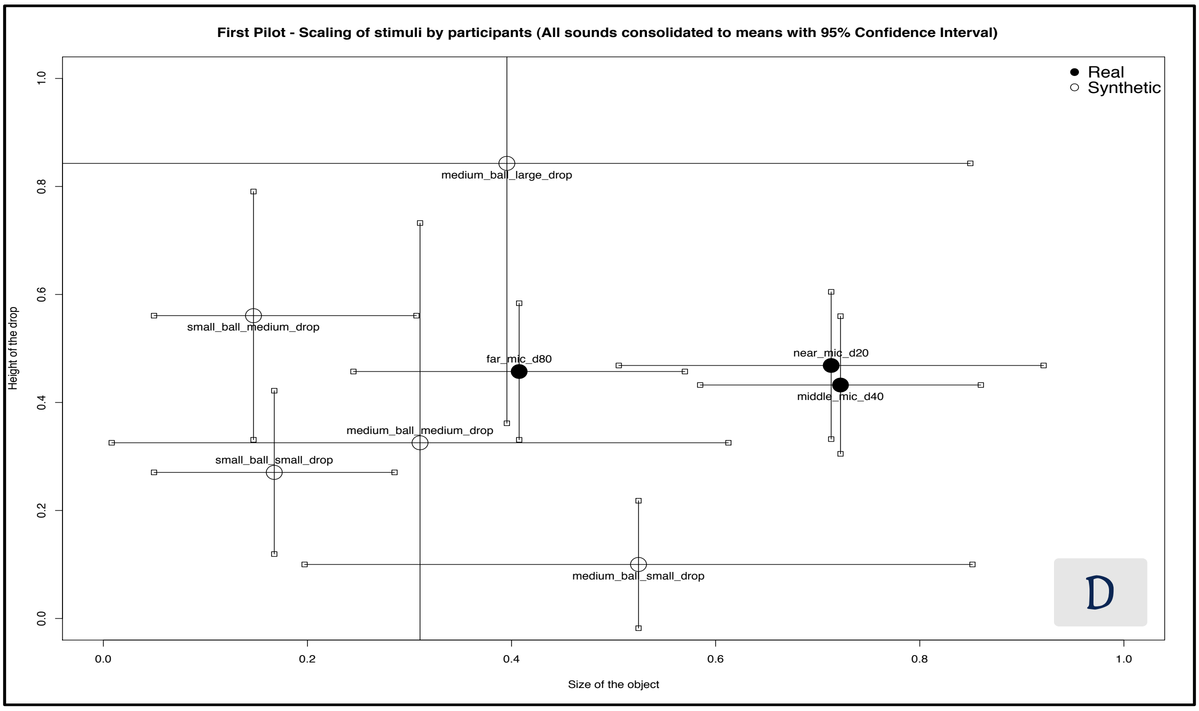


Fig. 4.4: *The perceptual scaling averages of the participants in the first pilot. The top figure, 'C', consolidates the results for the synthesised sounds (for both glass & wood) with trends, 1 & 2, representing the participants scaling of the synthesised sounds. The bottom figure, 'D', represents the 95% confidence intervals showing the range of variation across participants.*

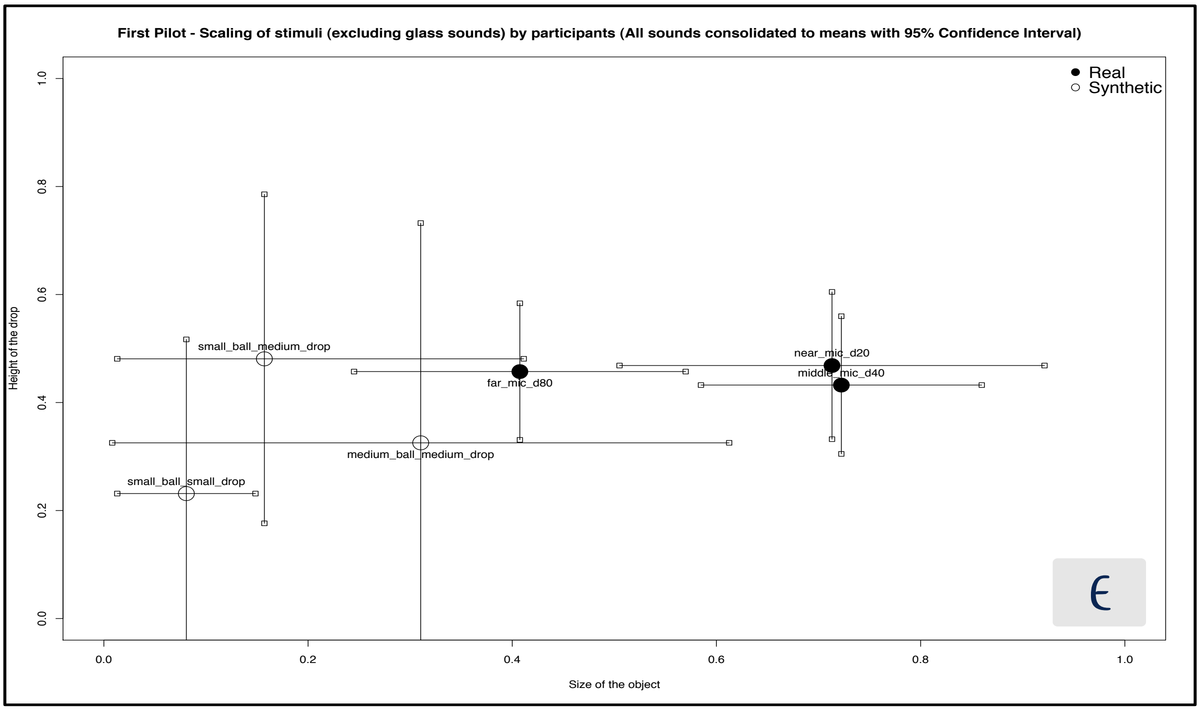


Fig. 4.5: *The perceptual scaling averages of the participants in the first pilot. 'E' represents the 95% confidence intervals but with the synthetic glass sounds removed. This highlights how the wooden synthetic sounds had a range of variation but it can be seen to be much closer to the real sounds.*

Drilling further down into the data and investigating the individual participant's scaling and "tagging" information as shown in Figure 4.6. It shows that two participants (participant two and participant three) made a particular distinction between the real and synthetic sounds. This could indicate that other factors are affecting their scaling judgements. This could be anything from the realism to the recording conditions in the real sounds which are not present in the synthesised sounds as suggested by Carello et al. (2003). The next stage of the analysis looked at how each sound was scaled.

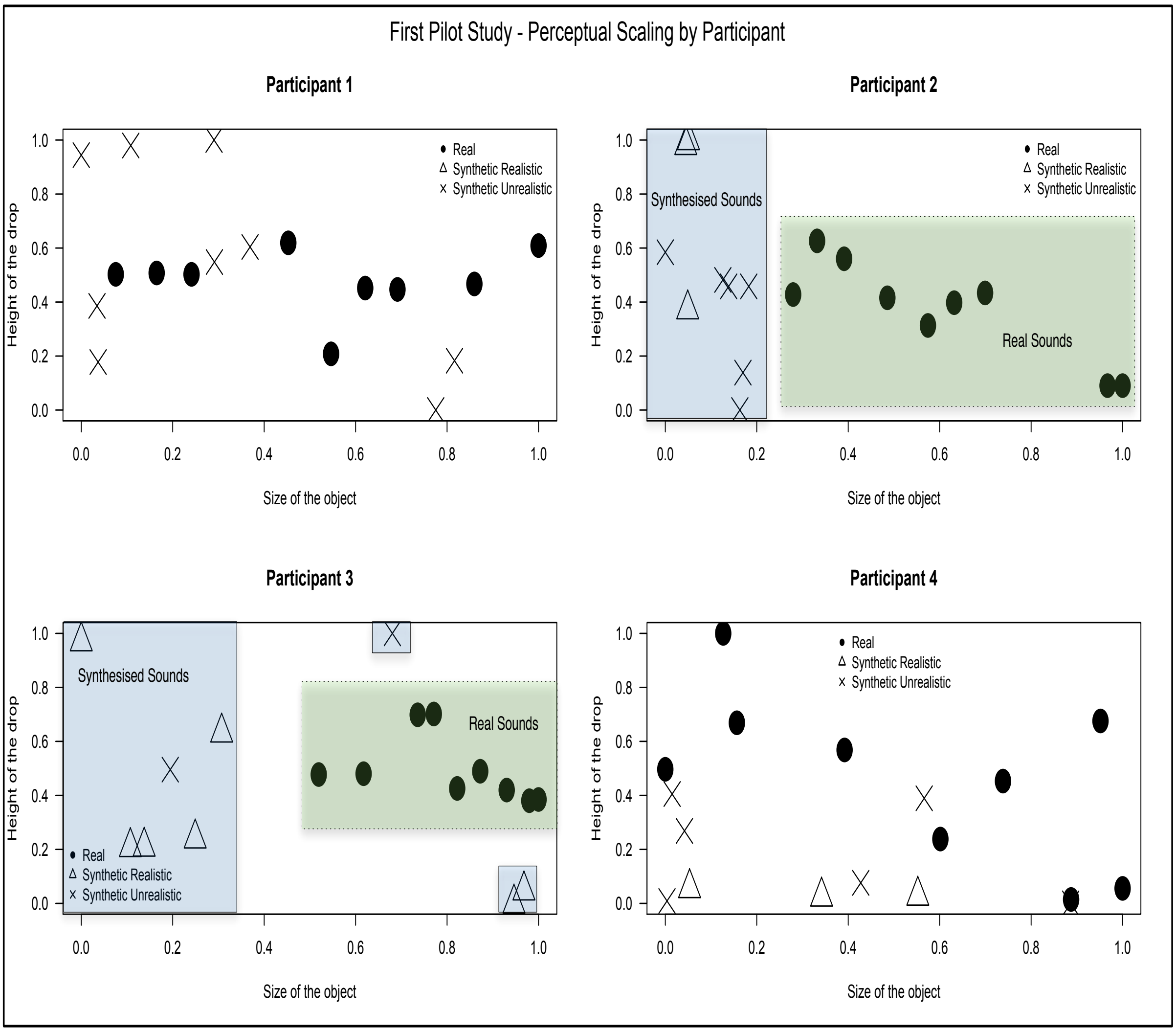


Fig. 4.6: *The perceptual scaling and "tagging" information for each of the participants in the first probe.*

The results of the perceptual scalings and "tagging" choices are shown in Figure 4.7, sorted by stimuli with green backgrounds representing real sounds and blue backgrounds representing synthetic sounds. The height of the real sounds was perceived correctly but that the size estimation varies between participants. The clustering of the height estimation was as expected given the stimuli were all designed to have the same height. The synthetic glass sounds can be seen to have a much greater spread than the synthetic wood sounds. The fact that participants were able to achieve the complex task of scaling sounds in a multidimensional space with two distinct parameters was important. It means that this is a usable method by participants. The difference in size estimation could be due to the different distances in the recordings of the real sounds or due to some other recording conditions present during the recording of the sounds. Interpolation using linear regression after transformation was carried out on the scalings in Figure 4.7 to explore participant's proportion judgements where stimulus magnitudes were estimated. This analysis explored linear regression after transformation as two parameter function in the form of a power law.

$$y = ax^b \tag{4.3.1}$$

This was suggested by Stevens (1974), who made the assertion that participants in a magnitude estimation task use a ratio scale such that x and y are values on this scale, there exists a constant k such that x = ky. In this analysis, k can be found as the slope of the linear regression. Examining Figure 4.7 it can be clearly seen that it does not remain constant for either of the sub groups (real or synthetic) when analysed. The results in this study weakly support signal detection theory and findings by Luce (1990) that contexts such as recording or background noise will ensure that magnitude estimation function does not cleanly match a power function such as that suggested by Stevens (1974). As the numbers of participants and scalings are limited in this study however it is difficult to make any concrete statements about power function relationships and magnitude estimate.

The synthesised sounds are shown with the blue background and the '-pd-X.wav' naming in Figure 4.7. These images show that the participants agreed in the scaling task for at least one of the two dimensions. Two of the synthesised sounds, *w-s-h-s-gl-pd-1.wav* and *w-m-h-m-wd-pd-2.wav*, had noticeably spread perceptual scalings. In order to focus on the variance of the sounds, barplots for both the size and the height were created with annotations representing the individual scaling judgements for the particular scale. Figure 4.7 provides the details to supplement the overview shown in Figure 4.7 with the specific size and height scaling details.

The individual scaling of the stimuli and the ranges of scaling are shown in Figure B-2 in Appendix B.1 and show the perceptual scaling of size and of height presented some interesting results. The synthesised sounds, in particular the sounds of *w-s-h-s-gl-pd-1.wav* and of *w-m-h-m-wd-pd-2.wav*, show how they were spread across the evaluation space, while *w-s-h-m-wd-pd-3.wav* was only marginally spread and was judged as realistic by all the participants. This spread effect across the evaluation space could be related to the lack of realism provided by the synthesised sounds. A good example of a uniform judgement across the participants is *w-s-h-s-wd-pd-4.wav*. The remaining five other synthesised sounds were all judged to be uniform in one of the scaling dimensions. The height scaling was judged as uniform in *w-m-h-l-gl-pd-2.wav*, *w-m-h-s-gl-pd-3.wav*, and *w-m-h-s-gl-pd-4.wav*. The size scaling was judged as uniform in *w-s-h-m-gl-pd-5.wav* and *w-s-h-m-wd-pd-1.wav*. There was a varying degree

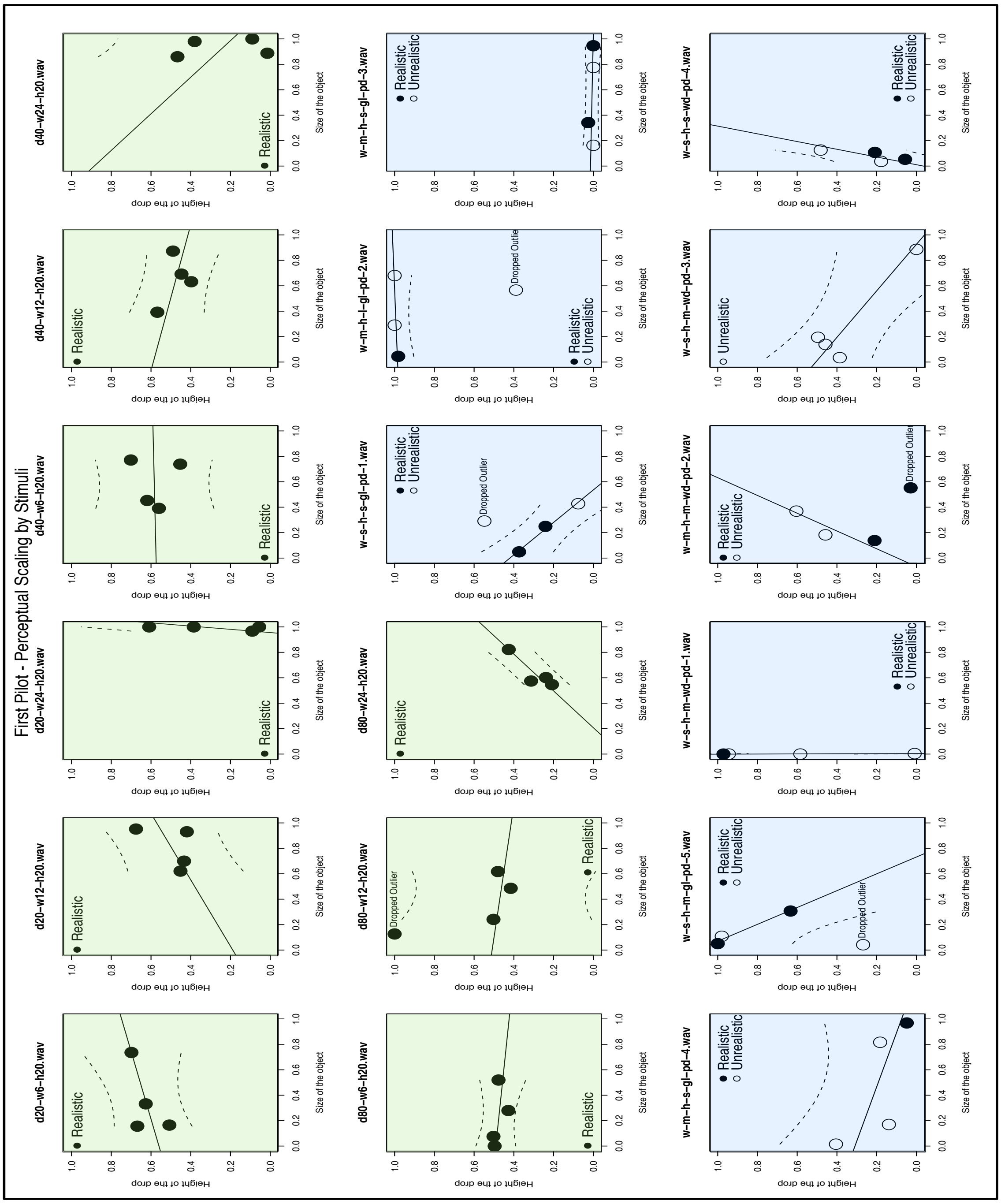


Fig. 4.7: *The perceptual scaling and "tagging" information for each of the stimuli in the first pilot. The dotted lines indicate the 95% confidence intervals.*

of uniformity across the participants in their judgement of the realism of the synthesised sounds. One sound, *w-s-h-m-wd-pd-3.wav*, was deemed as unrealistic by all participants. The sounds, *w-m-h-l-gl-pd-2.wav* and *w-m-h-s-gl-pd-4.wav* were judged as unrealistic by three of the participants, while the remaining synthesised sounds were judged unrealistic by two of the participants. The real sounds were all tagged as realistic by the participants. These results could be due to a number of factors but one possibility is the lack of reverberation in the synthesised sounds.

The results of the questionnaire show that participants found the task to be non-trivial (Question 1) and the ease of use of the application was above average (Question 2). The sounds were judged to be realistic (Question 3) and of a high quality (Question 4). In certain instances the Sonic Browser playback could result in a delay of up to 0.3 seconds in playback, this was found to be acceptable but noticeable by the participants (Question 5). Participants accepted the delay and it had no noticeable effect on the scalings produced. The questionnaire was a 7-point Likert scale with semantic labels and further details can be found in Appendix D.1.

During the debriefing at the end of each participant session, comments were solicited from the participants. This returned several interesting comments and observations. The lack of room acoustics or background recording noise in the synthesised sounds was highlighted by one participant who stated that some of the sounds did not have any “room effect”. Several of the participants found that the “speed of bouncing was directly related to the realism of the sounds”. There were three distinct strategies used for scaling the sounds in the study. These strategies were to “rough order the objects to the height scale first” or to “sort according to size initially” or to “sort them into real or synthesised sounds”.

Results and implications from the first pilot study for the second pilot study

The results of the first pilot study and a review of the procedures and stimuli used suggested a number of minor changes. The success of participants in estimating the height of the dropped ball and of the size of the dropped ball suggested that it would be possible to vary both aspects. The poor scaling of the synthetic sounds parameters may be problematic but a deeper study is required to explore if this is the case. In the second pilot both the size and height of the drop ball were varied. The lack of realism in the synthesised sounds could have been due to the fact that they were limited to a single mode. It is possible that the synthetic wood sounds with a single mode could potentially be less affected than the synthetic glass sounds, as normal glass sounds are quite acoustically rich. The results point to the need for a more focused selection of synthetic sounds with a single material with one or two modes to help to address the issues raised. There is still an open question similar to that raised by Carello et al. (2003) about the acoustic richness of the synthetic signal which may not include sufficient information for the task. The choice of a single material is informed by the work of Giordano (2003) where difficulties were found when material categories were similar. In order to reduce this complexity wood was chosen as the only material, which should be modelled in the next study. In the second pilot synthesised sounds with both one and two modes were used to determine if this would improve the perceived realism of the synthesised sound. Several participants had difficulties where two different materials were used to represent the synthesised ball. In the second pilot the material used to represent the synthesised ball was only wood rather than the wood and glass mix of sounds used in the pilot. Hypothesis 3 and the difference in microphone recording distances did have an effect, however it was chosen to limit this variable to a single fixed distance of 80 cm to reduce the factors being analysed.

4.3.2 The second pilot

The main part of the study stimuli had stimuli with variable height and size unlike the first pilot. This study also used a wider set of synthesised sounds when compared to the first pilot study. The synthesised sounds had one or two modes but used only wood to represent the material of the dropped ball. The following hypotheses were made with regard to the expected results of this study. These are similar to, and built upon, those made in the first pilot study.

Hypothesis 1. *The synthesised sounds would convey perceptually distinct sizes and heights.*

Hypothesis 2. *The scalings of the height and size of the real and of the synthesised sounds would be consistent within the same type of sound.*

Hypothesis 3. *The scalings of the height and size of a synthetic sound is not affected by its perceived realism.*

Hypothesis 4. *The two mode synthesised sounds would convey height and size information as well as recordings of real ball dropping sound events.*

Participants

There were 5 participants (2 males, 3 females) in the second pilot; all were either postgraduate students or researchers. None of the participants had taken part in the first study. In pre-screening for the study, all reported to having no hearing or sight problems and all had prior musical training. The musical training varied but averaged 8 years, with a minimum of 6 and a maximum of 10 years. Two of the participants required glasses for reading, none of the participants reported any hearing problems.

Stimuli

Several of the stimuli used in the first pilot study were included in the second pilot, however the number of synthesised sounds used in relation to real sounds was changed and several parameters of the synthesised sounds were changed. In addition to changing several of the parameters, sounds *w-s-h-s-wd-m2-pd-1.wav* to *w-m-h-l-wd-m2-pd-6.wav* inclusively used the two mode synthesis model as opposed to the other synthesised sounds which used the single mode synthesis model. In the second pilot the synthesised sounds used only wood as the material for the dropped object. Observations during the first pilot study showed evaluation difficulties by participants where both wood and glass had been used as the material. The stimuli used in the second pilot are presented in Table 4.3.

The stimuli are labelled using abbreviations where *w-s-h-s-real-1.wav* represents a recording of a real sound event of a small ball being dropped from a small height. The abbreviations used in for the sounds in Table 4.3 are as follows *w* represents the weight of the ball, *h* represents the distance dropped. The real sounds contain *real* as part of their filename and the synthesised sounds are indicated by *pd*. The synthetic sounds use the additional abbreviations of *wd* to represent wood, of *m1* to indicate a single mode synthesis model, and *m2* to represent a two mode synthesis model.

The particular settings and values used to generate the particular synthesised sounds from the impact model are shown in Appendix C.1 in Table C-2 and in Table C-3, for the single mode and for the two mode model parameters respectively. The value of the parameters were selected during informal trials in order to select the widest possible stimuli set with a particular focus on the perceptual quality of the sounds. The impact model, its main features and the meanings of its parameters are discussed in further detail in Rath and Fontana (2003).

Procedure

The procedure for the second pilot was the same as the first pilot with three distinct stages. The first was the perceptual estimation task, the second was the "tagging" of unrealistic sounds by the participants, and the final stage was the use of a questionnaire on the participant's perceptions of the performed tasks and of the Sonic Browser application. The testing was carried out in an isolation room and stimuli were presented over stereo headphones using the Sonic Browser application.

Results and Observations

The sequence in Figure 4.9 and in Figure 4.10 shows the perceptual scaling (similarity ratings) of the participants. The exploratory nature of this pilot study means that the results are indicative rather than firm findings and where hypotheses are supported; these are merely

Sound file	**Sound type**
w-s-h-s-real-1.wav	Real
w-s-h-m-real-2.wav	Real
w-s-h-l-real-3.wav	Real
w-m-h-s-real-4.wav	Real
w-m-h-m-real-5.wav	Real
w-m-h-l-real-6.wav	Real
w-s-h-m-wd-m1-pd-1.wav	Synthesised- Limited to 1 Mode
w-m-h-m-wd-m1-pd-2.wav	Synthesised- Limited to 1 Mode
w-s-h-m-wd-m1-pd-3.wav	Synthesised- Limited to 1 Mode
w-s-h-s-wd-m1-pd-4.wav	Synthesised- Limited to 1 Mode
w-s-h-s-wd-m1-pd-5.wav	Synthesised- Limited to 1 Mode
w-m-h-s-wd-m1-pd-6.wav	Synthesised- Limited to 1 Mode
w-s-h-s-wd-m2-pd-1.wav	Synthesised- 2 Modes
w-s-h-m-wd-m2-pd-2.wav	Synthesised- 2 Modes
w-s-h-l-wd-m2-pd-3.wav	Synthesised- 2 Modes
w-m-h-s-wd-m2-pd-4.wav	Synthesised- 2 Modes
w-m-h-m-wd-m2-pd-5.wav	Synthesised- 2 Modes
w-m-h-l-wd-m2-pd-6.wav	Synthesised- 2 Modes

w - s/m - size of object h - s/m/l - height of drop of object
wd - synthesised wood
m(1/2) - number of modes used in synthesised sound model
real - recording of real dropping sound
pd - 1/2/3/4/5 - PureData synthesised sound settings

Table 4.3*:* List of the real stimuli and of the synthesis stimuli used in the second pilot.

qualitative findings. Both of the figures are broken into two sub-images with particular aspects of the data highlighted to focus on particular aspects of interest. In image ‘A’ of Figure 4.9, an overview of the scalings are shown by sound type and indicates support for Hypothesis 1. The trends, 1 & 2 show the expected scaling for real sounds where as the height of the drop increases, this is reflected in the scalings of the participants. The second image ‘B’ consolidates the synthesises sounds together to focus on the perceived size and height. It shows a fair match between the height and the size scalings from the participants with regard to the synthetic sounds.

The trends, 1 & 2 show that while there is some degree of scaling in trend 2, this is not reflected in trend 1. Trend 1 shows no discernible pattern. Trend 2 shows the scaling trend for the medium sized synthesised sounds, however this trend would indicate that participants have inversed the scaling with ball size being scaled by height. The result is not strong enough to support Hypothesis 2. These results show that there is a trend with regard to the height of the drop for real sounds and apart from this there are no other clear scaling patterns. In image ‘C’ of Figure 4.10 the 95% confidence intervals for the consolidated sounds are shown. In order to focus more on the differences between the one and two mode synthesis sounds, the 95% confidence intervals for this subset of sounds is presented as image ‘D’ in Figure 4.10.

Hypothesis 4 was not supported and the results show that even the more complex two mode synthetic sounds do not convey information about size and height of drop as well as real recording of such events. The second pilot showed that two mode synthetic sounds used did “*cartoonified*” (Fernström, Brazil and Bannon, 2005) to a certain extent. However, more study on what are the key salient perceptual attributes for these types of dropping sounds are required in order to produce synthetic sounds that can convey information as well as real sounds to listeners.

The perceptual scaling and “tagging” information sorted by participant for the second pilot is shown in Figure 4.8. The “tagging” information is illustrated using different shapes. The black circles are used to represent the real sounds. These were all judged as realistic by all the participants, while the triangle outline and the “X” symbols were used to represent the synthesised sounds, respectively judged as realistic and unrealistic by the participants. In Figure 4.8, the first three participants show some clear distinctions in scaling between real and synthesised sounds. However, only participant three limited the scaling of synthesised sounds, where they alone felt that all the drops of the synthesised sounds were small.

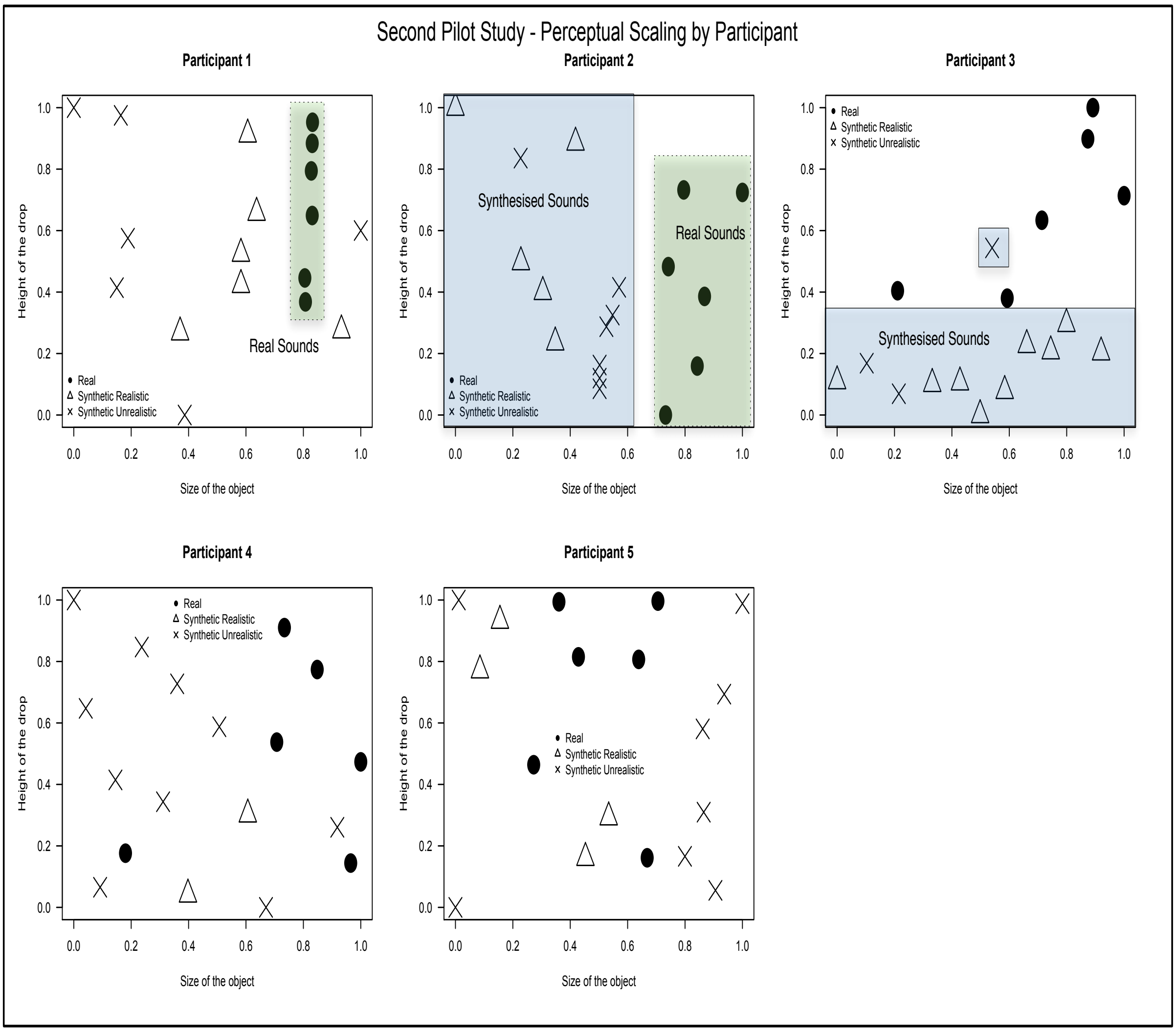


Fig. 4.8: *The perceptual scaling and "tagging" information sorted by participant for the second pilot study.*

Examining the realistic and unrealistic sounds indicates support for Hypothesis 3 that a sound's realism does not affect its perceptual scaling. An observation from both the pilot and the second pilot was that even when the real sounds consisted of dropping steel balls sounds, the participants still referred to the balls as being made from wood. This results indicate that there is a larger influence from the surface material for particular cases which follows the suggestion of Giordano and McAdams (2006). In a similar manner to the first pilot, the classification made by participants was grouped according to the type of sound. In particular, two participants (participant 1 and participant 2) only made minor judgements on the size of the real sounds and their judgements focused on the height aspect of the sounds.

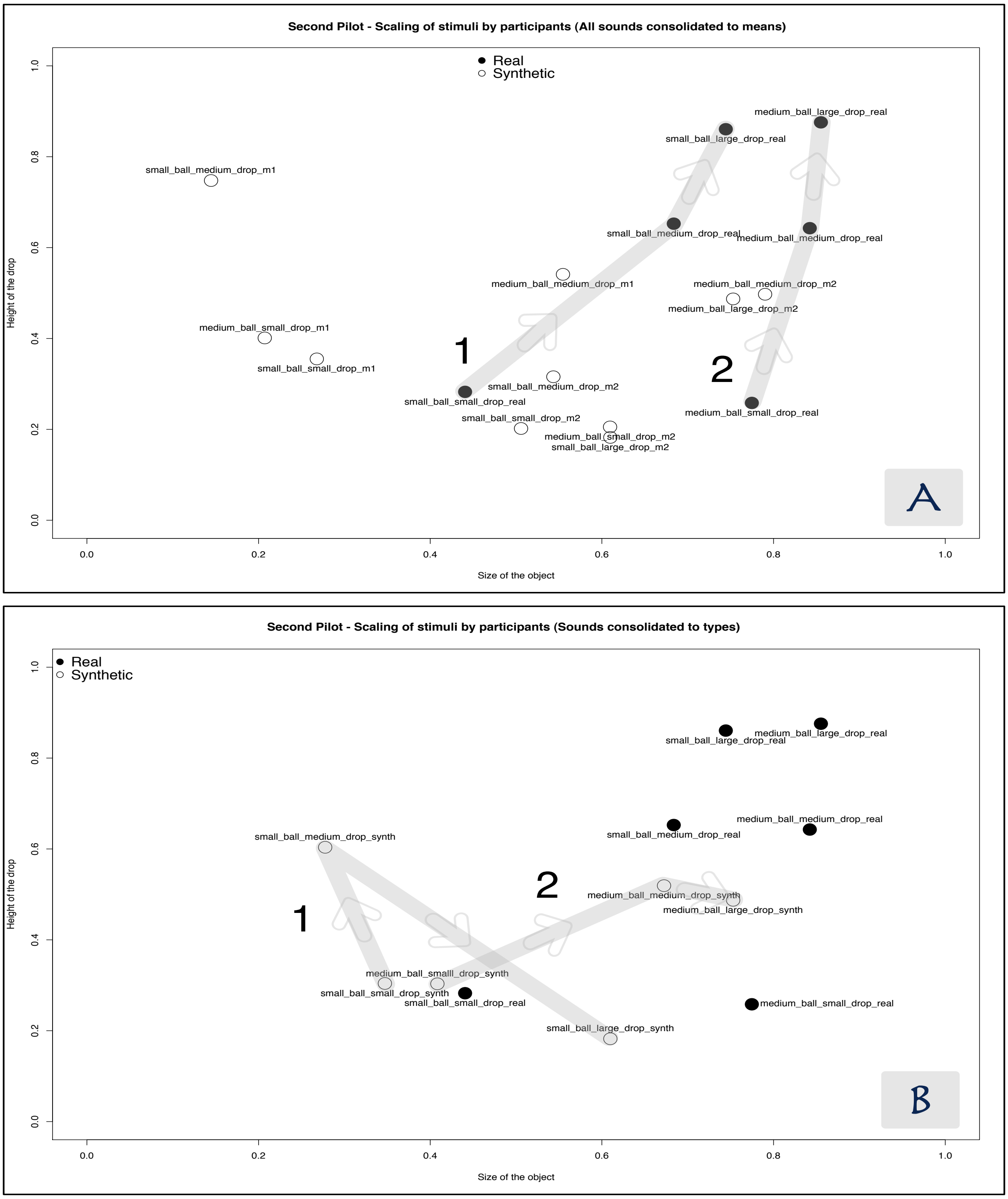
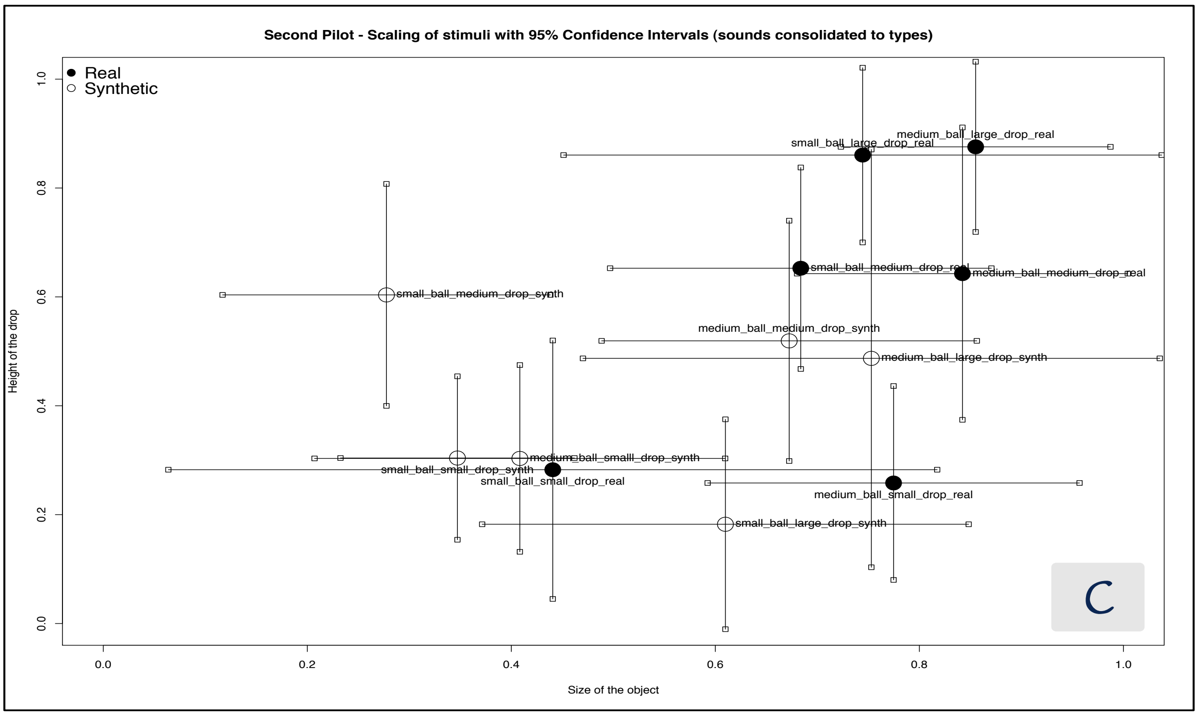


Fig. 4.9: *The perceptual scaling averages of the participants in the second pilot. From top to bottom has 'A' which represents the overview and 'B' the synthesised sounds consolidated together (1 & 2 mode). The trends, 1 & 2, in 'A' represents the scaling of the real sounds in the study. In 'B', the trends 1 & 2 represent the scaling for the consolidated synthesised sounds.*



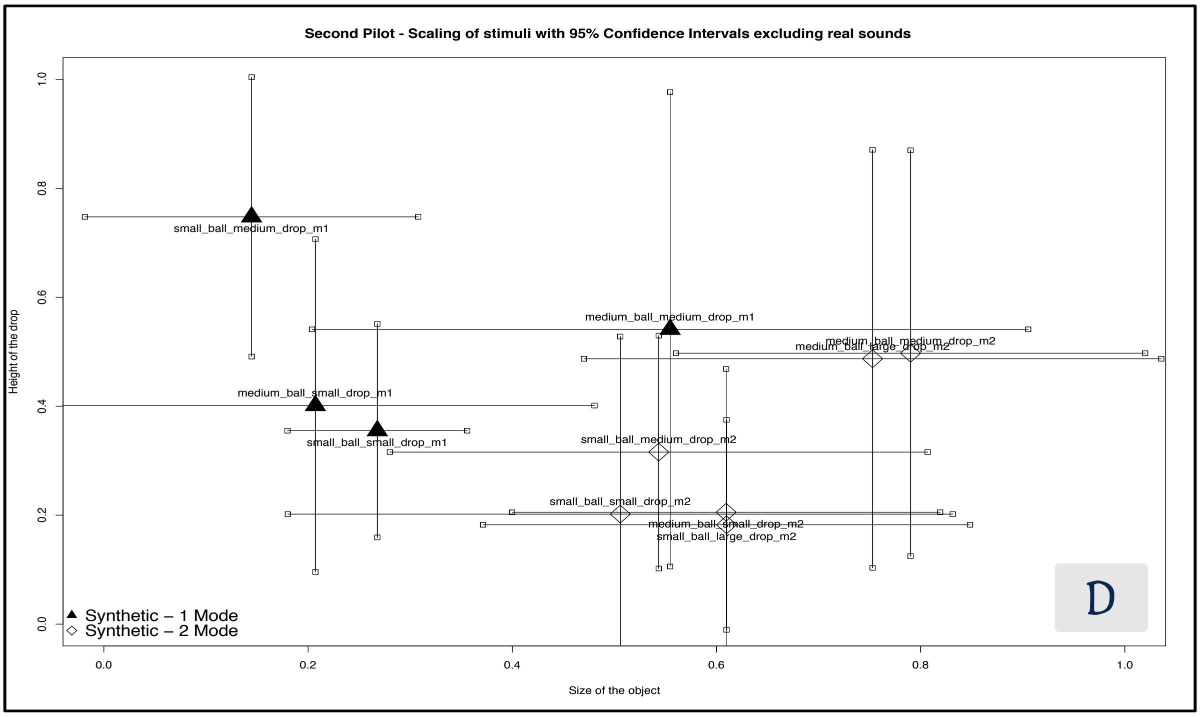


Fig. 4.10*: The perceptual scaling averages of the participants in the second pilot. From top to bottom has 'C' the 95% confidence intervals for the sound and 'D' the 95% confidence intervals with the real sounds removed.*

The results of the perceptual scalings and "tagging" choices are shown in Figure 4.11, sorted by stimuli. The estimations vary between participants but seem to be consistent on at least one of the properties in many of the stimuli. Interpolation using linear regression was carried out on the scalings in Figure 4.11 to explore participant's proportion judgements where stimulus magnitudes were estimated. In a similar finding to the first pilot study, the results did not cleanly match a power function. As the numbers of participants and scalings are limited in this study however it is difficult to make any concrete statements about the possibility of a power function relationship for magnitude estimate.

In this part of the study, the main focus was on the synthesised sounds and the individual scaling of the stimuli and the ranges of scaling for the stimuli, as shown in Figure B-4 in Appendix B.1. In a similar result to the first pilot, the users did agree in at least one of the perceptual scaling dimensions. Four of the stimuli were judged uniformly by all participants, *w-s-h-m-wd-m1-pd-1.wav* and *w-s-h-s-wd-m1-pd-5.wav* had strong unity in both dimensions while *w-s-h-s-wd-m1-pd-4.wav* and *w-s-h-m-wd-m2-pd-2.wav* had a fair degree of uniformity in both dimensions but to a somewhat less degree. Four of the stimuli were judged uniformly in one dimension (*w-s-h-m-wd-m1-pd-3.wav*, *w-m-h-s-wd-m1-pd-6.wav*, *w-s-h-s-wd-m2-pd-1.wav*, and *w-s-h-l-wd-m2-pd-3.wav*). The last four stimuli were found to be dispersed with the stimuli (*w-m-h-m-wd-m1-pd-2.wav*, *w-m-h-s-wd-m2-pd-4.wav*, and *w-m-h-l-wd-m2-pd-6.wav*) being somewhat dispersed while the stimuli, *w-m-h-m-wd-m2-pd-5.wav*, was found to be dispersed. In this realism task of this study, none of the synthesised stimuli were tagged as unrealistic by all the participants. The maximum consensus amongst participants was by 3 participants. Similarly to the first pilot, all of the real sounds were judged to be realistic.

The questionnaire, a 7-point Likert scale with semantic labels and further details can be found in can be found in Appendix D.1. The results of the questionnaire show that participants found the task to be non-trivial (Question 1) and the ease of use of the application was below average (Question 2). The sounds were judged to be only of average realism (Question 3) and of a average quality (Question 4). The issues with the synthesised sounds are potentially due to their lack of room acoustics and to the presence of a "buzz tail" at the end of a number of the two mode synthesised sounds. This was stated verbally as being distracting by one participant. This was due to a technical issue in the synthesis model. In certain instances, the Sonic Browser playback could result in a delay of up to 0.3 seconds in playback, this was found to noticeable and annoying by the participants (Question 5). The delay did not have any noticeable effect on the scalings produced, it a source of irritation to one of the participant's but did not effect their performance.

During the debriefing at the end of each participant session, comments were solicited from the participants. Several participants commented that they found it problematic to decide upon a scale and "whether to start with big or small sounds". A number of the participants found

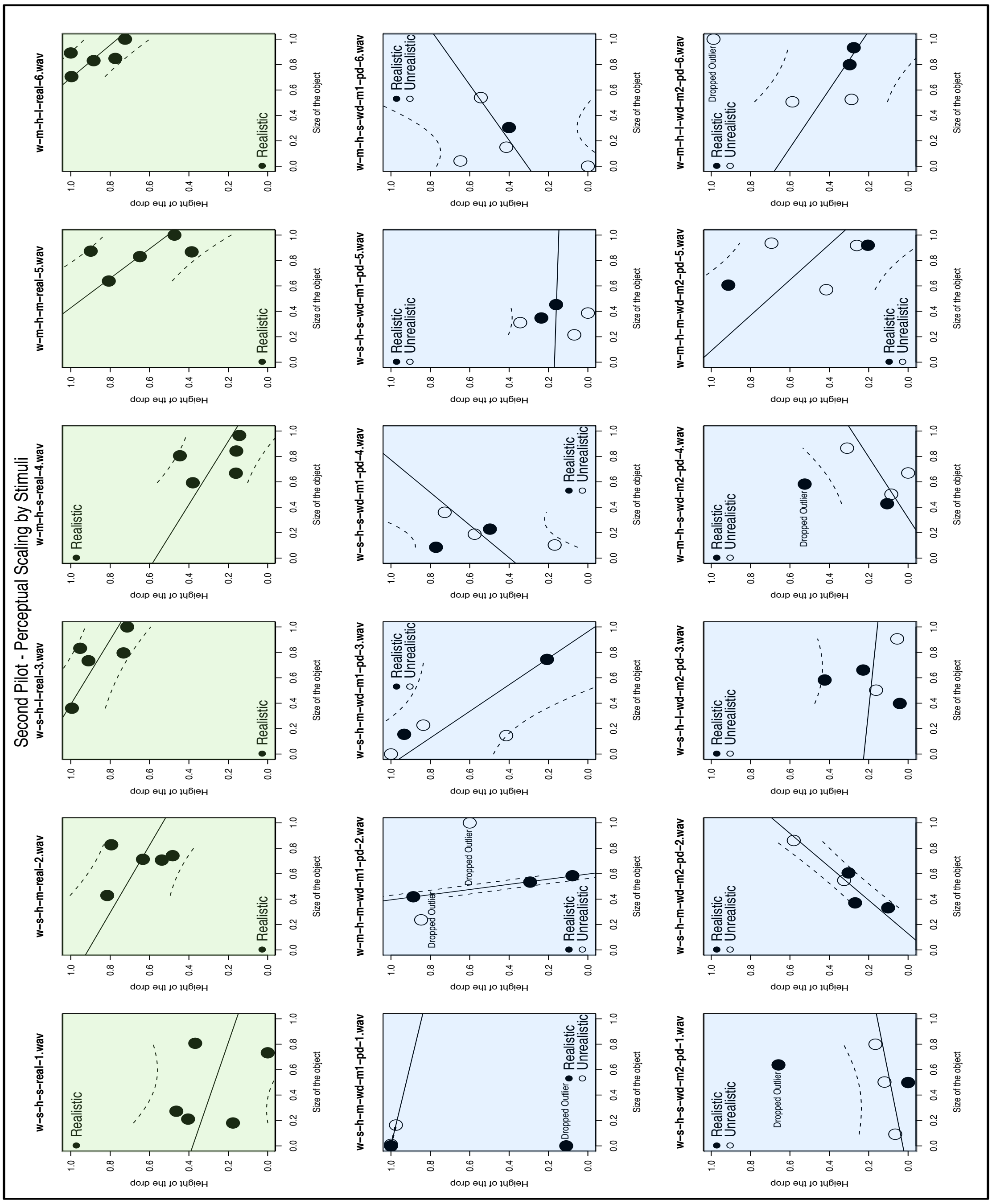


Fig. 4.11: *The perceptual scaling and "tagging" information for each of the stimuli in the second pilot. The dotted lines indicate the 95% confidence intervals*

that it was “much easier to judge size over height”. The “speed of repetition as a characteristic of height” was found to be a useful concept in classifying the height of a particular sound. One issue that occurred in the study and not in the first pilot was a “metallic zips distraction/confused the classification of sounds”. This referred to the ending of a number of the two mode sounds as mentioned previously. Several of the participants commented that a “detailed comparison without reference points was very difficult and would be much easier in a single dimensional space”. This point highlights the cognitive load in the task of scaling the sounds within a 2-dimensional space. The aura functionality of the Sonic Browser was found to be useful as “it allows me to see if it is higher or which is lower by using pitch. The aura now gives me a comparison for similar sounds”. The aura functionality was not instrumented so it is not possible to determine if this effected the results excluding the debriefing comments made by the participants. The different materials and surfaces within in the stimuli set were commented on by the participants who stated that they found the “different surfaces are very noticeable”. The final and important observation was from two different participants who felt that the stimuli set consisted of “three divisions (small, medium, and large) and that it was very hard to compare between divisions but it was easy to compare within a division”. On further discussion, they stated that the divisions were the three types of sounds used within the stimuli set, real, synthesised with one mode, and synthesised with two modes.

Results and implications from the second pilot

The results showed that participants were able to scale sounds using a 2-dimensional scale, however the results suggest that real over synthesis sounds should be used for Auditory Displays, given the wider perceptual variances found in the scaling of synthetic sounds. In order to lessen difficulties when comparing synthesised sounds, stimuli should be limited to a single material. The material and the number of modes used to synthesise it may also affect the perception of the sound. The addition of reverberation or similar “room acoustics” should be explored for synthesised sounds to improve their realism. Participant’s comments suggested that a number of reference points be added to assist participants in making their comparisons. Holyoak and Mah (1982) have suggested that adding “reference points” can assist the ability of people to spread their finite discriminatory ability across a given magnitude range. This is an area that should be investigated in future studies to see if referenced points can improve this approach. The findings and implications from both parts of this study show the potential this approach as one alternative to classical techniques.

The rationale for two small pilot probes rather than a large study was to find out if the new technique could work and to see if the available synthetic sounds were ready for more detailed exploration and testing. The pilots showed that the new technique using a computer for sorting and scaling of the stimuli was engaging and offers a viable alternative approach when compared with pairwise comparison. The synthetic sounds and the results of their scaling highlight that more work is needed on these types of synthesised sounds prior to larger scale experimental studies. Issues such as the lack of room acoustics, the presence of a “buzz tail”, and the possible lack of “acoustical richness” are possible factors when effected the synthetic sounds used in these pilot studies. The lack of synthesised sounds, which adequately dealt with these issues and the results of the studies in this chapter point to the need for these issues to be addressed prior to conducting new studies.

The results from both studies show that there is a large perceptual variance in interpretation of the perceived physical parameters of synthetic sounds. It highlights the need for more work on this type of synthetic sound, as the variance is much larger than with real sounds. It suggests that real sounds should be used in favour of synthetic models for Auditory Displays. The future potential of these models with computational and real time control aspects, which could deal with the real time interactive sound of interactive Auditory Displays suggests that more detailed work is to develop more perceptual robust models. There is more work required if these types of models are to be used in Auditory Displays. Designers need to be aware of this issue as the use of real recorded sounds may require a large library to ensure that there is a sufficient number to cover a range of sizes and materials. Additional recordings may be required to provided alternatives to each of the recorded sounds to help ensure that a single sound or group of sounds are not repeated frequently as this could be annoying for listeners.

4.4 Limitations of the Pilot Studies

In the following paragraphs, a number of issues regarding the limitations of the studies presented in this chapter are presented. They are discussed here to highlight potential areas for consideration prior to use of this type of experimental method. It also raises some of the potential issues that need to be considered when using a particular method.

4.4.1 Multidimensional Scaling Approach used in the Pilot Studies

In this chapter an approach using multidimensional scaling and realism judgements was presented as a method to gain an understanding of the participant’s judgements regarding synthesised sounds. This study and the methods presented in this chapter provide an understanding of how participants scaled sounds, both real and synthesised. The stimulus representations in this study are based on on measures of stimulus similarity (Shepard, 1974, 1980) using a dimensional approach (Goldstone, 1999). The dimensional approach is where stimuli correspond to a point in a multidimensional space, and the distance between their representative points measures the similarity between any two of the stimuli. The participant’s representations in this multidimensional space, may result in different patterns of similarity when analysed using different distance metrics (Shepard, 1991). In this study it may be the case that participants gave more 'ordinal' judgements meaning for sounds that they were able to state which sound was higher or larger but were not able to quantify the higher or larger amount.

4.4.2 Use of Questionnaires in the Pilot Studies

The results from the questionnaires in this chapter supported the experimental findings but the use of questionnaires in a study needs to be considered. It is important to ensure that the phenomenon under investigation has not simply been called into existing by having been enquired about through post-event measures (see Slater (2004) for a more detailed exploration of this issue). The use of questionnaires in this study provided quantitative measures. They should be combined with interview or debriefing data. This type of combinational approach can limit the potential for the phenomenon being called into existence by enquiring about it. In the case of comparing real sounds, there is a stock of experience for participants against which to judge a particular experience which should assist in gauging the comparisons of synthesised sounds. However, it is possible best to use questionnaires as "hypothesis generators" (Slater, 2004, pp. 492) rather than mechanisms for producing conclusions.

4.4.3 Population used in the Pilot Studies

These studies used university students, which raises a question about what is a good representative choice for the general population. This is a widely debate argument in many disciplines (Peterson, 2001, Sears, 1986). The discussion here is only aimed at providing a cautionary reminder. This is relevant to this thesis as a warning not to draw unfounded inferences or make generalities from the results presented. In particular the work of Calder et al. (1981) points to the purpose of the research as determining the suitability of university students for use as the general population of the study. Calder et al. (1981) divides research into "*effects application*" and into "*theory application*" research. The two types of research have different goals with effects application aiming to obtain results that can apply directly to a real world situation, while theory application aims to obtain scientific theory and making it applicable for future studies in the real world. In the case of this thesis and its studies, the aim is theory generalisation and the calibration of methods for use by Auditory Display Designers. Readers are referred to Bech and Zacharov (2006) for a more detailed introduction into the selection of suitable subjects for auditory experiments.

4.4.4 Linear Regression used in the analysis of the Pilot Studies

A form of simple linear regression was used to interpret the scaling of sounds to explore the stimulus magnitude of participant's proportion judgements. In the case of the current study, it is likely that the other contexts are affected the scalings such that a multiple linear regression or more advanced technique would be required to gain a better understanding about the magnitude estimations used by participants for their scalings within this study. The result was limited by the number of participants and may have been affected by other contexts.

4.5 Applying results of the studies in practise to three hypothetical domains

In Section 2.5, three domains or hypothetical Auditory Display applications were introduced. As part of the review of each method presented in this thesis, this section presents a short review of how this particular method can help address certain design challenges for Auditory Displays.

Mobile device Auditory Display for messaging/notifications Taking an idea from the Shoogle application (Williamson et al., 2007) and the use of shaking to determine the number of messages, one sound design could be the use of marble sounds as the sounds used for this hypothetical scenario. The method in this chapter would allow for the physical characteristics of a sound whether it is dropping or hitting to be explored with regard to the perceived size of marble involved. The mappings of this type of sound could be to message size, to the receipt time of the message, or to the priority of the received message. The value of the presented method is that it provides an approach for determining if the expected mappings of the sound, whether natural or synthesised are actually the same or similar to those as perceived by the user. This provides a critical validation of the sound mapping and can highlight a possible inappropriate mapping.

Network or processing monitoring using an Auditory Display The naturalistic soundscape in an Auditory Display has been used for stock monitoring by Mauney and Walker (2004). Taking this idea and suggesting river flowing sounds as part of the sound design, this method can be used determine if the desired meaning was easily conveyed to users using this mappings of events or amounts. The sounds can be rapidly compared and scaled within the listener's perceptual space. This can assist the designer in ensuring what the listener perceives as the largest sound is indeed the sound used by the designer and so on depending on the mapping being explored. Complex monitoring environments such as factory or process monitoring could also benefit from ensuring better identification and less confusion for sounds that may occur concurrently. If these sounds are being used in such a context, the user can ensure they hear more of the information and should perform better as they are more informed as to the current state of the process.

Interactive table surface and its related Auditory Display The increasing growth in interactive surfaces such as the Reactable (Jordà et al., 2005) and the possibilities for expanding Auditory Displays within this type of interface shows another scenario where this method can provide the designer with useful feedback. One potential auditory design would be to use scratching or friction sounds to represent the gestural interactions on the surface. The visual objects represented on the surface could also have related auditory aspects, i.e. one object may be related to a metal sound so that as it is grasped or dragged it sounds like a piece of metal being interacted with. Details such as file size, priority, file type or other parameter could be linked to the perceived physical size of the sound. This type of auditory interaction and in particular, the sounds used for the interactions could be examined using the scaling method to ensure that the perceptual mappings perceived by the user are the same or similar to those intended by the designer. The scaling method allows for the rapid exploration of a number of possible mappings to select the best mapping between parameter and sound property.

The three domains above have shown a small sample of possibilities using this technique, in particular with regard to perceptual scaling of sounds. The study also highlights the further possibility of tagging certain sounds to indicate favourites, realism, or other binary properties that the designer might additionally wish to explore. The next section details the specific contributions from this chapter to Auditory Display. This section has shown how broadly applicable this approach can be to one common type of design problem faced by Auditory Display designer, that of ensuring the intended mapping of the set of designed sounds is actually that which is perceived by the listener.

4.6 Conclusions

The work in this chapter used both synthesised and real sounds to examine the physical and perceptual aspects of a particular type environmental sound, that of a dropped ball. Gygi and Shafiro (2007, p. 3160-3161) asserts this type of research can be used as a foundation for "*developing perceptually motivated synthesis models*" and for "*designing new sounds for representing complex information structures*". The indicative results from the pilots concentrated on understanding how the sounds used were scaled by listeners in comparison with their physical dimensions, how realistic the participants found the sounds, and explored one and two mode synthetic sounds. The real sounds showed consistent scaling trends and were easily scaled along the 2-dimensional perceptual dimensions of the height of drop and of the size of object. The synthetic sounds were found to be unrealistic in the majority of cases and had a wider variance in their scaling of perceptual dimensions. The best results of the synthesised sounds were simple one mode wood-like material sounds, suggesting this should be the starting point for further developments. These synthesised sounds were found to be the best as they did not contain any "buzz tail", which irritated and distracted listeners. The overall results indicate that current synthetic sound models are not yet ready for wider deployment in real Auditory Displays. The advantages of synthetic sounds and the existing trends in research do suggest they are a potential area for future inclusion in Auditory Display but this will require additional work to better link the perceptual dimensions of the sounds to what is perceived by listeners. Clear results from the studies are that mixing synthetic sounds, which contain different modes, causes scaling difficulties for participants. The results of the studies did show that sound realism did not prevent participants from extracting and evaluating meaningful information from a synthesised sound. The differences in scalings between one and two mode synthetic sounds, between the materiality of the synthetic sounds (e.g. glass versus wood), and the exploratory nature of the pilot studies limit the wider application of the results. These results point to the need for studies exploring the modal differences (e.g. 1 versus 2 mode) in the perceptual scaling of synthetic sounds and for newer models with more acoustical richness (Carello et al., 2003).

The exploratory results in this chapter do not fully support the possibility for sound cartoonification (Fernström et al., 2005), where certain perceptually salient features of a sound are taken and exaggerated. This exaggeration is the auditory equivalent to the visual exaggerations used in children's cartoons. The area of sound cartoonification and of synthesised sound models has been discussed in greater detail by Rath and Fontana (2003). Examining the real and synthesised sounds and their results has shown that participants were able to clearly distinguish between the materials and the events involved. This is an area for a more detailed exploration as other dimensions not examined in the studies could influence the results from this chapter. These observations are based on a relatively small stimuli set and a limited number of participants, it is difficult to generalise these results but the trends found are worthy of further in depth studies. The success of the method for exploring the perceptual scaling of sounds and its applicability within a design process is a major addition to the methodological toolbox of Auditory Display designers. The method can allow designers to use a computer based interactive sorting – scaling interface for participants and allows for the rapid scaling sounds within a 2-D perceptual space. The studies in this chapter have show that synthetic sounds created using the same modelling approach as those used in the studies will allow listeners to extract some perceptual information from the sounds.

The approach detailed in this chapter benefits researchers creating new sound models, in particular those designed at levels similar to the basic event, or the derived process, or the simulated examples from the Sound Object taxonomy as shown in Figure 2.3 in Chapter 2. For example, a researcher creating a splashing model could use this approach to determine the salient features the listeners associate with splashing and what criteria they used to organise this type of sound. The approach will allow researchers to explore this type of situation where more one variable changes and should help participants keep a better overview of the entire stimuli set. Future work should focus on exploring if this approach does provide a better overview, exploring means for deeper verification of perceived perceptual properties of sounds, and how the approach from this chapter deals with issues such as fatigue and comparison inconsistencies when compared to traditional approaches. A comparison between pairwise comparison and the interactive sorting – scaling approach presented by the Sonic Browser is another area for future studies. The exploratory results of these studies provide positive indicators that the Sonic Browser did help in reducing fatigue but this requires further exploration.

The results presented in this chapter suggest that it is may not be possible to use synthesised sounds of impact/bouncing events for an Auditory Display to reliably provide quantifiable information in multiple dimensions. These studies provide a better understanding of this type of synthesised sound by providing a method to verify parameter-based synthesis models. This verification can help ensure the creation of Auditory Icons (either real or synthetic) with acoustic commonalities to the objects they are designed to represent and that are able to successful communicate quantifiable information to users. This type of synthesised sound could allow an auditory designer to use efficient artificial sounds that are at least as informative and as useful as real world sounds. However, the results of this chapter suggest that while parametrically controlled synthesis models may allow for more computationally efficient and dynamically controllable sounds, more work and better perceptual mappings could ensure their use in Auditory Displays. The wider interaction space these models provide is a strong argument for wider research into this type of synthetic sound and the problems highlighted by the work in this chapter.

This method is important for Auditory Displays as it provides an understanding of how people hear and think. These are the first step towards the design of Auditory Displays that communicate information successfully. Chapter 3 highlighted event similarity as an important factor in grouping and classification. This indicates in many groupings and their related classifications that an implicit knowledge of both the object and action creating the event is used to judge the similarity. This recognition of sound events and type of implicit knowledge has been suggested in ecological psychoacoustics research (Gaver, 1993a,c). In this chapter the focus was concentrated on the grouping aspect but the following chapters will further explore issues of source identification and of sound identification with regard to the classification aspects.

This study has looked at research addressing one of the three tasks suggested by Walker and Kramer (2004) for Auditory Display research that of "*the simple perception of sounds*" in the environment. The next chapter focuses on a second part of this task, the parsing of sounds into their sources or streams. The third task is more complex and is focused at understanding cognitive and associative processing or how a listener makes meaning. This task is important as it ensures the Auditory Display's sounds and their intended meaning is actually what is conveyed to the listener. The work in Chapter 6 presents an approach which can be used to address this task.

The next chapter presents an investigation into concurrent Auditory Icons and provides one method for Auditory Display designers to address a problem posed by Walker and Kramer (2004, p. 167) who emphasised the importance of knowing "*how individual sounds will blend in or stand out from the growing acoustic crowd*". The method provides both identification measures and a confusion metric, which helps designers choose the appropriate sounds for concurrent auditory presentation. In the next chapter, this method shows how a designer can ensure the best select Auditory Icons that improve their separability from other sounds while maintaining their own distinctiveness.

Chapter 5

Understanding Concurrent Auditory Icons: Investigating An Object-Action Duality For Improving Sound Identification

Words Commonly Used To Describe Sounds
BANG CLINK JANGLE RUSTLE THUMP
BARK CLUCK KNOCK SCREAM THUNDER
BELLOW CLUNK MEW SCREECH TICK
BLARE CRACK MOAN SHRIEK TINKLE ....
- Peterson & Gross, Handbook of Noise Measurement, p. 217

Auditory scene analysis (Bregman, 1990), auditory alarms (Stanton and Edworthy, 1999) and Brewster's work on Earcons (2002) have shown that truly artificial Auditory Displays with symbolic mappings can work successfully. There are a lack of methods available to help designers perform empirical investigations into the everyday sound scene, in particular of Auditory Displays using everyday sounds. Concurrent audio presentation offers advantages such as increased bandwidth and faster presentation but can suffer from disadvantages such as the sounds can interfere with each other, confused interpretations, or perceptual masking. This can occur even where sounds are played in a sequence. Moore and Hedwig (2002, p. 331) point out that the segregation of the sounds is based on the "*degree of perceptual difference*" between them. They point out that even a small perceptual difference "*may be sufficient to improve performance*" for tasks which require some degree of sound segregation. Moore and Hedwig (2002, p. 331) further suggest that larger perceptual differences could produce "*obligatory stream segregation*". If Auditory Icons are to be used in Auditory Displays as the communication mechanism, they need to be able to keep pace with interactions as they occur. This can be problematic if only sequential sounds are used as the system will wait for a sound to finish before playing the next sound or may end up playing the sound for a just completed interaction, either of the situations will not provide any advantages to a user. The use of simultaneous sounds is explored in this chapter as one approach to dealing with these issues. This work goes a step further by asking if object and action properties can be used to classify sounds prior to their use and if this is effective in improving the identification of concurrent Auditory Icons. The method presented can help Auditory Display designers rapidly investigate the suitability of a set of sounds for concurrent audio presentation. A summary of how the methods in this chapter can benefit designers is given in Table 5.1. Designing Auditory Displays where the sounds have larger perceptual differences should assist in the segregation of sound. This chapter explores an approach that looks at the perception and prior categorisation of Auditory Icons, and at the source and the sound identifications of listeners to improve their identification. Understanding how everyday sounds blend together gives a better understanding of the identification of an everyday sound scene. Previous work by others on Earcons (Brewster, 1994, McGookin, 2004), on auditory scene analysis (Bregman, 1990), and on sinusoidal tones (van Noorden, 1975) have all proven to be valuable in highlighting particular aspects of sound that are relevant for the design of Auditory Displays. The work in this chapter provides methods which present concrete measures that can be used to determine the identification and confusion of sounds when concurrently presented. Similar research on everyday sounds, in particular on complex Auditory Displays using concurrent sound presentation has not been as comprehensive and the work in this chapter provides a starting point from which to address this deficiency.

Informational	Identification measures of concurrent everyday sounds. Confusion metric for concurrent everyday sounds.
Inspirational	Metaphors and descriptors in the listener's own words of individual sounds. Descriptions and interpretations of how listeners perceived concurrent sound combinations.
Difficulties	Masking of sounds in certain conditions. Potential perceptual merging of similar sounds in certain conditions. Low degree of sound categories overlap in third study may have influenced results.
Contributions	Method for identification of appropriate combinations of concurrent sounds. Verification that overlapping categories of action / object cause identification difficulties. First exploration of concurrent everyday sounds for use in Auditory Display.

Table 5.1*:* Summary of informational and inspirational aspects of the methods and techniques from this chapter and also the difficulties and contributions from this chapter.

Prior work in Auditory Display has often concentrated on situations where a single event or message, sometimes complex, is being conveyed but with increasingly more complex technology and interfaces there is a growing need for the ability to convey multiple events or messages simultaneously. The lack of guidelines and research in the area of conveying multiple events or messages using Auditory Icons led us to investigate concurrently presented Auditory Icons. In the context of this thesis, concurrent Auditory Icons are defined as the playing of several Auditory Icons together and simultaneously, to build more complex and compound Auditory Icons. This is inspired by how real sounds work in the world. Researchers such as Papp (1997), Brewster (1994), and McGookin and Brewster (2004) have investigated the concurrent audio presentation issue in a more formal manner. Their work has, however focused on Earcons. The research in this thesis uses methods based upon this work but it is focused on Auditory Icons rather than Earcons.

The previous work on Earcons by Papp (1997), by Brewster (1994), and McGookin and Brewster (2004) found that the limit of identification for Earcons was most positively influenced by using staggered onsets between the Earcons and designing multi-timbre Earcons to enhance their identification or ability to 'stand out'. McGookin and Brewster (2004) explored up to four concurrent Earcons and found that in many conditions only two of the four could be easily identified, however for certain mappings, three of the four Earcons were identifiable. These results suggest that at most three concurrent Earcons can be used and opens up the question of how many can be used for concurrent presentation if instead of Earcons, Auditory Icons are used. Papp (1997) proposed a prediction model for user identification error rates based upon a number of factors including average overlap. Average overlap referred to the sum of the lengths of all the Earcons divided by the time that they had to be played within. This approach was based on a solid mathematical foundation, but it has found little acceptance in the wider field of Auditory Display. A reason behind the possible low acceptance of this approach is that the source code of the prototype and the thesis dissertation document are not electronically available. Papp (1997) studied one, two, and three concurrent Earcons. The relatively low number of concurrent Earcons that can be uniquely identified and the studies in this chapter focus on exploring if everyday sounds could have higher numbers of concurrent sounds correctly identified. Brewster (1994) explored parallel Earcons but found that subjects required more cognitive resources to recognise the meanings of the Earcons and this might potentially lower the performance rates on other tasks.

This chapter focuses on two questions - the identification and the confusion of the simultaneous everyday sounds. The investigation focused on how the identification of simultaneously presented Auditory Icons are affected when sounds have a prior classification based on the objects and actions occurring in the sound. In each of the explorations in this chapter, two distinct sets of sounds are presented. The first set uses a prior classification of the sounds, where the sounds are selected to ensure that no sound has the same objects or actions occurring within the sound. This ensures that no two sounds in a listening condition will have the same action or object, e.g. only one hitting sound or only one glass sound. The second set of sounds has no constraints and were randomly selected from the particular stimuli pool. This meant that in a listening condition of random sounds, a number of glass or hitting sounds could be present. The conditions presented a range of near simultaneous everyday sounds with a 300 ms onset-to-onset delay between the start of each sound. The range was from three to ten near simultaneous everyday sounds, within each study there was no reuse of sounds between the conditions or between the sets. The studies explored if prior classification based on object and action properties improved the listener identification of Auditory Icons with a specific focus on concurrent or simultaneous presentation of Auditory Icons. The studies in this chapter also investigated the related descriptions and issues of confusion where a sound may be well identified by a listener, i.e. the individual heard a sound but confused it with some other object or action. A causal uncertainty metric was introduced and measured how the sounds within the conditions may be produced by different causes.

The theories of Bertin (1983) and Bregman (1990), in vision and audition respectively, suggest a two level hierarchy of information with a global level for overall analysis and a local level which focuses attention on particular details. These levels are shown in Table 5.2. In the case of audio and of Bregman's (1990) theory, it suggest that sounds are segregated into different auditory streams. This segregation theory of audition suggests streams are categorical and exclusive. It further suggests that judgements within a stream of elements are easy but similar judgements between different streams are much more difficult. This is important for Auditory Display designers, as undetectable elements add no information or value so it is important that when sounds occupy the same space and time that they be heard as separate entities.

Levels of Information	Visual	Auditory
local	single item or element	single element within stream
intermediate	subsets and groups of elements	a single stream
global	all you see (entire set of elements)	all you hear (auditory scene)

Table 5.2*:* Levels of information, derived from Bertin (1983) and Bregman (1990) as applied to vision and audition.

Another view of audition comes from Gaver (1993a, 1993c) and the previously discussed idea of everyday listening with its related taxonomy and hierarchical sonic events. This type of listening would suggest that perceived similar sounds would be heard as the same element. This suggests that identification could be improved in Auditory Displays by using sounds from different categories. The methods in this chapter help in the identification of which of the sounds 'stood out' and which of the sounds were not heard or misidentified. While not explored further in this thesis, such insights offer a greater insight into whether or not particular sound designs will be accepted and how people react to the particular sounds. This relates to Schafer's (1977) concepts of 'sound romances' and 'sound phobias' which are sounds that people have positive or negative reactions to, respectively. Additionally, from a methodological viewpoint, the studies investigated the extension of existing methods to see if they can be used with concurrent everyday sounds as previously the methodological focus was on sequential rather than on simultaneous audio. The sounds explored are hybrid events as defined by Gaver (1988) or as simulation examples within the Sounding Object's taxonomy (Houix et al., 2007b). Concurrent sounds consist of many interacting sounds and the study of how these sounds interrelate and how to combined them successfully is of great importance to designers wishing to use them in their Auditory Displays. In this work, it is hypothesised that the best combinations will be sounds, which do not contain any similar object or action components. This will ensure that the sounds are more distinct and identifiable to listeners. This work can be seen as the one step higher in this taxonomy when compared to the previous chapter and its study of dropping sounds.

Designers of Auditory Displays, especially novice designers make naïve psychological assumptions about what sounds were being listened to in their Auditory Display. Experienced designers often make educated guesses based upon their skill but even they may be wrong with the result that the Auditory Display will communicate information in a manner that is demanding or confusing for listeners. The approach presented in this chapter shows how designers can obtain accurate data about what sounds a user was actually hearing and identifying. Further work is needed to find out more about presenting sound in Auditory Display and this chapter concentrates on the aspect of presenting concurrent everyday sounds.

5.1 Exploring the identification of concurrent Auditory Icons

The study in the previous chapter examined if real and synthetic Auditory Icons could provide information on a listener's perceptual scaling and realism judgements for the stimuli presented. This work raised questions about how the participants classified the sound and whether knowledge of such categories could be used to improve an Auditory Display. In this chapter, source identification and sound identification, are explored to provide an approach for selecting appropriate sounds for use where many sounds can occur simultaneously. The motivation for exploring simultaneous sounds came both from the experimental interface used in Chapter 4 and was inspired by the cocktail party effect (Cherry, 1953, Cherry and Taylor, 1954). This effect highlights a listener's ability to selectively attend to multiple different simultaneously occurring streams or sounds. There is a lack of research and method in Auditory Display for exploring this question, particular on the topic of Auditory Icons where no previous work exists. This work uses an approach, that elicits a set of free text descriptors provided by the participants, which can offer potential mappings and metaphors. This information can be used to assist designers and complement the identification and confusion information that is generated about the sounds.

Concurrent presentation of Auditory Icons can cause them to interfere with one another and may result in a cacophony. This can make it difficult to distinguish any sound individually but to-date there have been no studies that have investigated the extent of how the identification of Auditory Icons can be impaired by their concurrent presentation. McGookin (2004) conducted such studies for Earcons while Brungart et al. (2002) and Brungart and Simpson (2002) investigated concurrent speech presentation. These studies have shown that where the number of concurrent audio items is increased, the total proportion of identified audio is, not surprisingly reduced. The work by McGookin and Brewster (2004) on Earcons suggested using at least a 300 ms onset-to-onset gap. They found using an onset gap between the starts of concurrently presented Earcons improved identifiability. The work in this chapter explores the impact on Auditory Icon identification by increasing the number of concurrently presented Auditory Icons while maintaining a 300 ms onset-to-onset gap between sound onsets.

Monophonic sounds were used in this chapter's studies to reduce the number of variables in the problem domain. Unlike visual objects, when two sounds are placed near or overlapping each other in a physical space, it will be much more difficult for a person to switch between them. The sounds may even partially or completely mask one another. This makes it difficult to determine issues such as how far apart the sounds need to be, how many sounds can be presented, or even how similar the sounds can be. The area of concurrent spatialised sound presentation as it is applied to the design of Auditory Displays is outside the scope of this thesis. Spatialisation and a listener's ability to localise or identify the location in space of a sound source or sources is a complex research area that has many complex variables that need to be modelled (Best, 2004, Shinn-Cunningham and Ihlefeld, 2004, Afonso et al., 2005, Simpson et al., 2007). Extending the work to explore how spatialisation affects the research questions in this thesis is an area for future work. This work provides a baseline, which can then be used to compare with comparable spatialised studies.

5.1.1 Prior methods and approaches for the identification of concurrent Auditory Icons

Causal Uncertainty

Entropy is a measure determined by events and the probability of the occurrence of those events. Ludwig Boltzmann (1872) first used it in statistical mechanics as a mechanism for measuring the number of microscopic ways a defined macroscopic state can be reached. This concept was later incorporated into information theory as developed by Shannon (1948) as a mechanism defining the relationship between the received information and the probability of the event being observed. As well as developing an information measure I, Shannon also developed a measure H for the mean information called the entropy of the system. This entropy measure has three fundamental properties:

1. It is at its maximum where all events have the same probability of occurring. This is the most undefined state of the system.
2. It is at its minimum when only one event has occurred. This is the perfect state of the system where no information can be added.
3. The entropy function is positive, continuous and symmetric.

Shannon's entropy measure H inspired Ballas et al. (1986) to develop the concept of *causal uncertainty*. Causal uncertainty is a measure of how a single sound may be produced by different causes. A H value for a sound is calculated based on the number and frequency of the different identifications for the given sound. Ballas and colleagues found these values were stable for different examples of a particular sound (Ballas et al., 1987), when correlated with identification time for a sound (Ballas et al., 1986), when compared with subjective ratings of uncertainty (Ballas and Howard, 1987), or when compared with the rating of the identifiability of a sound (Ballas, 1989).

These values were found to be consistent across many categories of participants who varied from secondary school students, university students to older listeners (Ballas et al., 1987, Ballas and Barnes, 1988). Ballas's research found that the H values had a greater correlation with identification time than with either percentage correct or with the number of alternative identifications of the sound. This suggests that for everyday sound identification there is some type of parallel processing of alternatives or other information processing occuring. In analysing the descriptions Ballas's method of causal uncertainty (1993) was used. Ballas et al. (1986) found that the identification time for everyday nonspeech sounds was a linear function of the logarithm of the number of alternative interpretations of a sound. H_{CU} is a measure of causal uncertainty for sound i, where p_{ij} is the proportion of all responses for sound i sorted into event categories j and n is the number of categories for responses to sound i as shown below in equation 5.1.1. Ballas' method gives a more informative measure of how easy it is for users to identify everyday sounds and provides stable results with two or more listeners.

Using the equation for H_{CU} results in a 0 when all participants agree on a particular response in a given listening test. Taking 11 participants as an example, where the responses are equally split between two alternatives, the causal uncertainty is 1.0. However, if the participants' responses are skewed where 10 of the 11 responses agree but a single response is different, then the causal uncertainty is 0.4395. In the case where all 11 responses are different, then the causal uncertainty is 3.4594. Using the causal uncertainty approach has the advantage of being able to illustrate unity, degree of split, or skewed responses easily. These types of responses would normally require several figures to illustrate the same trend in participant responses as opposed to using a single figure for causal uncertainty.

$$H_{CU} = \sum_{i}^{n} p_{ij} \log_2 p_{ij} \tag{5.1.1}$$

Equation 1: Causal Uncertainty (Ballas et al., 1986)

Understanding Auditory Icons using Causal Uncertainty

In order to further the understanding of people's perception of auditory events, a listening test approach was adopted as used by other researchers (Ballas, 1993, Gaver, 1988, Vanderveer, 1979) for examining the same topic. The recorded sounds used in this thesis and its studies were presented to participants using headphones, who responded in free-text format to what they thought each sound was. These text descriptions were often highly descriptive and accurate. Examples include "*marble ball dropped onto table*", "*bell ringing from medium sized bell*" and "*horse walking, the horse noise sounds dry, dark, even brown*". These studies used an approach based on Ballas' method of causal uncertainty as discussed in Section 5.1.1. This method has one important caveat when combined with the measure for the average proportion of correctly identified sounds. The method of causal uncertainty indicates confusion with *heard sounds* as described by participants in their free text responses. The distinction between *heard sounds* and *unheard sounds* used in this chapter is related to those sounds heard, recognised and identified by listeners versus those sound unheard and unrecognised by listeners. This is an important issue for causal uncertainty as sounds that are not heard by listeners can negatively effect the calculation of this measure.

The interpretation of participants free-text responses generates the list of which sounds a particular heard and which sounds they did not hear or identify. These unidentified or *unheard sounds* will not have a textual response from the participant. If all participants do not hear the same sound, the result of using the causal uncertainty measure will return zero. This zero must be interpreted in light of the low number of responses from the participants, otherwise the causal uncertainty measure will show a high degree of unity for participant's identification of the sound, where in fact the listeners did not hear the sound. This caveat is important, and means that both the causal uncertainty results and the results of the correctly identified Auditory Icons must be taken into account when determining which sounds are best suited for use. A number of heuristics and methods were combined in the categorisation process of the free-text responses.

Marcell's (2007) equivalence judgement method and Vanderveer's (1979, p. 88) heuristic for correct response method were used in the classification of free text descriptors from listeners. The heuristic used by Vanderveer (1979, p. 88) rates the response by asking "*does the expression (the written response) refer to the correct (target) event ?*" or "*What is the class of events that would be correctly described by this expression .. does that class include, and is it limited to, the kind of event that actually took place ?*". In a handful of cases, stream merging occurred and two descriptors were created from the original descriptor when appropriate. An example of the mapping process is shown in Figure 5.3. The six criteria suggested by Ballas (1986) for categorisation of free-text responses of everyday sounds were also used as part of the categorisation process.

1. Phrases using exactly the same noun and verb should be placed in the same category.
2. Phrases using nouns and verbs that are synonyms should be placed in the same category.
3. Phrases describing the same physical scene, as would be used to describe a scene in a movie script, should be placed in the same category.
4. A phrase missing a verb such as in the response "door", should be set aside until the first pass was completed. These phrases should be placed into the most frequent category which uses the noun contained in the phrase.
5. Responses which are not specific enough to be categorised,(e.g., "object hitting another object" or "item falling"), should be excluded from the sorting.

5.1.2 Study 1 - Pilot Study examining three, four, five, and six concurrent Auditory Icons

This pilot study investigated concurrent presentation of Auditory Icons in two conditions, the first used the random selection of Auditory Icons and the second used prior classification to select sounds where no two sounds in a listening condition had the same action or object properties. Example sets of Auditory Icons could not for example, contain several banging sounds or several sounds from the same type of object such as glass objects or motor vehicles. In this particular pilot study, the number of concurrent Auditory Icons being presented varied from three to six. The previous studies in Earcons (McGookin, 2004) suggested that four concurrent Earcons was the optimal number for concurrent presentation. This pilot study explores a similar range to determine the optimal number of everyday sounds for concurrent presentation. The following hypotheses were made with regard to the expected results of the first study that examined three to six concurrent Auditory Icons.

Hypothesis 1. *The participants perform better in identifying the sounds in conditions which had been selected to ensure that no two sounds in a condition had the same object or action properties based on a classification of the sound's descriptors.*

Hypothesis 2. *The performance of participants with regard to identification would degrade as the number of Auditory Icons presented increased.*

Hypothesis 3. *The identification performance of participants would degrade more in conditions which permitted sounds to have similar object or object classifications when compared to the other conditions that prevented this.*

Hypothesis 4. *The object and actionhood of sounds are salient criteria used by participants for the identification of sounds.*

The main aim was to see if simultaneous complex sounds could be identified by listeners, as there would be no point in using them if not. The hypothesis was that Auditory Icons would be effective at communicating information simultaneously. The overall recognition rates for Auditory Icons would be high and higher again when the sounds used had been selected to ensure that no object or action descriptor overlapped. The prior classification of Auditory Icons would make them easier to remember and discriminate between them.

Participants

11 participants were recruited from the postgraduates at the University of Limerick. All participants reported normal hearing and had normal or corrected to normal vision. Written consent was obtained prior to the pilot study from all participants. These participants had not taken part in the previous experiments.

Stimuli

72 high-quality monophonic sounds (44.1 Kilohertz 16-bit) everyday sounds (durations between 0.4 and 28.5) were used in this investigation. 40 sounds were taken from a local sound

collection in the University of Limerick[1] and 32 of the sounds were selected from a commercial sound effects CD collection (Hollywood-Edge, 1990). The sounds used were complex, dynamic and informational events with different temporal patterns (Jenkins, 1985, McAdams, 1993). These sounds were edited to a fixed duration allowing for the "*sound event*" or "*sound object*" (Port et al., 1995) to appear to occur naturally (Wightman and Jenison, 1995). This editing ensured that things such as obvious fades or linear volume ramping at the start of a sound were avoided and only the essence of the event was kept. The sounds used in this pilot study are described in Table 5.3 and represent a wide range of environmental sounds, to pose both easy and difficult identification problems.

The sounds had been classified with particular focus on two categories, the object category of the sound and the action category of the sound. An interesting result from related research Fernström et al. (2005) was that actions of sounds are better identified than the objects involved in a sound. Research from Ballas and Howard (1987) found that semantic context in sound interpretation is an important factor and that auditory perception is directed towards awareness of the sources of sounds i.e. the events producing sounds. They also stated that the function of auditory perception is to recognise events rather than processing acoustic patterns. Events are defined for this pilot study to consist of actions, objects and context. The pilot study explored the action and object properties in the context of the identification of concurrently presented Auditory Icons. The work by McGookin and Brewster (2004) on Earcons suggested using at least a 300 ms onset-to-onset gap to improve identifiability. This start-to-start gap between two concurrent or almost concurrent sounds can prevent merging of the two sounds into a single stream. This suggestion and the idea that better identification can occur in simultaneous conditions where no pair of sounds within a condition have the same action or object descriptors were key aspects of the studies in this chapter.

[1] Interaction Design Centre, University of Limerick, Ecological Sounds Collection - http://www.idc.ul.ie/mikael/sounds/ecosound.zip

	Prior classification sounds and their descriptions (no sounds in condition with same actions or objects)		***No prior classification (randomly selected) sounds and their descriptions***
ID	*Description*	*ID*	*Description*
1	Water filling a glass bottle from a tap in a kitchen	1	Dishwasher in operation
2	Ball bouncing three times	2	Pouring water out of glass bottle into a sink in a kitchen
3	A person running up carpeted stairs in a wooden hallway.	3	Water splashing slowly onto tiles in a kitchen
Degree of Overlap	Objects 0% - Actions 0%		Objects 100% - Actions 100%
4	A motorbike passing the house and driving up through an estate.	4	Washing hands in sink
5	A person brushing their teeth	5	Rapid Door Knocks
6	Knocking on a door	6	Stirring water with a metal spoon in a cup in a kitchen
7	An electic alarm block buzzing	7	Bare fingers drumming on a wooden table in a kitchen.
Degree of Overlap	Objects 0% - Actions 0%		Objects 50%-50% - Actions 75%
8	Several birds singing in a rural setting	8	Several birds singing in a rural setting
9	A glass window breaking	9	Rain falling through the trees
10	Water flowing quickly from a bath tap in a tiled bathroom.	10	Walking on sticks outdoors under trees and breaking them.
11	A person knocking on a wooden kitchen door.	11	Tractor returning from the fields
12	Rattling a metal chain several times times in bare hands.	12	Wind rustling through the leaves on a summers day.
Degree of Overlap	Objects 0% - Actions 0%		Objects 80% - Actions 40%
13	Running on a concrete surface	13	Many glasses clinking in toast
14	Sawing a piece of wood in a kitchen on a wooden table.	14	Putting plate on top of another in a press and closing the press after.
15	Ball bouncing three times	15	Two glasses clashing together 5 times.
16	Using a vending machine and getting an item.	16	Breaking 3 glasses in succession against cement brick.
17	A motorbike starting, revved and driven off out of the estate.	17	Sliding cups into a cupboard after drying them in a kitchen.
18	Several glasses clinking off each other	18	Breaking 2 ceramic cups in succession of a cement block.
Degree of Overlap	Objects 0% - Actions 0%		Objects 100% - Actions 33%-66%

Table 5.3*:* The sounds used in the first pilot study with descriptions - 3, 4, 5, 6 Auditory Icons. The sounds on the left are those, which had been classified to ensure that no object or action properties overlapped in any particular condition. The sounds on the right where randomly selected.

Experimental Platform - Technical Details The system was designed for data and application logging, the components of the system are programmed in C# and run on the Windows NT and XP platforms. This system and its source code, as well as the configuration files used are all available as free software under the GNU General Public License version 3 or greater.

The application provided for sound and condition randomisation, the recording of the participants' responses, and application logging. A copy of the program and the results from the participants can be found online or in accompanying DVD for this thesis in Appendix L.1.

Training

The focus of this training phase was to familiarise the participants with concurrent presentation of Auditory Icons. A training interface allowed participants to stop, start and loop up to seven Auditory Icons simultaneously whose descriptions were provided on-screen. The training interface is shown in Figure 5.1. The stimuli used in the training phase were not used in the later tests. The participants spent ten minutes listening to the training stimuli using the interface after a short introduction on its operations. They were provided with the task sheet describing the experiment and a verbal explanation was given explaining how the tasks were to conducted using the experimental platform. Participants listened to the sounds (in mono) using headphones while interacting with the system.

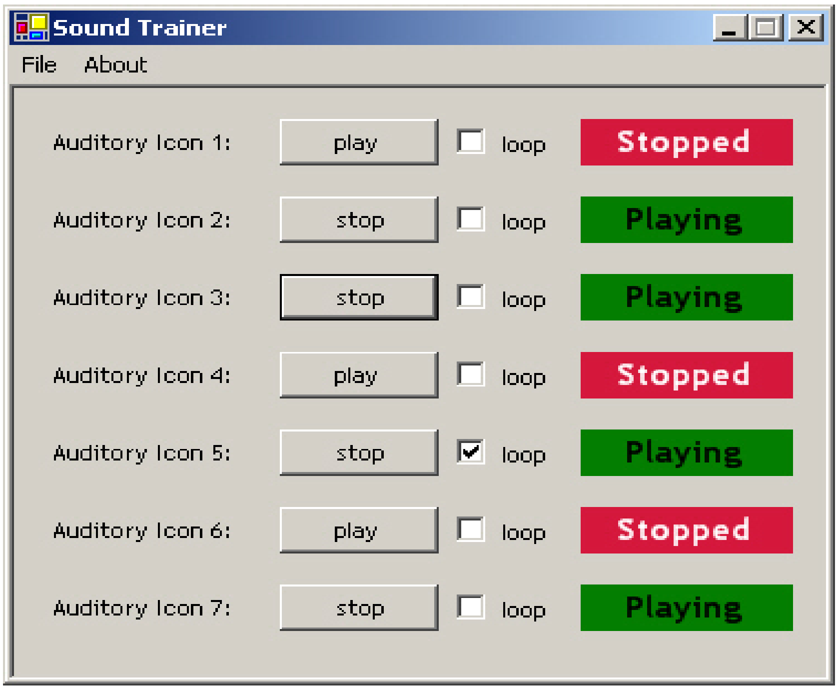


Fig. 5.1: *A screen shot of the training interface used by participants to familiarise the participants with concurrent presentation of Auditory Icons.*

Study Design

Using a within-subjects design, the stimuli were presented in random order within each condition and the task order was counter-balanced for the conditions (three, four, five or six concurrent Auditory Icons). Each set of stimuli was presented randomly and as a single block for the particular condition.

Procedure

The participants listened to the recorded sounds (mono) in random order using headphones, responding in free-text format to what each sound was, using the interface shown in Figure 5.2. The additional text boxes in Figure 5.2 were provided to ensure that participants had adequate space to enter all the responses they felt they heard in a condition. This was provided as it was envisaged that participants may enter a number of descriptions that corresponded to a single event.

The conditions varied from three to six sounds being concurrently presented. Each condition was played four times in succession. Participants had no control over the playback of sounds and could only use the interface to input their descriptions and move to the next step. The interface below used ten description boxes, as the complex scenes with multiple sounds may lead to richer descriptions than just one descriptor per recorded sound.

In a similar fashion to Vanderveer (1979, p. 83) non-directive instructional guidelines were given to the participants. The instructions to participants can be seen in Appendix E.1. They were asked to describe a sound but they were not told to describe its agent or source, the action or event, or its perceptual qualities. A single rater codified the action and the object descriptors elicited from participants using a two stage process.

The coding of the object and action descriptors was performed by a single rater following the approached used by Fernström et al. (2005), by Houix et al. (2007a), and by Tardieu et al. (2008). It used a two stage classification, which took approximately four days. In the first stage, the categories created by the rater did not have to be kept consistent. The second stage, grouped these existing classifiers to reduce the number of criteria while keeping a consistency to the internal measures the categories had used. Houix et al. (2007a) analysed the correlations between classifiers and found a degree of dissimilarity in rater classification for only a small number of sounds. The effect of prior experience with training and the correlation results from earlier studies would suggest that using a single codifier or sorter should not be a major issue but further studies are required to clarify this issue.

The methods used were application data logging and post study questionnaires. The application data logging captured the sound description entered by the participant per condition. Participants used the questionnaires to rate the ease of use and perceived difficulty of the task. To determine the number of Auditory Icons correctly identified by participants, the following method was used. The sets of descriptors from participants were collected and analysed. This analysis compared the descriptors for each condition to the sounds presented in the particular condition as shown in Figure 5.3. In cases where a participant's descriptor matched

Fig. 5.2: *A screen shot of the dialogue used by participants to fill in the descriptions of parallel sounds presented in the three, four, five, and six concurrent Auditory Icon conditions.*

approximately one of the existing sounds in the particular condition and where that particular sound had not already been matched with a previous description, the number of correctly identified Auditory Icon was increased by one, and the Auditory Icon description was marked as allocated. This is shown in image 'A' of Figure 5.3, this process was continued for each description as shown in images 'B' and 'C' until all the descriptions were linked to one of the existing sounds. In the case where no matching descriptor could be found, the descriptor was not counted and discarded. In a couple of rare instances, a description was merged with another descriptor where stream segregation of the particular sound had occurred in the mind of the particular listener. Stream merging also occurred in a handful of cases and two descriptors were created from the original descriptor when appropriate. As discussed in Section 5.1.1, Marcell's (2007) equivalence judgement method and Vanderveer's (1979) heuristic for correct response method were used in classification of free text descriptors from listeners in a similar type of a task. The process of how the mapping between the set of Auditory Icons presented and the set of participant free text responses was handled is shown in Figure 5.3 for four Auditory Icons.

Due to the differing numbers of Auditory Icons to be identified in each condition, a direct numerical comparison between the Auditory Icons and the correctly identified Auditory Icons would not be possible. An alternative comparison can be made using the average number of Auditory Icons identified per participant and converting this average into a percentage of the number of Auditory Icons that were concurrently presented. For example, in the four concurrent condition, if on average three of the sounds were correctly identified, the average percentage identified was calculated as (3/4)*100 = 75%.

Results

The results show that the high performance of participants in both of types of conditions presented. There is an indication that in the six concurrent Auditory Icon condition with no classification, the identification performance of listeners drops (50%) versus the same condition with prior classification (89.4%). The percentages of overlap for the categories with no prior classification are shown in Table 5.4. The overlap in these conditions should kept in mind when reviewing the average proportion of correctly identified Auditory Icons.

Average proportion of correctly identified Auditory Icons For each set of (three, four, five or six) concurrently presented Auditory Icons, the set of Auditory Icons presented and the set of participant responses to those Auditory Icons were compared. The categories derived from the individual participant results are shown in Table E-3 and in Table E-4 in Appendix E.1. The responses from participants for the action and object conditions with prior classification are shown in Tables E-9, E-10, E-17, and E-18. The responses from participants for the action and object conditions with no prior (random) classification are shown in Tables E-5, E-6, E-13, and E-14. The average proportion of correctly identified Auditory Icons across all participants is presented graphically in Figure 5.4.

The individual results of each participant was plotted and is shown in Figure 5.5 is shown highlighted as 'A' while the indicator of a drop off in performance is highlighted as 'B'. The individual results give a better indication of the spread within conditions than an averaging of all the participants. The highlighting in Figure 5.5 of 'A' does show that the majority of participants are clustered in the upper ranges of scale from 50% to 100% correct identification. This confirms the task's relatively good identification results were consistent excluding a small number of outliers. These outliers could be due to the particular sound not being heard, the sound being masked, or the sound may have been confused. The overall trends in the results have shown the beginnings of a performance decline in the six concurrent Auditory Icons with no classification. The hypotheses made prior to this pilot study were supported for low numbers (6 or less) of concurrent Auditory Icons. This small number of Auditory Icons points out the need for further exploration with higher numbers of concurrent Auditory Icons, which the next pilot study deals with. The results of this pilot study also verify hypothesis 4, that object and actionhood were used as salient criteria for identification by listeners.

The distribution was not normally distributed which required the use of Kruskal Wallis and Dunn multiple comparison tests (Crawley, 2005) to determine if any of the difference

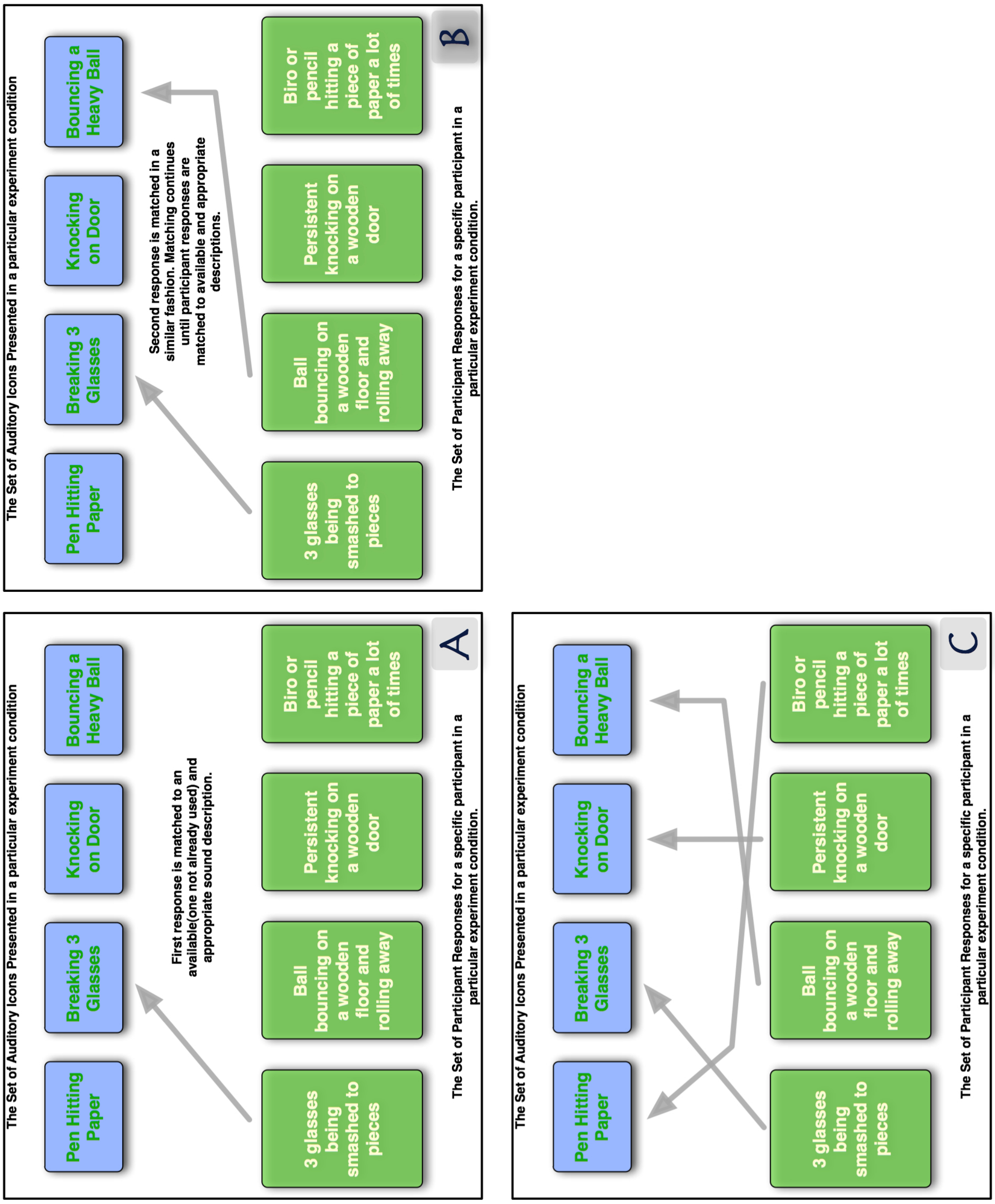


Fig. 5.3: *An example from the four Auditory Icon condition of how the set of participant responses was mapped to the set of Auditory Icons presented in order to determine the number of correctly identified Auditory Icons. This mapping used the heuristics and methods discussed in Section 5.1.1.*

Number of concurrent Auditory Icons (AIs) presented	Overlap of Actions	Overlap of Objects	Average overlap for number of presented AIs
3	100%	100%	100%
4	75%	100%	88%
5	80%	100%	90%
6	100%	100%	100%
Average overlap for condition	**88.75%**	**100%**	

Table 5.4*:* Percent of overlapping sound categories within the concurrent auditory presentation conditions in the first pilot study.

shown in Figure 5.4 were statistically significant. The findings from this analysis showed that conditions with more sounds are more difficult to identify. The exception to this finding is the three Auditory Icon condition with prior classification, which was more difficult to identify either of the four concurrent Auditory Icon conditions. This was due to potential masking and a more detailed explanation of this outlier is given in Section 5.1.2 with the possible masking shown in Figure 5.9. The results of the Kruskal Wallis analysis are shown in Table 5.5. Six differences were found, as expect both of the three simultaneous Auditory Icon conditions are found to be easier than the six Auditory Icon condition with constraints (No Prior Classification 6). Similarly, it was found that the six simultaneous Auditory Icons with prior classification (Prior Classification 6) was significantly easier to identify than the six simultaneous Auditory Icons with no classification or random selection (No Prior Classification 6). These results suggest that Auditory Icons identification grows worse as the number presented is increased. These results motivated the second pilot study, which looked at seven, eight, nine, and ten concurrently presented Auditory Icons to determine if performance would continue to drop in conditions without classification. The high performance rates (of 85%+) of the conditions with prior classification support the hypothesis that prior classification does improve identifiability of sounds.

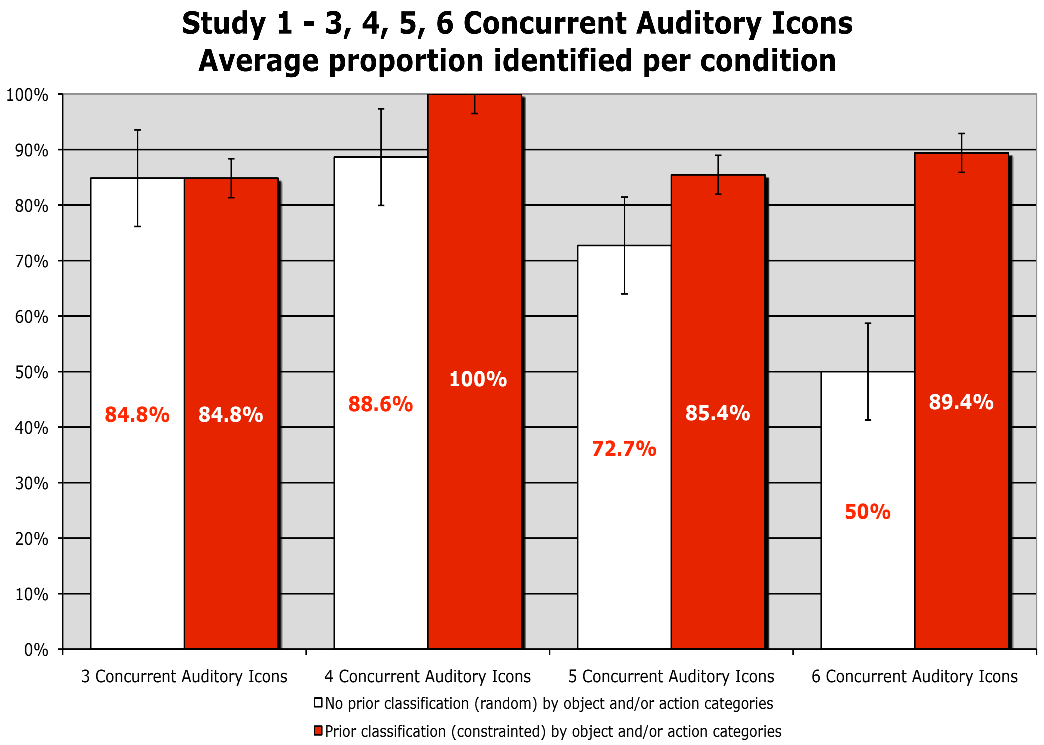


Fig. 5.4: *Graph showing the average proportion of Auditory Icons identified for the three, four, five and six Auditory Icon conditions.*

Conditions	*Mean Rank Difference*	*P-value*	*Conditions*	*Mean Rank Difference*	*P-value*
Prior Classification 3 vs. No Prior Classification 3	−1.273	ns P>0.05	Prior Classification 4 vs. No Prior Classification 4	18.636	ns P>0.05
Prior Classification 3 vs. Prior Classification 4	−17.455	ns P>0.05	Prior Classification 4 vs. Prior Classification 5	23.273	ns P>0.05
Prior Classification 3 vs. No Prior Classification 4	1.182	ns P>0.05	**Prior Classification 4 vs. No Prior Classification 5**	39.545	** P<0.01
Prior Classification 3 vs. Prior Classification 5	5.818	ns P>0.05	Prior Classification 4 vs. Prior Classification 6	14.818	ns P>0.05
Prior Classification 3 vs. No Prior Classification 5	22.091	ns P>0.05	**Prior Classification 4 vs. No Prior Classification 6**	58.091	*** P<0.001
Prior Classification 3 vs. Prior Classification 6	−2.636	ns P>0.05	No Prior Classification 4 vs. Prior Classification 5	4.636	ns P>0.05
Prior Classification 3 vs. No Prior Classification 6	40.636	** P<0.01	No Prior Classification 4 vs. No Prior Classification 5	20.909	ns P>0.05
No Prior Classification 3 vs. Prior Classification 4	−16.182	ns P>0.05	No Prior Classification 4 vs. Prior Classification 6	−3.818	ns P>0.05
No Prior Classification 3 vs. No Prior Classification 4	2.455	ns P>0.05	**No Prior Classification 4 vs. No Prior Classification 6**	39.455	** P<0.01
No Prior Classification 3 vs. Prior Classification 5	7.091	ns P>0.05	Prior Classification 5 vs. No Prior Classification 5	16.273	ns P>0.05
No Prior Classification 3 vs. No Prior Classification 5	23.364	ns P>0.05	Prior Classification 5 vs. Prior Classification 6	−8.455	ns P>0.05
No Prior Classification 3 vs. Prior Classification 6	−1.364	ns P>0.05	No Prior Classification 5 vs. Prior Classification 6	−24.727	ns P>0.05
No Prior Classification 3 vs. No Prior Classification 6	41.909	** P<0.01	No Prior Classification 5 vs. No Prior Classification 6	18.545	ns P>0.05
			Prior Classification 6 vs. No Prior Classification 6	43.273	*** P<0.001

Table 5.5: Kruskal-Wallis and Dunn multiple comparison test results for the three, the four, the five, and the six concurrent Auditory Icons conditions.

Examining the results for correct identification with prior and with no prior classification shows that Auditory Icons identification grows more difficult as the number presented is increased beyond four Auditory Icons. The drop in correct identification is not as severe in conditions using prior classification versus those using random or no classification. The high

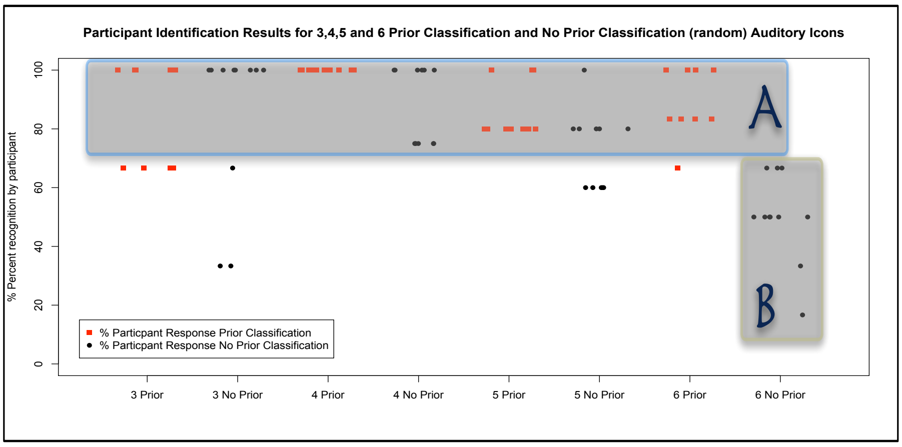


Fig. 5.5: *The identification of the Auditory Icons for the three, four, five and six Auditory Icon conditions by individual participants. Each dot represents a single participant's identification result for the particular condition. Trend 'A' shows that majority of participants have high individual identification rates with the exception of a few outliers. Trend 'B' shows a drop-off in performance in identification where six auditory icons are presented randomly or without any prior classification.*

(90%) correct identification in the six concurrent auditory icon condition using prior classification suggests that more situations with even more concurrent Auditory Icons should be studied. This is done in the next pilot. In order to provide greater detail about the difficulties in this pilot, the next sections look at an analysis of the participant's textual descriptors using the causal uncertainty method. This provides details on which of the sounds were confused.

Results and Observations for Participant Object Descriptions Of 3 To 6 Concurrent Auditory Icons The set of participant responses (SPR) to the presented Auditory Icons were sorted and categorised using the two stage classification process described in Section 6.2.1, as well as evaluated for correctness. From the responses, the object segments of the texts were extracted and categorised as shown in Table E-7 and in Table E-11 in Appendix E.1, such as what objects/materials were involved in the interaction. The responses from participants for the object conditions with prior classification are shown in Table E-9 and in Table E-10. The responses from participants for the object conditions with no prior (random) classification are shown in Table E-5 and in Table E-6.

The causal uncertainty results of the object descriptors for the 3 to 6 conditions are shown in Table E-8 and in Table E-12 in Appendix E.1. Interpreting Figure 5.6 shows each of the sounds has an associated bar in the chart whereas the less confused sounds or those with a zero causal uncertainty result did not have any associated bar in the chart. This allows for the tables to be easily interpreted as any bar graphs in results show the presence of confusion with regard to an individual Auditory Icon's interpretation by participants. The red bar charts represent Auditory Icons in conditions with prior classification, white bar charts represent Auditory Icons in conditions with no prior classification, the single divider lines represent the groups of the 3, 4, 5, and 6 Auditory Icons. The cumulative causal uncertainty displays the overall confusion for the particular sound scene. This scene consisted of the concurrently presented Auditory Icons for each condition. Examining Figure 5.6, the summary of results of the causal uncertainty measures, there are two sounds with high causal uncertainty.

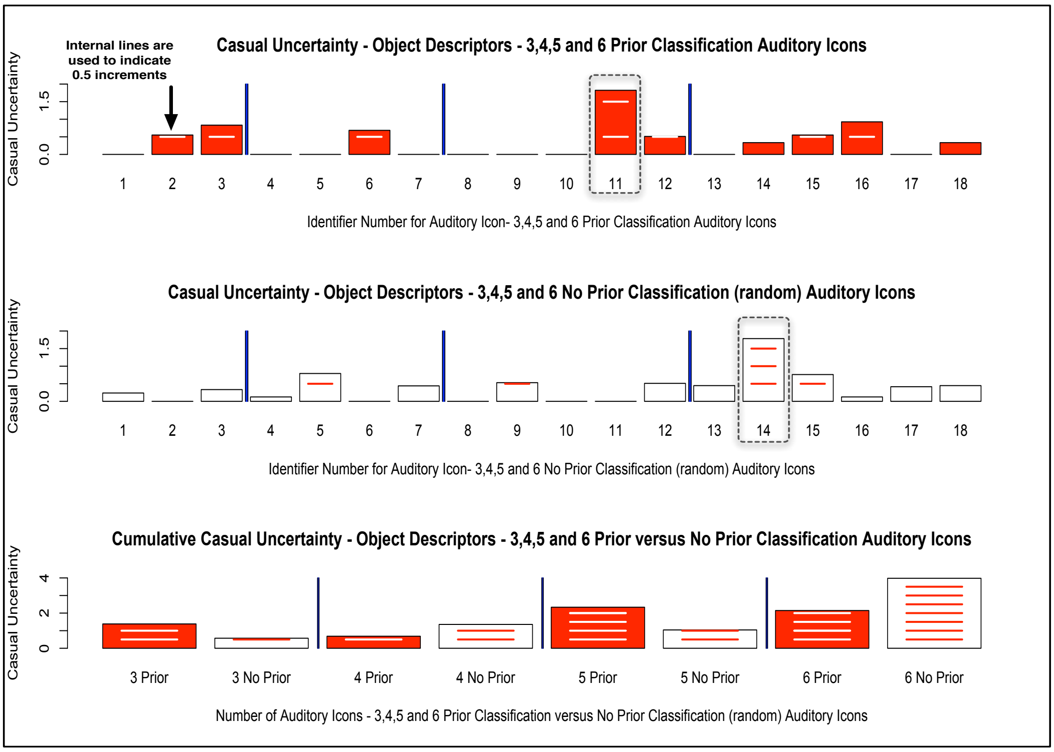


Fig. 5.6: *The analysed object descriptors results of participants, 3 to 6 Auditory Icons with prior and with no classification. The two sounds of interest, sound 11 and sound 14 are highlighted.*

Examining the first of the sounds, *sound 11* in the five concurrent Auditory Icons condition with prior classification. This was the sound of a person knocking on a wooden kitchen door. This sound had four distinct interpretations as shown in Table E-7 and a causal uncertainty of 1.7841 with all eleven participants having heard the sound. The participant's descriptors indicate that the sound was clearly heard but there was confusion as to what was being hit or knocked. This result would indicate that the sound was confused but did not contribute to the identification problems. Looking at Table E-7, it is most likely that *sound 12* which was a metal chain being rattled and which was only identified by three participants that caused the most identification problems.

The second sound with a high causal uncertainty was *sound 14* in the six concurrent Auditory Icons condition with no prior classification. This was a sound of putting a plate on top of another in a press and then closing the press after this. This sound had four distinct interpretations as shown in Table E-11 and a causal uncertainty of 1.8231 with only nine of the eleven participants having heard the sound. The participant's descriptors indicate that the sound was clearly heard but there was confusion as to what was moved or placed. This confusion and the lower rate of identification show that this sound was problematic. A further examination of Table E-11 shows that the sounds immediately before this and after it were both poorly identified (only 2 and 3 participants respectively heard them). Looking at Figure 5.7 which considers the sonograms and waveforms of the sounds, we can seen that there is a high possibility of merging due to close onsets and of material similarities between ceramic and glass. There is a strong indication from the results that *sound 17*, the sound of cups being put into a cupboard could have been the merged result as it had low confusion measure and a good identification as 7 of the 11 participants identified the sound.

The results of average identification of Auditory Icons identified per condition as shown in Figure 5.4 and the causal uncertainty results as shown in Figure 5.6 are helpful in providing a overview of how participants identify the sounds and which of the sounds are confused or misinterpreted. The phrasing and text provided in the participant's free text responses are also useful after analysis as the metaphors used in the descriptions can aid in designing mappings using the particular sound or combinations of sounds. These results would suggest that *sound 14* in the 6 concurrent Auditory Icons condition was affected by the lack of prior classification. The similarity between it and the other sounds in the condition should be further explored. This result supports the earlier hypothesis about ensuring the identifiability of sounds by ensuring only sounds with distinct object and actionhoods be used in concurrent presentation to maximise identifiability and to reduce confusion.

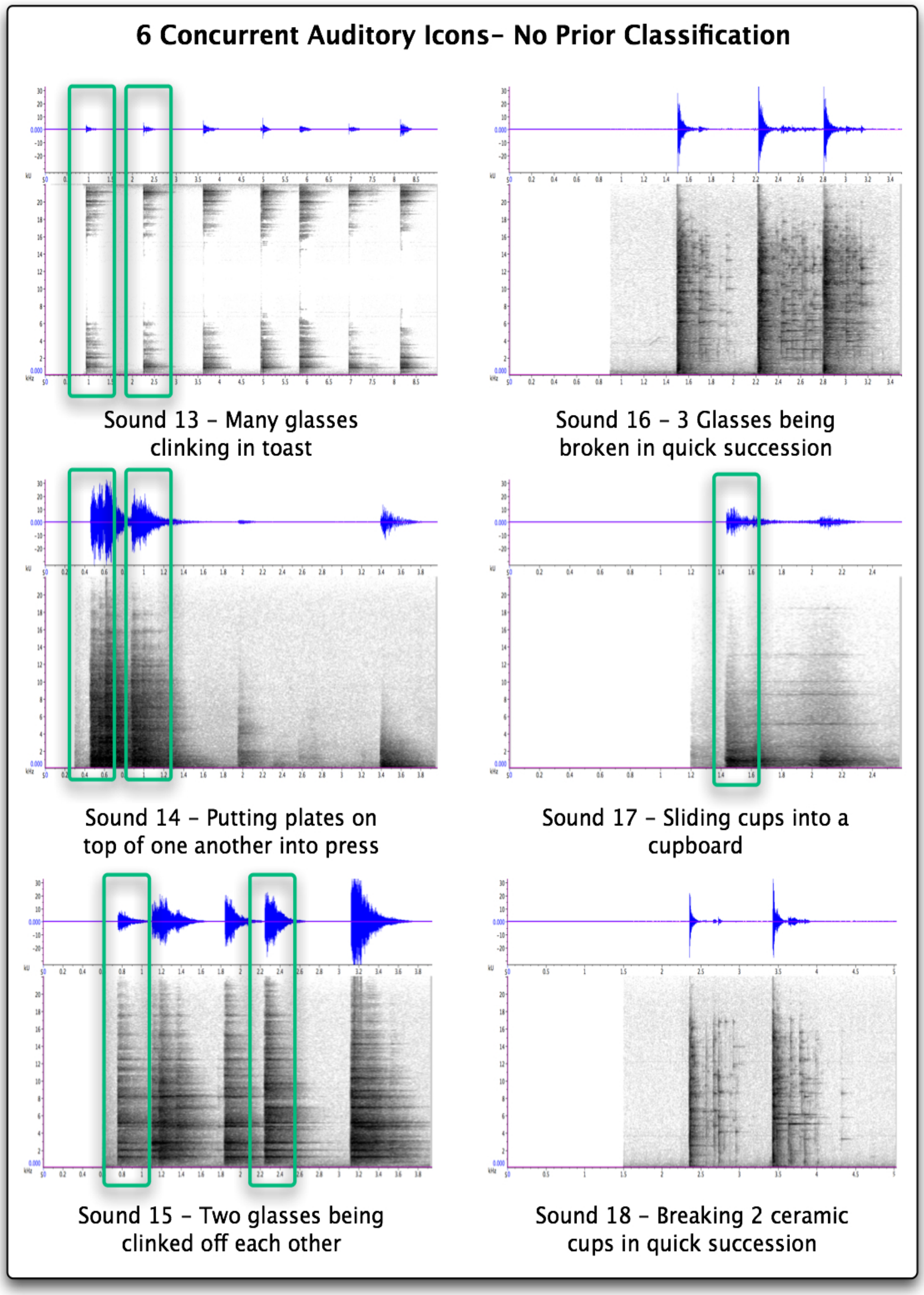


Fig. 5.7: *The six concurrently presented Auditory Icon condition with no prior classification. The highlighted boxes indicate close onsets of sounds with the same or similar material properties. This may have resulted in perceptual merging of the sounds by the listeners. The sonograms used had an FFT size of 512 with a 50% overlap using a Hanning window, from 20Hz to 4.7KHz.*

Results and Observations for Participant Action Descriptions Of 3 To 6 Concurrent Auditory Icons In a similar fashion to the previous object description analysis, the participants' responses were extracted and categorized by action segments of the texts as shown in Table E-15 and in Table E-19 in Appendix E.1. The responses from participants for the action conditions with prior classification are shown in Table E-17 and in Table E-18. The responses from participants for the action conditions with no prior (random) classification are shown in Table E-13 and in Table E-14.

Figure 5.8 presents the summary of results of the causal uncertainty results for 3 to 6 concurrent Auditory Icons with regard to the action descriptors. The causal uncertainty results of the action descriptors for the 3 to 6 conditions are shown in Table E-16 and in Table E-20 in Appendix E.1. Examining the 3, 4, 5, and 6 concurrent Auditory Icon conditions shows some interesting results especially noticeable is the 4 concurrent Auditory Icons with prior classification condition as all the sounds in this condition have a perfect (0) causal uncertainty result. This indicates the sounds and their actionhoods were clearly interpreted by the participants and examining Table E-16 confirms this as all eleven participants identified the sounds. There are two other interesting details in Figure 5.8, the first is the group of *sound 2* and of *sound 3* and the second is *sound 16*. These sounds were in the three and six conditions with prior classification respectively.

Examining the group of *sound 2* or "*Running Up Stairs*" and *sound 3* or "*"Bouncing Heavy Ball*" in Figure 5.8 suggests that there may have some impact type similarities which has added to the uncertainty of their interpretation. In Table E-16, we see that both sounds were confused with causal uncertainties of 0.8454 and of 1.0224 respectively. It also shows that only slightly more than half of the participants heard *sound 3*. The third sound in this condition, *sound 1* or "*Water Filling A Glass Bottle*" does not offer any such suggestion when considered with the other sounds in the condition. One participant commented for this condition saying that the "*sudden thudding sounds were interrupting the following sounds*". This indicates masking may be occurring for *sound 2* and *sound 3*. Examining the two sounds and their sonograms as shown in Figure 5.9 shows that it is possible that *sound 2* may have masked the quieter *sound 3*. In addition, both of the sound's respective waveforms show a number of very close onsets. This could have lead listeners to merge the two sounds together as one perceptual unit. A final consideration is that the similarity of the materials of the objects and of their actions may have increased the difficult for interpreting this set of sounds. This result highlights how this approach can provide useful details, which can facilitate an Auditory Display designer in the selection of identifiable and distinct sounds. In this case, it would have pointed out to the designer that at least *sound 2* should be replaced from the group

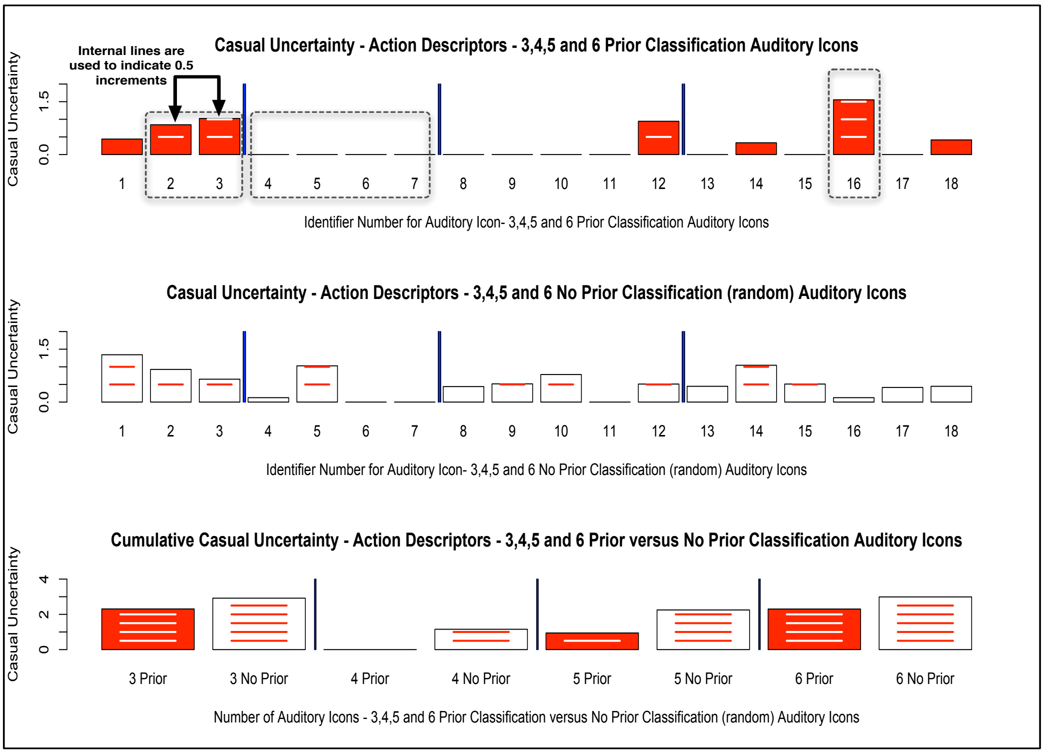


Fig. 5.8: *The analysed action descriptors results of participants, 3 to 6 Auditory Icons with prior and with no classification. The three groups of interest, sound 2 & 3, the four concurrent Auditory Icons with prior classification and sound 16 are highlighted.*

of sounds and possibly *sound 3* as well.

The second result of interest is focused on the 6 concurrent Auditory Icons with prior classification condition and on the *sound 16*. This sound had a causal uncertainty of 1.5531, the sound's description is that of using a vending machine and getting an item as shown previously in Table 5.3. Ten of the eleven participants heard this sound but confused its identification. The confusion may be due to this sound being heard in context with the other sounds in the condition. Perceptual merging could be a possibility when looking at *sound 15* and *sound 16* in Figure 5.10 except the results in Tables E-15, E-16 show that while the sound prior to it was a dropping sound, only two of the ten participants who heard *sound 16* classified it as a dropping sound. This result would suggest that *sound 16* may be masked to some degree or the actionhood of the sound was confused when heard in context with the other sounds within the condition or possibly a combination of both.

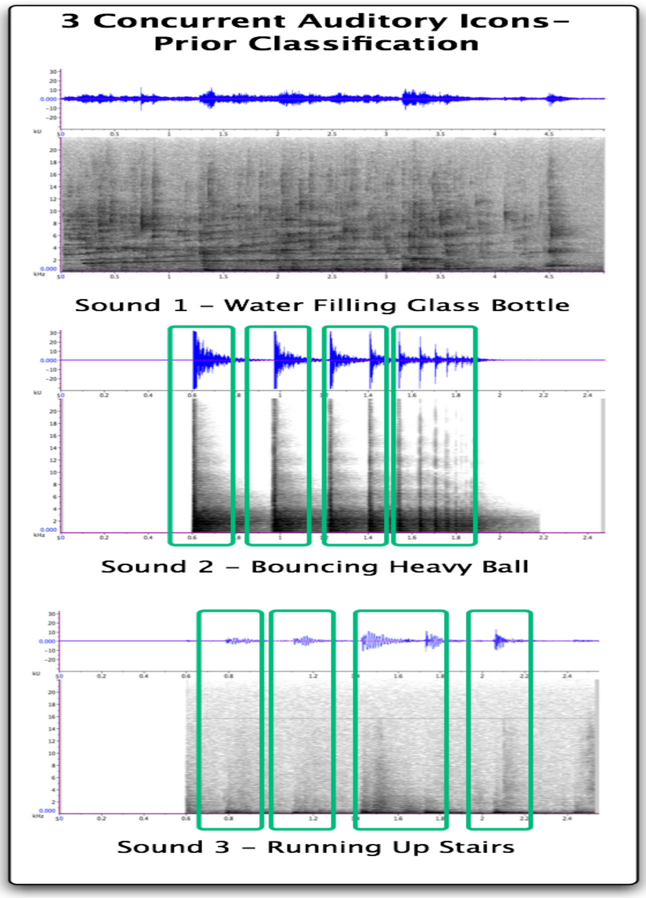


Fig. 5.9: *The three Auditory Icon condition with prior classification. The highlighted boxes indicate close onsets of sounds with the same or similar action and / or object properties. This may have resulted in perceptual merging or masking of the sounds by the listeners. The sonograms used had an FFT size of 512 with a 50% overlap using a Hanning window, from 20Hz to 4.7KHz.*

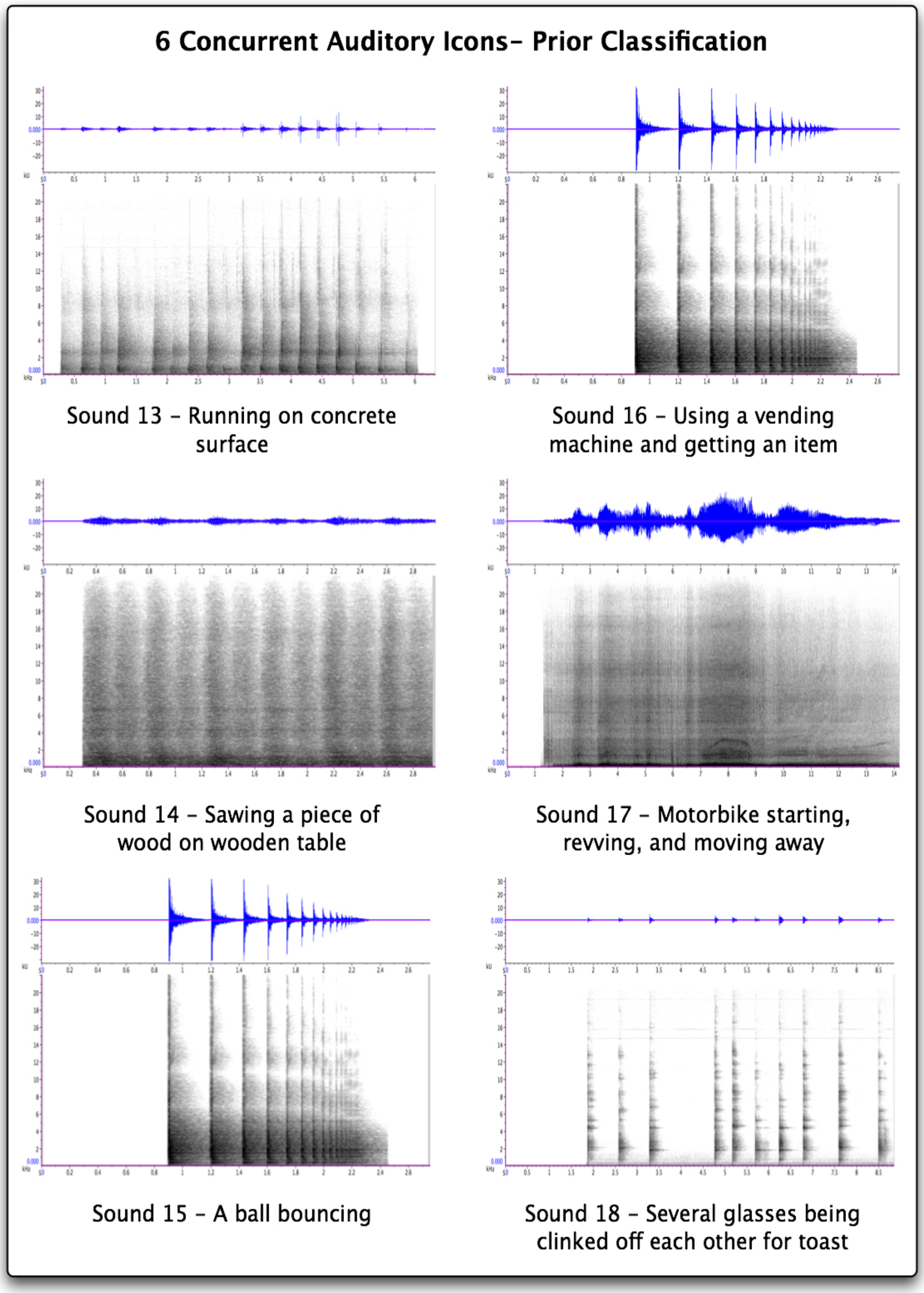


Fig. 5.10: *The six concurrently presented Auditory Icon condition with prior classification. The sonograms used had an FFT size of 512 with a 50% overlap using a Hanning window, from 20Hz to 4.7KHz.*

These results would suggest that *sound 16* is not suitable in the present selection of sounds for the 6 concurrent Auditory Icon condition and that the use of *sound 2* and *sound 3* in the 3 concurrent Auditory Icon condition should be reconsidered. The results for the action causal uncertainty for 3, 4, 5, and 6 concurrent Auditory Icons did support the hypothesis that it would be more difficult to interpret the actionhood of sounds in conditions when similar sounds overlap in an auditory scene.

Post study participant debriefing As part of this pilot study, a post study debriefing was used to collect data from the participants. The debriefing session raised several interesting points as indicated by comments from the participants like "*When you have the water going down a plug hole, so there was a sound like hitting a tap that started it and stopped it but I wasnt sure if it was part of the same action*" and "*Some sounds I heard once and even when concentrating on other sounds would drown out the other sounds I was trying to concentrate on, like the alarm clock or the stream bubbling sounds*". This type of comment did not directly affect the results for the first pilot study but did highlight how complex a person's interpretation of the auditory scene is. The point about the alarm clock and the stream bubbling sounds indicate that continuous or repeating sounds could irritate and distract listeners. There is a variety of research in auditory alarms (Patterson, 1989, Patterson and Datta, 1999, Bliss and Dunn, 2000) that deal with this and similar issues. In the context of this research, continuous or repeating sounds were not examined in detail and should be an area for future research. A more relevant type of comment can be seen where one participant stated that "*sudden thudding sounds were interrupting the following sounds*". This was related to the three concurrent Auditory Icons condition with prior classification and pointed out that masking may be occurring. Further analysis showed this was not the case but it was a useful comment that gave focus to the data exploration in that condition. Overall, the comments helped point out a number of salient issues with the first pilot study from the participants, such as masking and merging of sounds in particular conditions.

Results and implications from the first pilot study The results of the first pilot study show that even for the highest number of concurrently presented Auditory Icons it is relatively easy for participants to identify the sounds. The results are shown in Figure 5.4 and it can be seen that the 6 concurrent Auditory Icons with prior classification based on action and object categories still had an identification rate of approximately 89%. This result indicates that there is the potential for the presentation of many more concurrent Auditory Icons as the task was relatively easy for participants. It highlights how human listeners have a high proficiency in everyday listening skills. In order to explore the possibility for the presentation of many more concurrent Auditory Icons the next pilot study will concentrate on looking at situations of 7 to 10 concurrent Auditory Icons. The first pilot study has also shown how causal uncertainty can assist in determining confusion and how visualisation of the conditions using waveforms complemented with sonograms can highlight masking and merging issues. The results of this pilot study prompted the idea of a second small-scale pilot study to establish rough estimates of seven to ten concurrent Auditory Icons. These estimates would contribute to the selection of an appropriate range for a larger scale exploratory probe.

5.1.3 Study 2 - Pilot study examining seven, eight, nine, and ten concurrent Auditory Icons

This second pilot study investigated concurrent presentation of Auditory Icons with the aim of assessing the rough estimates of seven to ten Auditory Icons in a similar fashion as the first study. It investigated concurrent presentation of Auditory Icons in two conditions, the first used the random selection of Auditory Icons and the second used prior classification to select sounds where no two sounds in a listening condition had the same action or object properties. This pilot study explores the range of seven to ten Auditory Icons to determine the optimal number of everyday sounds for concurrent presentation. The hypotheses for this pilot were essentially the same as those in the first pilot study.

Hypothesis 1. *The participants perform better in identifying the sounds in conditions which had been selected to ensure that no two sounds in a condition had the same object or action properties based on a classification of the sound's descriptors.*

Hypothesis 2. *The performance of participants with regard to identification would degrade as the number of Auditory Icons presented increased.*

Hypothesis 3. *The identification performance of participants would degrade more in conditions which permitted sounds to have similar object or object classifications when compared to the other conditions that prevented this.*

Hypothesis 4. *The object and actionhood of sounds are salient criteria used by participants for the identification of sounds.*

The main aim was to see if simultaneous complex sounds could be identified by listeners, as there would be no point in using them if not. The hypothesis was that Auditory Icons would be effective at communicating information simultaneously. The particular reason for this pilot was to attempt to gauge the range or end point before identification performance dropped to a point at what the use of concurrent Auditory Icons would not be feasible. The hypothesis was that overall recognition rates for Auditory Icons would deteriorate with more concurrent sounds but that this degradation of identification performance would be lessened where prior classification had been used.

Participants

5 participants were recruited from the postgraduates at the University of Limerick. All participants reported normal hearing and had normal or corrected to normal vision. Written consent was obtained prior to the pilot study from all participants. All the participants in this study had not taken part in the earlier study.

Stimuli

A different set of 72 high-quality monophonic sounds (44.1 Kilohertz 16-bit) everyday sounds (durations between 0.4 and 28.5) to the first pilot study were used in the investigations. These sounds were selected from the BBC sound effects CD collection (BBC, 2006). These sounds used were selected in a similar manner to the first pilot study. The sound were chosen to represent complex, dynamic and informational events with different temporal patterns (Jenkins, 1985, McAdams, 1993) and edited to a duration allowing for the "*sound event*" or "*sound object*" (Port et al., 1995) to appear to occur naturally (Wightman and Jenison, 1995). In the same manner as the first pilot study the sounds had been classified with particular focus on two categories, the object category of the sound and the action category of the sound. In a similar manner to the first pilot study, the analysis concentrated on the properties of the events, the actions, the objects, and the context of the events. The 300 ms onset-to-onset gap between sound onsets, mentioned in the first pilot study was kept. The set of descriptions for the sounds provided by the participants using their free text responses were also collected. The sounds used in this study are described in Table 5.6 and in Table 5.7, for the 7 & 8 and the 9 & 10 Auditory Icon conditions. In a similar fashion to the first pilot study the sounds selected encompassed a wide range of environmental sounds, to pose both easy and difficult identification problems.

	Prior classification sounds and their descriptions (no sounds in condition with same actions or objects)		***No prior classification (randomly selected) sounds and their descriptions***
ID	*Description*	*ID*	*Description*
1	Rain water in manhole	1	Sweeping floor with brush
2	Dentist drilling and grinding teeth	2	Coffee brewing and filtering
3	Front door modern- A person enters	3	Shower with person- shower curtain
4	Vacuum cleaner- rewinding flex cord	4	Water splashing
5	Pouring Beer	5	Kitchen table being set
6	Venetian Blinds - down and up	6	Water boiling
7	Plates dropped on floor	7	Match being struck and lit
Degree of Overlap	Objects 0% - Actions 0%		Objects 43%-57% - Actions 14%-57%-28%
8	Brushing pants	8	Aquarium with bubbles pump
9	Toaster	9	Window sliding glass opening/closing
10	Light switch - on/off	10	Tab pulled off soft drink can and poured
11	Cooking oatmeal/porrige	11	Person showering in shower
12	Clipping nails	12	Wringing out water from rag
13	Hammering on sheet metal	13	Brushing clothes
14	Eraser on paper	14	Drapes opening/closing
15	Bicycle on asphalt- cycles off- passes- comes in/stops	15	Zipper up/down
Degree of Overlap	Objects 0% - Actions 0%		Objects 50%-12%-38% - Actions 38%-50%-12%

Table 5.6*:* The sounds (1 - 15 of 34) and their descriptions for the 7 and for the 8 concurrent Auditory Icon conditions. The sounds on the left are those, which had been classified to ensure that no object or action properties overlapped in any particular condition. The sounds on the right where randomly selected.

	Prior classification sounds and their descriptions (no sounds in condition with same actions or objects)		*No prior classification (randomly selected) sounds and their descriptions*
ID	*Description*	*ID*	*Description*
16	Beating on heavy door copper clad	16	Waterfall small
17	Window cleaning	17	Waves - long on pebble beach
18	Rice cripsies in bowl adding milk	18	Raining heavily on roof and pavement
19	Shaving electric razor Philips	19	Waves - small on pebble beach
20	Drapes closing then opening	20	Water pouring
21	Whistling tea kettle	21	Plops water
22	Scrubbing floor	22	Rain medium- splattering close-by
23	Small dog barking	23	Milk container opened and poured into tall glass
24	Binder being used	24	Ketchup squirt
Degree of Overlap	Objects 0% - Actions 0%		Objects 89%-11% - Actions 100%
25	Ice cube dropping into glass	25	Hail on window
26	Cloth being torn / being cut	26	Wind across a lakeshore
27	Child laughing - 8 months old	27	Brook/Stream flowing
28	Ice thaw in ocean bay with waves	28	A crackling fire
29	Dog barking- medium	29	Hammering nails- 1 1/4
30	Writing with pencil	30	Bicycle stand up/down
31	Chair on castors	31	Foorsteps on dry snow with creaking- leaving -normal pace
32	Bicycle on asphalt passing with engaged dynamo	32	Water pouring into a container
33	Cigarette lighter	33	Hammering nails
34	Footsteps wooden floor- male walking stairs up/down	34	Footsteps on dry snow with creaking- leaving -fast pace
Degree of Overlap	Objects 0% - Actions 0%		Objects 10%-30%-10%-10%-20%-20% Actions 40%-10%-30%-20%

Table 5.7*:* The sounds (16 - 34 of 34) and their descriptions for the 9 and for the 10 concurrent Auditory Icon conditions. The sounds on the left are those, which had been classified to ensure that no object or action properties overlapped in any particular condition. The sounds on the right where randomly selected.

Experimental Platform - Technical Details The application platform used in this pilot study is similar to the application used in the previous study. Minor modifications where made to the user interface to allow for the additional recording of up to fifteen descriptors by participants. The core functionality, data logging, and application logging were as described in the previous pilot study. A copy of the program and the results from the participants can be found in accompanying DVD for this thesis in Appendix L.1.

Training

The focus of this training phase was to familiarise the participants with the concurrent presentation of Auditory Icons. A training interface allowed participants to stop, start and loop up to seven Auditory Icons whose descriptions were provided on-screen. The training interface is shown in Figure 5.1. The stimuli used in the training phase were not used in the later tests. The participants spent ten minutes listening to the training stimuli using the interface after a short introduction on its operations. They were provided with the task sheet describing the experiment and a verbal explanation was given explaining how the tasks were to conducted using the experimental platform. Participants listened to the sounds (in mono) using headphones while interacting with the system.

Design of the Second Pilot

The second pilot kept the procedure of the first pilot study. The study involved a within-subjects like design where the stimuli were presented in random order within each condition and the task order was counter-balanced for the conditions (seven, eight, nine or ten concurrent Auditory Icons). Each particular set of stimuli was presented randomly as a block for the particular condition and particular subject. This approach was taken from the first pilot study and was kept as it reused the functionality of the existing software. The reason of keeping the similar design was to provide a rapid exploratory pilot study without the need to re-code the software and re-design the procedure used.

Procedure

In a similar fashion to the first pilot study, the participants listened to the recorded sounds (mono) in random order using headphones, responding in free-text format to what each sound was, using the interface shown in Figure 5.11. The conditions varied from seven to ten sounds being concurrently presented. Each condition was played four times in succession. Participants had no control over the playback of sounds and could only use the interface to input their descriptions and move to the next step. The interface below used fifteen description boxes, as the complex scenes with multiple sounds may lead to richer descriptions than just one descriptor per recorded sound.

In a similar fashion to the previous approach, non-directive instructional guidelines (Vanderveer, 1979, p. 83) were given to the participants. The instructions to participants can be seen in Appendix F.1. They were asked to describe a sound but they were not told to describe its agent or source, the action or event, or its perceptual qualities. In the same fashion as the first pilot study, a single rater codified the action and the object descriptors elicited from participants using a two stage process.

In the same manner as the first pilot study, application data logging, participant observation, and post study questionnaires were used to gather the results. The set of participant responses to the presented Auditory Icons were sorted and categorized, as well as evaluated

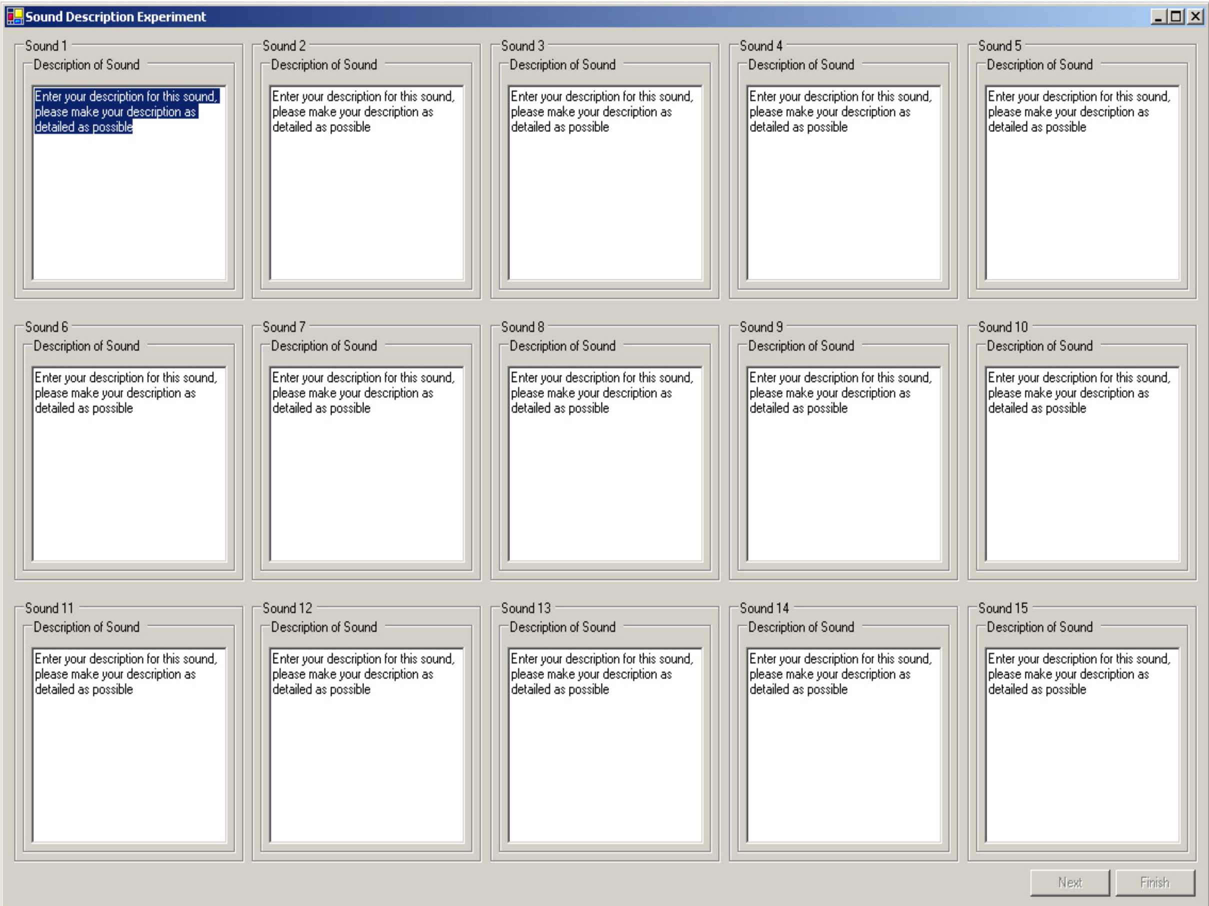


Fig. 5.11: *A screen shot of the dialogue used by participants to fill in the descriptions of parallel sounds presented in the seven, eight, nine, and ten concurrent Auditory Icon conditions.*

for correctness. In the same fashion as the first pilot study the number of Auditory Icons correctly identified by participants was determined by analysing the participant's responses. For each set of (seven to ten) concurrently presented Auditory Icons, the set of Auditory Icons presented and the set of participant responses to those Auditory Icons were compared. If the description of an Auditory Icon from a participant's response matched an Auditory Icon in the set presented, and if that Auditory Icon has not already been identified and matched with a previous description, the number of correctly identified Auditory Icon was increased by one, and the Auditory Icon description was marked as allocated.

Results

The sounds in conditions with prior classification showed a good performance from participants with over 74% identification in three of the four conditions. The results can be seen in Figure 5.13. These are similar to the results of the first pilot study, which suggested that prior classification improves identification. The exploratory trends in the second pilot study showed distinct individual identification performances in Figure 5.14. There is an outlier to this trend with the result of the eight concurrent Auditory Icons with prior classification. It seems that many of the sounds may have been masked, an example of this is how *sound 10* may have been masked by *sound 11* as shown in Figure 5.12. It also indicates that *sound 13* may have also masked a number of sounds. The sounds *9, 10, 12,* and *15* were very poorly recognised as many participants simply did not hear these sounds. Prior classification cannot prevent masking and it is probable that this caused the poor result for this particular condition.

The last two conditions, nine and ten concurrent Auditory Icons, show a widening gap between those sounds that used classification and those that did not. These results have shown that at ten concurrent Auditory Icons, the condition without classification is becoming a cacophony as listeners can only identify 36% of the sounds. This meant that on average between 3 to 4 of the ten sounds presented were identified and shown in Figure 5.14. The condition with prior classification of the ten concurrent Auditory Icons shows a good identification performance from listeners with them identifying approximately 74% of the sounds. These results suggest that nine concurrent sounds are a good end point for studying Auditory Icons, as beyond this point identification is increasingly more difficult where classification has not been used. The percentages of overlap for the categories with no prior classification are shown in Table 5.8. The overlap in these conditions should kept in mind when reviewing the average proportion of correctly identified Auditory Icons.

Number of concurrent Auditory Icons (AIs) presented	Overlap of Actions	Overlap of Objects	Average overlap for number of presented AIs
7	71%	100%	86%
8	62%	75%	*69%*
9	100%	100%	100%
10	80%	90%	85%
Average overlap for condition	**78.25%**	**91.25%**	

Table 5.8*:* Percent of overlapping sound categories within the concurrent auditory presentation conditions in the second study.

Average proportion of correctly identified Auditory Icons For each set of (seven, eight, nine or ten) concurrently presented Auditory Icons, the set of Auditory Icons presented and the set of participant responses to those Auditory Icons were compared. The individual participant results are shown in Table F-3 and in Table F-4 in Appendix F.1. This pilot shows similar trends as shown in the first study with two trends for sounds with and without prior classification. The pilot has highlighted a number of outliers. Object and action causal uncertainty analyses were carried out to see if more insight could be gathered by using triangulation to explore the data from a different angle.

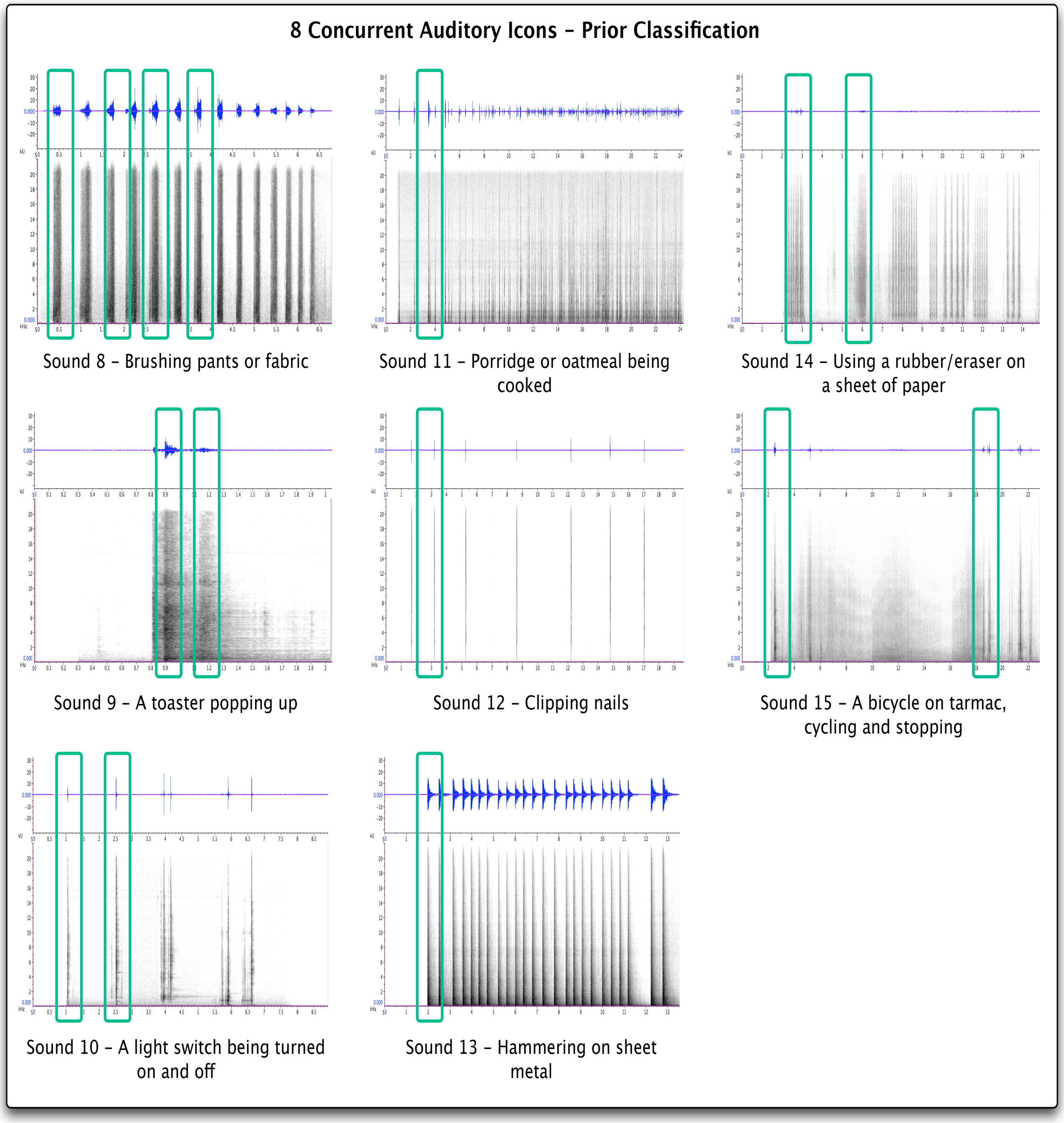


Fig. 5.12: *The eight concurrently presented Auditory Icon condition with prior classification. The highlighted boxes indicate close onsets of sounds with the same or similar action / or object properties. This may have resulted in perceptual merging or masking of the sounds by the listeners. The sonograms used had an FFT size of 512 with a 50% overlap using a Hanning window, from 20Hz to 4.7KHz.*

Reviewing Figure 5.13 and Figure 5.14, one potential anomaly stands out. This is the case of the eight concurrent Auditory Icons with classification; the next section on causal

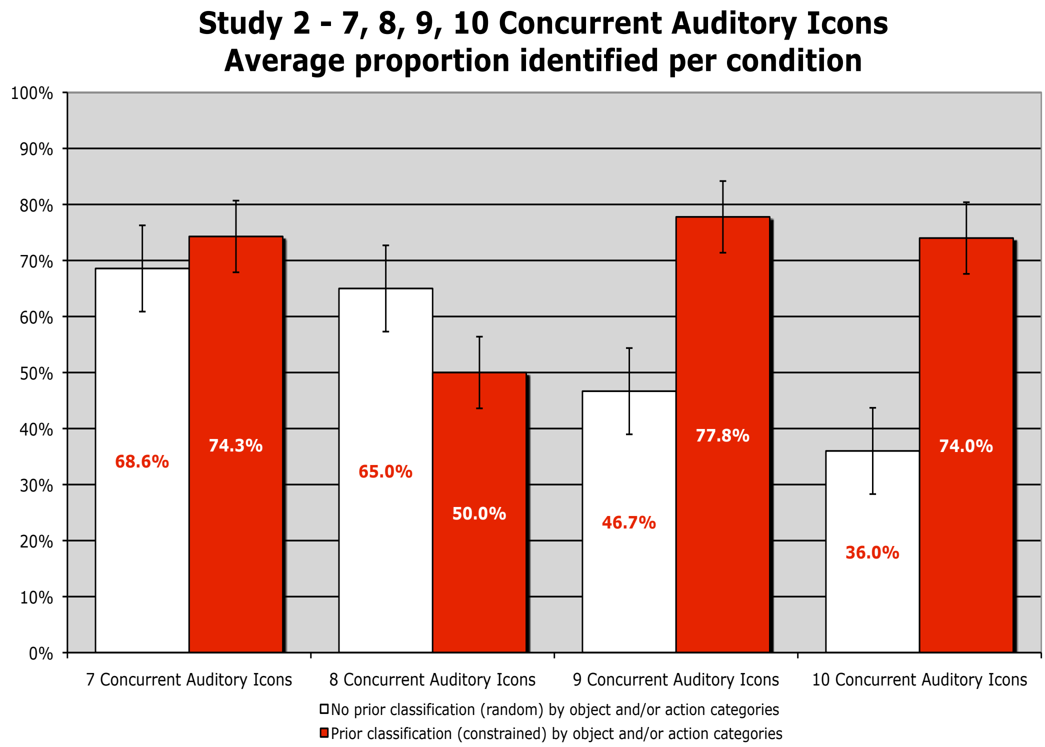


Fig. 5.13: *Graph showing the average proportion of Auditory Icons identified for the seven, eight, nine and ten Auditory Icon conditions. Notice the trends showing that classification does have an effect on identification performance. Prior classification conditions show relatively good identification whereas conditions with no prior classification show a deteriorating performance with a linear relation to the number of sounds being presented.*

uncertainty presents results, which combined with an examination of the sonograms of the sounds involved help to explain the possible reason of this anomaly. Excluding this the results show a general trend where listener identification performance is much improved in cases where classification has been used. The following sections on the analysis of the action and objecthood using the causal uncertainty method to explore the conditions and to provide more explanations explaining issues, which may have affected confusion and identification.

The distribution was not normally distributed which required the use of Kruskal Wallis and Dunn multiple comparison tests (Crawley, 2005) to determine if any of the difference shown in Figure 5.14 were statistically significant. The results of the Kruskal Wallis analysis are shown in Table 5.9. Two differences were found between the Prior Classification 9 vs. No Prior Classification 10 ($p<0.01$) and the Prior Classification 10 vs. No Prior Classification 10 ($p<0.05$) showing that it was easier to identify nine prior classified auditory icons than ten without classification (random selection) and that there was easier to identify ten auditory icons when prior classification is used.

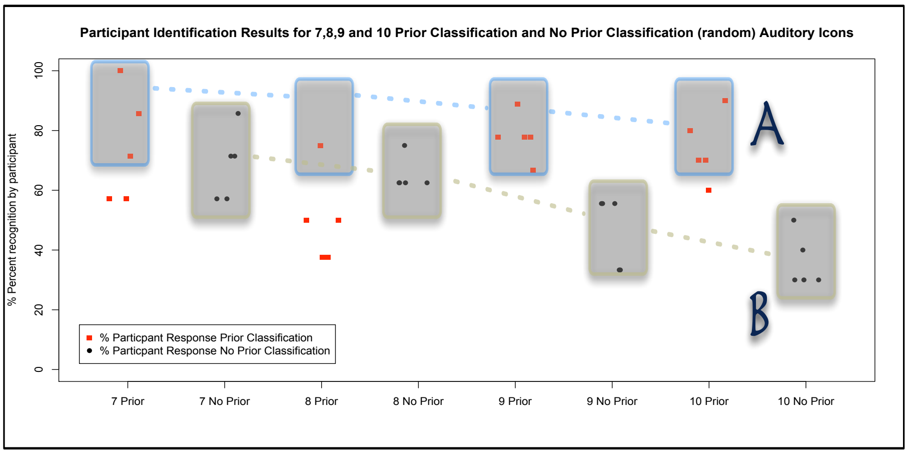


Fig. 5.14: *The identification of the Auditory Icons for the seven, eight, nine and ten Auditory Icon conditions by individual participants. Each dot represents a single participant's identification result for the particular condition. Trend 'A' indicates the reasonably good performance for participants in conditions with prior classification. Trend 'B' shows how listener's identification of sounds in conditions with no prior classification (random selection) starts to drop with the increasing numbers of sounds presented.*

Conditions	*Mean Rank Difference*	*P-value*	*Conditions*	*Mean Rank Difference*	*P-value*
Prior Classification 7 vs. No Prior Classification 7	2.400	ns P>0.05	Prior Classification 8 vs. No Prior Classification 8	−10.600	ns P>0.05
Prior Classification 7 vs. Prior Classification 8	14.800	ns P>0.05	Prior Classification 8 vs. Prior Classification 9	−19.500	ns P>0.05
Prior Classification 7 vs. No Prior Classification 8	4.200	ns P>0.05	Prior Classification 8 vs. No Prior Classification 9	3.100	ns P>0.05
Prior Classification 7 vs. Prior Classification 9	−4.700	ns P>0.05	Prior Classification 8 vs. Prior Classification 10	−16.100	ns P>0.05
Prior Classification 7 vs. Prior Classification 10	17.900	ns P>0.05	Prior Classification 8 vs. No Prior Classification 10	7.900	ns P>0.05
Prior Classification 7 vs. No Prior Classification 10	22.700	ns P>0.05	No Prior Classification 8 vs. Prior Classification 9	−8.900	ns P>0.05
			No Prior Classification 8 vs. No Prior Classification 9	13.700	ns P>0.05
No Prior Classification 7 vs. Prior Classification 8	12.400	ns P>0.05	No Prior Classification 8 vs. Prior Classification 10	−5.500	ns P>0.05
No Prior Classification 7 vs. No Prior Classification 8	1.800	ns P>0.05	No Prior Classification 8 vs. No Prior Classification 10	18.500	ns P>0.05
No Prior Classification 7 vs. Prior Classification 9	−7.100	ns P>0.05	Prior Classification 9 vs. No Prior Classification 9	22.600	ns P>0.05
No Prior Classification 7 vs. No Prior Classification 9	15.500	ns P>0.05	Prior Classification 9 vs. Prior Classification 10	3.400	ns P>0.05
No Prior Classification 7 vs. Prior Classification 10	−3.700	ns P>0.05	**Prior Classification 9 vs. No Prior Classification 10**	27.400	** P<0.01
No Prior Classification 7 vs. No Prior Classification 10	20.300	ns P>0.05	No Prior Classification 9 vs. Prior Classification 10	−19.200	ns P>0.05
			No Prior Classification 9 vs. No Prior Classification 10	4.800	ns P>0.05
			Prior Classification 10 vs. No Prior Classification 10	24.000	* P<0.05

Table 5.9: Kruskal-Wallis and Dunn multiple comparison test results for the seven, the eight, the nine, and the ten concurrent Auditory Icons conditions.

The results of the second pilot study show that even with ten simultaneous sounds, the identification results of Auditory Icons are still quite high when classification is used. The pilot studies informed a follow up probe, with the results from the first two pilot studies pointing to three, six, and nine concurrent Auditory Icons as the best conditions of everyday sound concurrency to explore. In order to provide greater detail about the difficulties in this pilot study, the next sections look at an analysis of the participant's textual descriptors using the causal uncertainty method. This provides details on the confused sounds and supported the hypothesis from the first study that listeners use both object and actionhood as salient criteria in identification.

Results and Observations for Participant Object Descriptions Of 7 To 10 Concurrent Auditory Icons The set of participant responses (SPR) to the presented Auditory Icons were sorted and categorized, as well as evaluated for correctness. From the responses of participants, the object segments of the texts were extracted and categorized as shown in Appendix F.1 in Table F-6 and in Table F-9, such as what objects/materials were involved in the interaction. The responses from participants for the object conditions, with prior and with no prior (random) classification, are shown in Table F-5 and in Table F-8.

The causal uncertainty results of the object descriptors for the 7 to 10 conditions are shown in Table F-7 and in Table F-10 in Appendix F.1. Examining Figure 5.15 shows two conditions with high causal uncertainties. The first condition is the 8 concurrent Auditory Icon with no prior classification, where the causal uncertainty had a cumulative rating above 7.

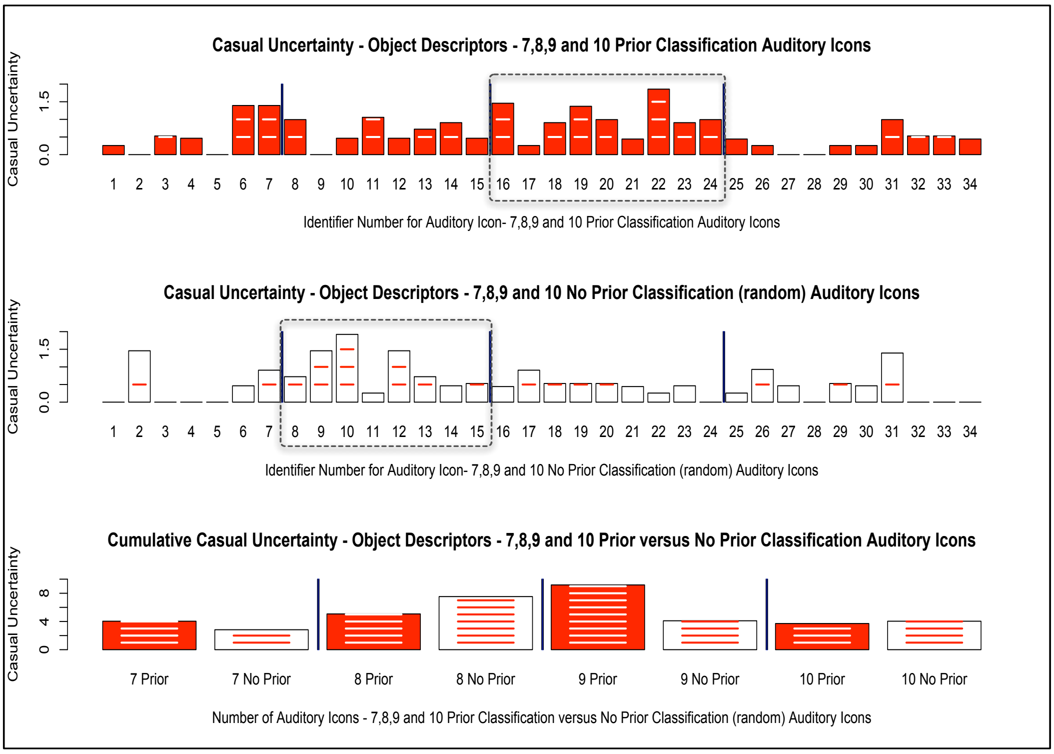


Fig. 5.15: *The analysed object descriptors results of participants, 7 to 10 Auditory Icons with prior and with no classification. The two groups of interest, the eight concurrent Auditory Icons with no prior classification condition and the nine concurrent Auditory Icons with prior classification are highlighted.*

Examining the 8 concurrent Auditory Icon condition with no prior classification shows one sound, *sound 10* with a causal uncertainty just below 2. Examining its descriptor, as shown previously Table 5.6, this is the sound of a tab being pulled from a soft drink can and the drink being poured. Considering this sound and the other sounds in the condition, it is possible that the liquid sound was somewhat confused with the other existing liquid type sounds and there is the potential that it may have been masked to a certain degree. A cursory examination of Figure 5.16 would suggest that *sound 10* may have been masked by the sounds *9* and *11*. The results for the sounds is shown in Table F-7 show that causal uncertainty measures of 1.4575, 1.9219, and 0.2575 respectively with identification of the sounds by four, five, and four of the five participants. These results show that the sounds were identified but confused rather than being masked. One of the participants remarked that "*I feel I can distinguish the sounds, but its what I make up*" which indicates in case like this instance of high confusion, a best guess will be made. This is supported by the theory of auditory scene analysis from Bregman (1990) where the best option will be taken for the identification of a sound scene until some other piece of information forces a re-evaluation.

The results of the second interesting condition, that being the 9 concurrent Auditory Icon with prior classification, had a causal uncertainty with a cumulative rating above 9. Three sounds in this condition had particular high causal uncertainty measures with *sound 22*, *sound 16*, and *sound 19* having measures above 1.5 for *sound 22* and above 1 for both *sound 16* and *sound 19*. The descriptors for the sounds is shown in Table 5.7 and considering *sound 22* in the context of the other sounds, there is a possibility that due to the other sounds in the condition that this sound may simply have been masked by *sound 21*. Examining the sonograms and waveforms for the sounds can highlight possible merging or masking. Looking at Figure 5.17 it can be seen in addition to *sound 21* and *sound 22* that *sound 16* may have be masked by *sound 17*. In this latter case, the waveform shows a number of very close onsets, which may have lead listeners to merge the two sounds together as one perceptual unit. In the case of such complex auditory conditions, it is difficult to determine exactly what problems are occurring with the identification of the sounds. This relates to a comment from a participant who said they found it "*hard to distinguish between the sounds, some are easier to distinguish but I wasn't sure on variants of the sound (if similar but different sound), I kept trying to link it to something in the real world which was difficult*". A number of other participants found it simply a "*a cacophony of sounds*" and difficult to identify any sounds in the more sound dense conditions. The methods in this thesis will highlight the appropriate and the inappropriate combinations of sounds allowing for the selection of alternative sounds for dense concurrent sound conditions. This is an advance on the *ad hoc* selection of Auditory Icons based on instinct and experience as this method can help both novice and experience Auditory Display researchers and designers.

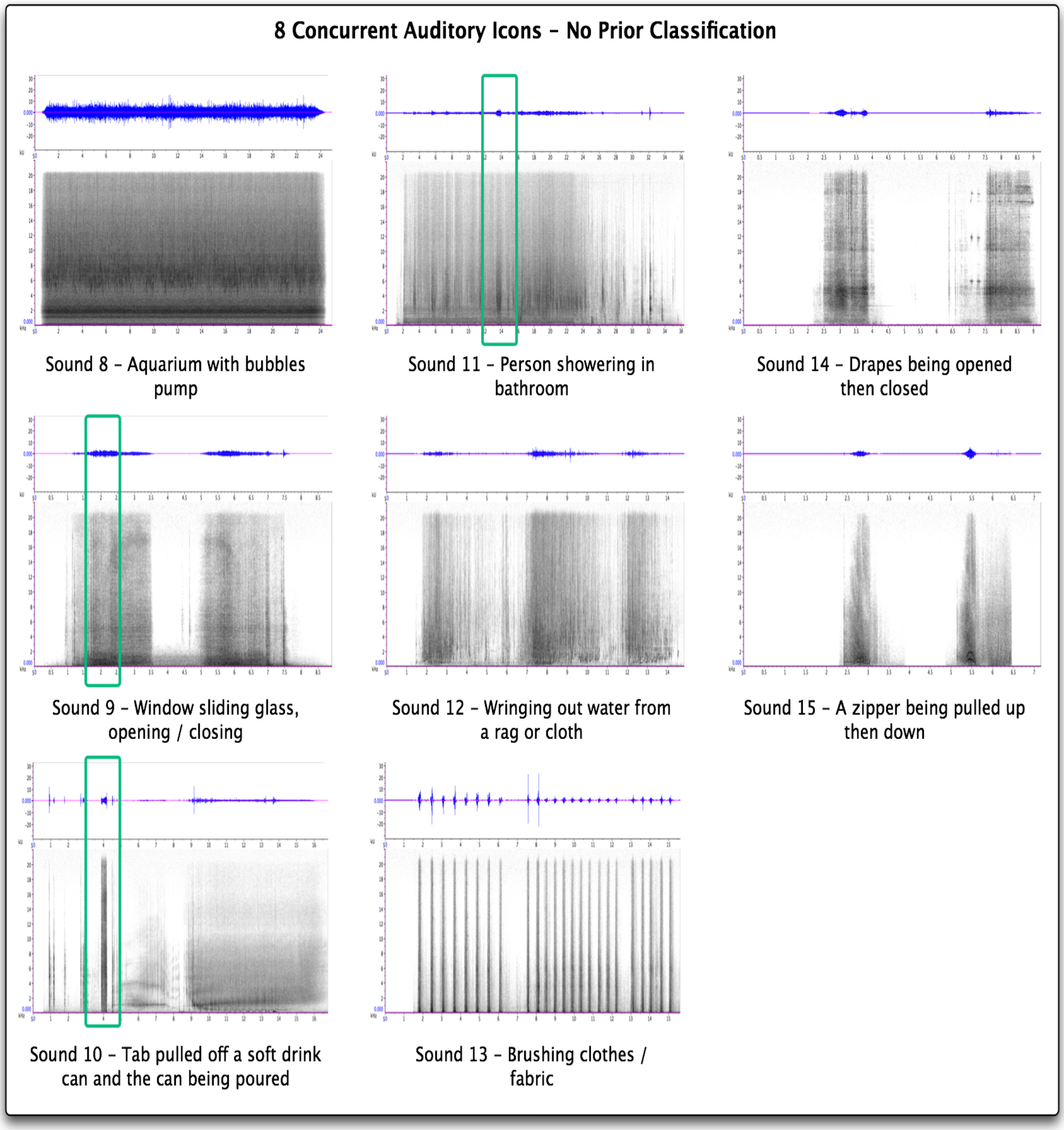


Fig. 5.16: *The eight concurrently presented Auditory Icon condition with no prior classification. The highlighted boxes indicate close onsets of sounds with the same or similar action and / or object properties. This may have resulted in perceptual merging or masking of the sounds by the listeners. The sonograms used had an FFT size of 512 with a 50% overlap using a Hanning window, from 20Hz to 4.7KHz.*

The results of the higher number of concurrent Auditory Icons (7, 8, 9, and 10) show there is a greater degree of confusion among listeners' identification of sounds. One interesting

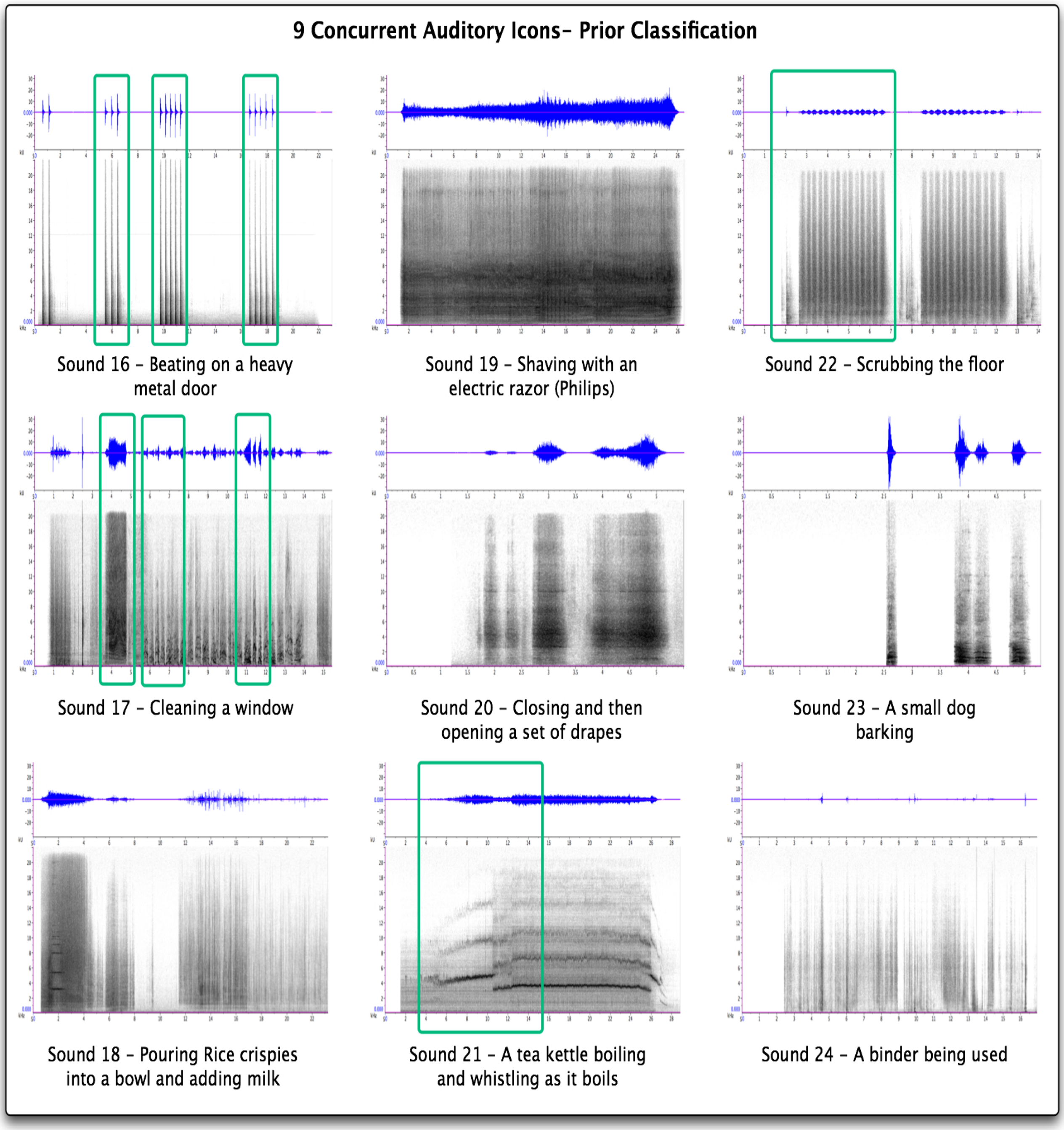


Fig. 5.17: *The nine concurrently presented Auditory Icon condition with prior classification. The highlighted boxes indicate close onsets of sounds with the same or similar material properties. This may have resulted in perceptual merging or masking of the sounds by the listeners. The sonograms used had an FFT size of 512 with a 50% overlap using a Hanning window, from 20Hz to 4.7KHz.*

finding was that the 9 non-overlapping condition had a much higher causal uncertainty than the corresponding 9 overlapping condition.

The non-overlapping condition had compound causal uncertainty measures of 1.021 for objecthood and of 0.772 for actionhood versus the overlapping condition with compound measures of 0.455 for objecthood and of 0.667. This would point towards a higher degree of confusion in the condition with prior classification but that the confusion of the sounds did not affect its identification (77.8%). The lower confusion measures for the overlapping confusion did not correlated to a higher identification, as the average identification was 46.7% for this condition. An explanation for this seeming contradiction can be found in Table F-9 and in Table F-10 for the nine Auditory Icons without prior classification. Looking at the results and in particular, at *sound 23*, which only a single person identified and *sound 24*, which was not heard by any of the listeners. This problem highlights one issue when using the causal uncertainty method; it works best only when listeners have identified some or all of the sounds. Examining the formula of causal uncertainty 5.1.1, it can be seen that a zero value for p, the number of identifications given by listeners, will mean that the result for the causal uncertainty will be zero. However, in this case of a consensus result it does not mean that all participants shared the same interpretation of the sound rather the listeners simply did not hear this sound or that they could not identify it. This shows it is important as part of an analysis when using this method to verify the number of sounds identified in a condition when considering a causal uncertainty result. The identification results (78%) suggest that the objecthood of the sounds should be exaggerated or made cleaner by putting it into context with related sounds. This can potentially ease the understanding of the sounds and their objecthoods.

Considering the causal uncertainty results of the 7, 8, 9, and 10 concurrent presentation conditions, objecthood is not as easily distinguishable as in the previous study and this can be easily seen in the second cumulative causal uncertainty graph in Table 5.15. It is important to note the scale used on the cumulative causal uncertainty graph in Table 5.15 as this graph has a range from 0 to 10. The length of this scale and the increases in both the number of concurrent sounds and the difficulty in interpreting their objecthood is evident. The information in Figure 5.13 shows the increasing difficulty in identification as the number of sounds increase is an additional indicator of this.

Results and Observations for Participant Action Descriptions Of 7 To 10 Concurrent Auditory Icons In a similar fashion to the previous object description analysis, the participants' responses were extracted and categorized by action segments of the texts as shown in Appendix F.1 in Table F-13 and in Table F-16. The responses from participants for the action conditions, with prior and with no prior (random) classification, are shown in Table F-11 and in Table F-14. The causal uncertainty results of the action descriptors for the 7 to 10 conditions are shown in Table F-12 in Table F-15 in Appendix F.1. Examining Figure 5.18 and the summary of results of the causal uncertainty measures for the 7, 8, 9, and 10 concurrent presentation conditions, the results show an increase in the confusion for sound action interpretation in both types of conditions, those with prior classification and those with no prior classification. These results re-iterate the caveat about the use of the causal uncertainty method without the identification results as shown in Figure 5.13 or in Figure 5.14. The lower identification rates for overlapping Auditory Icons (47% and 36%) versus the rates for non-overlapping Auditory Icons (78% and 74%) in both the 9 and 10 concurrent presentation conditions directly affect the causal uncertainty calculation as noted in the previous object description analysis.

The 9 overlapping Auditory Icon condition and the first seven sounds in this condition had varying degrees of confusion as shown by their respective causal uncertainty measures in Figure 5.18. One explanation for this comes from a participant who said ”*the water sounds were hard to contextualise*”. This makes sense examining the results in Tables F-15, and F-16 which showed that most participants heard only some of the sounds and that there was a number of different interpretations of those sounds.

Examining the 10 overlapping Auditory Icon condition and the *sounds 28*, *32*, *33*, and *34* as highlighted in Figure 5.18. These four sounds were all found to have a causal uncertainty of 0. These results need to be considered in terms for the 36% identification rate as an unheard sound will have 0 participant responses and a resulting causal uncertainty of 0. This means that these lower identification confusion ratings are due to the fact that many listeners did not hear the sounds rather than the sounds being clear and unambiguous. This highlights the importance of carrying out both types of analysis (identification and causal uncertainty) on a participant's written descriptions to establish clearly if the sound's actionhood (or objecthood) was clearly identified or if the sound was simply unheard as happened in the more dense concurrent presentation conditions.

Post study participant debriefing A debriefing was carried out after the second pilot study to collect data from the participants. Participants highlighted the particular difficulties they found in the more dense 7 to 10 Auditory Icon conditions. They felt it was like “*a cacophony of sounds*” and that it was “*hard to distinguish between the sounds, some are easier to distinguish but I wasn't sure on variants of the sound (if similar but different sound), I kept trying to link it to something in the real world which was difficult*”. This and similar comments such as “*the mind will substitute to what it thinks it is hearing*” or hearing “*certain related sound, creating a script where each feed into the other*” indicate how the participants find it easier to associate sounds to a particular sequence of events. In the case of conditions where unusual

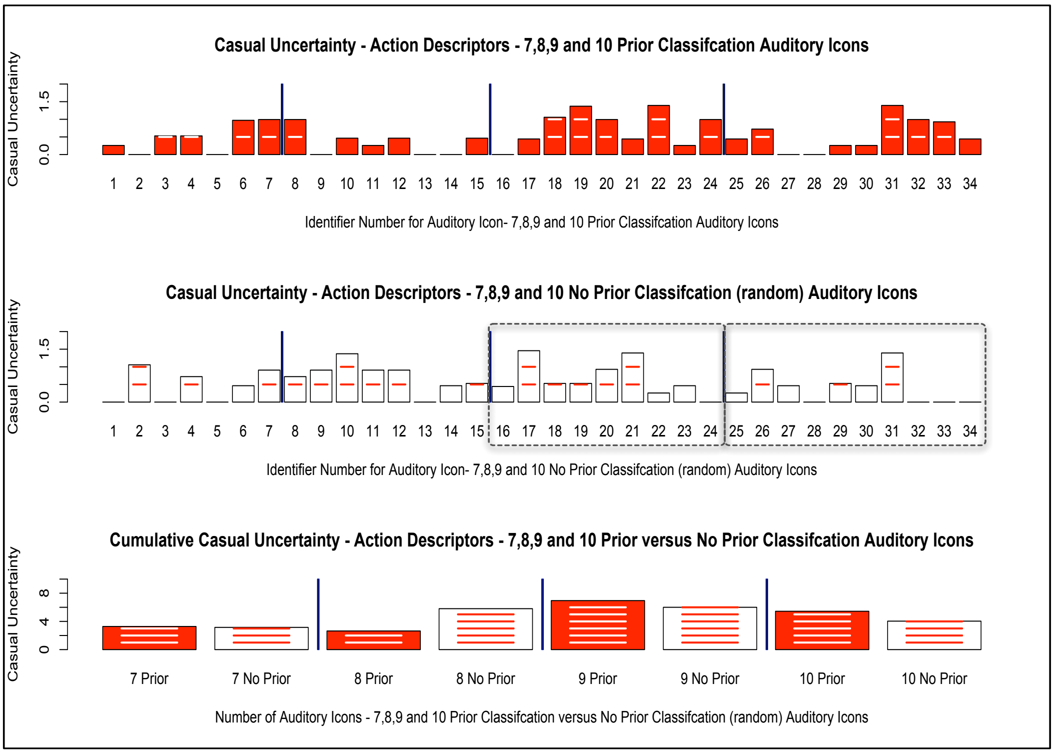


Fig. 5.18: *The analysed action descriptors results of participants, 7 to 10 Auditory Icons with prior and with no classification.*

or non-natural sequences of sounds occurred, the difficulty of identification increased. Participants comments helped indicate in this case how they listened to the auditory scene and how their approaches for how they organised the sounds in their head. However, the comments do not give a detailed insight into the mappings and categorisation used. This result was one of the rationales for carrying out the studies in the next chapter as a means of providing a more detailed understanding into the mappings and categorisations.

Results and implications from the second pilot study of 7 to 10 concurrent Auditory Icons The results from this study indicate that increasing the number of Auditory Icons increases the difficulty in identifying them. In particular, it shows that there is a distinct performance advantage between conditions with similar numbers of Auditory Icons but which have been classified rather than those, which were not classified. In particular, for 9 concurrent Auditory Icons the identification rate was still approximately 78%. This indicates that the task was achievable by participants, but that it was not trivial task. As a limit and given the results of this pilot study, it was decided that nine would be sufficiently difficult a number of sounds to identify and this was used as the maximum number of sounds to be presented in the next exploratory study. The

In a similar fashion to the first study the results of this pilot study indicates that prior classification can be an effective approach for sound selection when designing Auditory Displays with many concurrent sounds. Ballas's method of causal uncertainty (1986) can be used to further refine the sound selections by determining which sounds are most confused. These sounds can be removed and replaced with less confusing and more perceptually distinct sounds. Areas of possible masking may be highlighted by the use of waveform and sonogram views of the conditions. This type of visualisation can be used in addition to prior classification and causal uncertainty to help in determining sound selections.

5.1.4 Results of the pilot studies

The results of the action and object identification showed both are generally well identified even in the cases of conditions with higher numbers of concurrently presented sounds. Across the concurrent presentation conditions for the full range of 3 to 10 sounds the action identification of the Auditory Icons was better identified as shown when the results in Figures 5.6 and 5.15 (actions) and in Figures 5.18 and 5.8 (objects) are compared. It is important to note the results from the average identification rates as in the conditions with higher numbers of sounds there is the possibility of sounds being unheard or simply masked exists. This means that these lower identification confusion ratings for the causal uncertainty measure may be due to the fact that many listeners did not hear the sounds rather than the sounds being clear and unambiguous.

Future investigations could determine if applying acoustic comparisons such as computational auditory scene analysis[2] of the sounds within a condition to determine potential masking would improve the action or object identification of the sounds within the condition.

However, given the results of this study it appears to show that in many of the conditions it was a single sound or couple of sounds (in the conditions with 7+ Auditory Icons) that were responsible for the majority of participant confusion. This could indicate that the use of masking analysis or other CASA type technique may not be as important as determining the sounds with particularly high causal uncertainties. These sounds could then be replaced with similar but alternative sounds, which are more easily identified.

5.1.5 Limitations of the Pilot Studies

The goal of the first and second pilot studies was mainly to gauge and estimate if there was a potential for a further exploratory probe study with a focus on providing a better estimate

[2]There is a body of work on computational auditory scene analysis (Ellis, 1996, van der Kouwe et al., 2001, Wrigley and Brown, 2004) (CASA) attempting to provide acoustic comparisons for these issues. The research to date has proven inconclusive and does not offer any solutions that could be generalised to Auditory Icons at this point in time. CASA is a dynamic research fields but it is currently unable to take a composite sound, one that consists of several individual sounds, and answer how it will be processed and streamed by a listener.

with regard to the a maximum number of sounds that can be reasonably used for concurrent Auditory Icon presentation. A further issue relates to the sounds themselves as they were all monophonic high quality sounds (44.1 Kilohertz 16-bit), which may have affected their identification as the sounds did not contain stereo or spatial information for listeners.

The results from both pilot studies provided a useful estimation of the potential range that should be explored. This helped inform the choice of three, six, and nine concurrent Auditory Icons as the areas to explore in the exploratory probe. The pilot studies pointed to the need for a deeper exploratory and the participant pool size was more than doubled to help provide a greater depth when compared to the pilots. Another limitation is the applicability of the studies in this chapter. It will not be possible to claim, simply by highlighting an increase in participant's ability to identify concurrently presented Auditory Icons, that prior classification will lead to performance improvements in real interfaces using concurrent Auditory Icons. Real interfaces will require further studies to determine if there are workload issues or other performance issues that may arise. The issue of subjective workload needs to be considered as there may be combinations of Auditory icons that cause increased workload or that cause psychological distress (Hart and Staveland, 1988). A possible approach to dealing with this is to use the NASA Task Load Index (TLX) scales (Hart and Staveland, 1988, Gawron, 2000) to rate the subjective workload experienced by participants. However, this type of approach is more suited to exploring a real interface rather than a set of potential sounds as explored in these studies.

The coding of the object and action descriptors used in the pilot studies have two limitations. The coding was performed by this thesis's author only, and the categorisation of the participant's free text responses represented a simplification of the complex descriptions given. Important nuances might have been lost and some descriptions may have been "simplified". This categorisation was based on earlier work by Fernström et al. (2005), it also draws upon work by Houix et al. (2007b) and Houix et al. (2007a), which used similar categorisation and coding techniques. Houix et al. (2007a) analysed the correlations between classifiers and found a degree of dissimilarity in rater classification for only a small number of sounds. The correlation results from earlier studies and prior experiences with this approach suggest that using a single codifier or sorter should not be a major issue but further studies are required to clarify this issue.

These areas need wider research and exploration to understand if they contribute either negatively or positively to the results. In order to overcome the limitations, a third study using an exploratory probe was conducted. It concentrated on addressing a number of the issues raised. Three larger pools of stimuli were used and a new procedure was designed to address the issues raised.

5.1.6 Study 3 - Exploratory probe examining three, six, and nine concurrent Auditory Icons

This exploratory probe addresses a number of the limitations with the pilot studies by changing the procedure, increasing the stimuli pool size and increasing the number of pools from two to three. It uses a larger participant group to provide a better insight into the issue of identification of concurrent Auditory Icons. In this particular exploratory probe, the number of concurrent Auditory Icons being presented varied from three to six to nine in two conditions, one where the stimuli was limited to no overlapping action or object descriptors (constrained) and the second condition allowed random selection of stimuli from a pool with no conditions. This probe used three instead of two pools of stimuli with each pool consisting of 63 distinct stimuli. Stimuli did not overlap between pools. The hypotheses for this probe were essentially the same as those in the first and second pilot studies.

Hypothesis 1. *The participants perform better in identifying the sounds in conditions which had been selected to ensure that no two sounds in a condition had the same object or action properties based on a classification of the sound's descriptors.*

Hypothesis 2. *The performance of participants with regard to identification would degrade as the number of Auditory Icons presented increased.*

Hypothesis 3. *The identification performance of participants would degrade more in conditions which permitted sounds to have similar object or object classifications when compared to the other conditions that prevented this.*

Hypothesis 4. *The object and actionhood of sounds are salient criteria used by participants for the identification of sounds.*

This exploratory probe was a deeper exploration than the prior two pilot studies. It hypothesised that Auditory Icons would be effective at communication information simultaneously. In particular, it was hypothesised that using prior classification would improve the identification performance of the sounds.

Participants

26 participants were recruited from the wider community of postgraduates and staff at the University of Limerick. All participants reported normal hearing and had normal or corrected to normal vision. Written consent was obtained prior to the study from all participants. Two participants were dropped from the analysis as they had previously taken part in the earlier studies.

Stimuli

Three different sets (pools A,B, and C) of 63 high-quality monophonic sounds (44.1 Kilohertz 16-bit) everyday sounds (durations between 1.0 and 5.0 seconds) were used for this study. These sounds were selected from the BBC sound effects CD collection (BBC, 2006), the online creative-commons Freesound (www.freesound.org, 2007), a local sound collection in the University of Limerick[3], and a commercial sound effects CD collection (Hollywood-Edge, 1990). These sounds used were selected in a similar manner to the last study. The sound were chosen to represent complex, dynamic and informational events with different temporal patterns (Jenkins, 1985, McAdams, 1993) and edited to a duration allowing for the "*sound event*" or "*sound object*" (Port et al., 1995) to appear to occur naturally (Wightman and Jenison, 1995). In the same manner as the previous study the sounds had been classified with particular focus on two categories, the object category of the sound and the action category of the sound. In a similar manner to the previous study, the analysis concentrated on the properties of the events, the actions, the objects, and the context of the events. The 300 ms onset-to-onset gap between sound onsets, mentioned in the last study was kept. The set of descriptions for the sounds provided by the participants using their free text responses were collected and are shown in Appendix G.1 starting at Table G-18 for Participant 1 to Table G-43 for Participant 26. The sounds used in this study are described in Appendix G.1 in Tables G-1 and G-2 for Pool 'A', in Tables G-3 and G-4 for Pool 'B', and in Tables G-5 and G-6 for Pool 'C'. In a similar fashion to the previous studies the sounds selected encompassed a wide range of environmental sounds, to pose both easy and difficult identification problems.

Experimental Platform - Technical Details The application platform used in this study is similar to the application used in the previous study. The core functionality, data logging, and application logging were as described in the previous study. A copy of the program and study results from the participants can be found in accompanying DVD for this thesis in Appendix L.1.

Training

The focus of this training phase was to familiarise the participants with the concurrent presentation of Auditory Icons. A number of sets of concurrent Auditory Icons were presented to the listener using a standard media player. The stimuli used in the training phase were not used in the later tests. The participants spent ten minutes listening to the training stimuli and

[3]Interaction Design Centre, University of Limerick, Ecological Sounds Collection - http://www.idc.ul.ie/mikael/sounds/ecosound.zip

were then introduced to the platform used for the study. The training stimuli used in this study were layered sound's each representing a complex soundscape consisting of several complex concurrent events. The participants were provided with the task sheet describing the experiment and a verbal explanation was given explaining how the tasks were to conducted using the experimental platform. Users had headphones to listen to the sounds (in mono) while interacting with the system.

Design of Study

Using a within-subjects design, the stimuli were presented in random order within each condition and the task order was counter-balanced for the conditions (three, six, and nine concurrent Auditory Icons). Each set of stimuli was presented randomly and as a single block for the particular condition. Stimuli were selected from two of the three pools randomly selected for each participant. One of the pools was used for the constrained sounds (non-overlapping action or object descriptors) and the second pool was used for the randomly selected sounds. The selection of the sounds was carried out using an iterative search algorithm, in the case of the constrained sounds, this ensured that no two sounds in a particular condition would be the same sound or have any overlap in their action or object descriptors. The randomly selected sounds were chosen without replacement from the particular stimuli pool for each condition. Each pool consisted of 63 stimuli, which are described in Appendix G.1, with each participant hearing a total of 36 sounds, 18 from the first pool (constrained) and 18 from the second pool (random). The degree of overlap of the randomly selected sounds for this study is shown in Table 5.10, there was no overlap in the sounds used in the constrained conditions.

Procedure

In a similar fashion to the previous study, the participants listened to the recorded sounds (mono) in random order using headphones, responding in free-text format to what each sound was, using the interface shown in Figure 5.19. The conditions varied from three to six to nine sounds being concurrently presented. Each condition was played four times in succession. Participants had no control over the playback of sounds and could only use the interface to input their descriptions and move to the next step. The interface below used fifteen description boxes, as the complex scenes with multiple sounds may lead to richer descriptions than just one descriptor per recorded sound.

In a similar fashion to the previous approach, non-directive instructional guidelines (Vanderveer, 1979, p. 83) were given to the participants. The instructions to participants can be seen in Appendix G.1. They were asked to describe a sound but they were not told to describe its agent or source, the action or event, or its perceptual qualities. In the same fashion as the earlier studies, a single rater codified the action and the object descriptors elicited from participants using a two stage process.

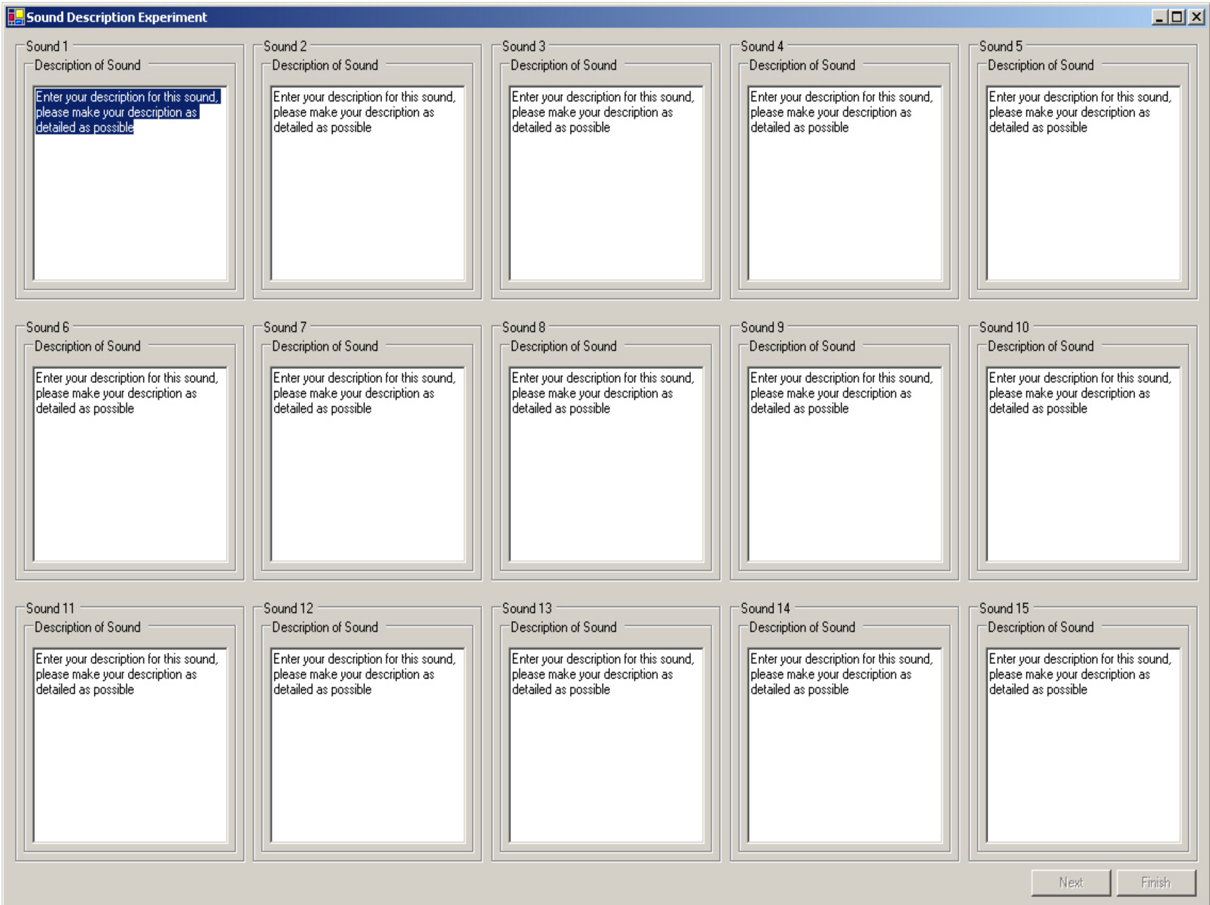


Fig. 5.19: *A screen shot of the dialogue used by participants to fill in the descriptions of parallel sounds presented in the three, six, and nine concurrent Auditory Icon conditions.*

In the same manner as the previous study application data logging, participant observation, and post study questionnaires were used to gather the results. The set of participant responses to the presented Auditory Icons were sorted and categorized, as well as evaluated for correctness. In the same fashion as the previous study the number of Auditory Icons correctly identified by participants was determined by analysing the participant's responses. For each set of (three, six, or nine) concurrently presented Auditory Icons, the set of Auditory Icons presented and the set of participant responses to those Auditory Icons were compared. If the description of an Auditory Icon from a participant's response matched an Auditory Icon in the set presented, and if that Auditory Icon has not already been identified and matched with a previous description, the number of correctly identified Auditory Icon was increased by one, and the Auditory Icon description was marked as allocated.

Results

The overall identification results are shown in Figure 5.20 and support the hypothesis that increasing the number of sounds in conditions with no classification increases the difficulty in identification. A breakdown of the results by participant and by condition is shown in Appendix G.1 in Table G-7. The sounds in conditions with prior classification show a better identification performance by participants. This supports the idea that prior classification improves identification. The overall trends in this study are more distinct when the individual identification performances are considered as shown in Figure 5.22. These results show a worse performance in comparison to the earlier studies and indicate that Auditory Displays should limit the number of concurrent Auditory Icons to range of between three (75%) and six (57%) concurrent Auditory Icons with prior classification. Figure 5.20 shows approximately a 7% identification improvement across the conditions where prior classification was used.

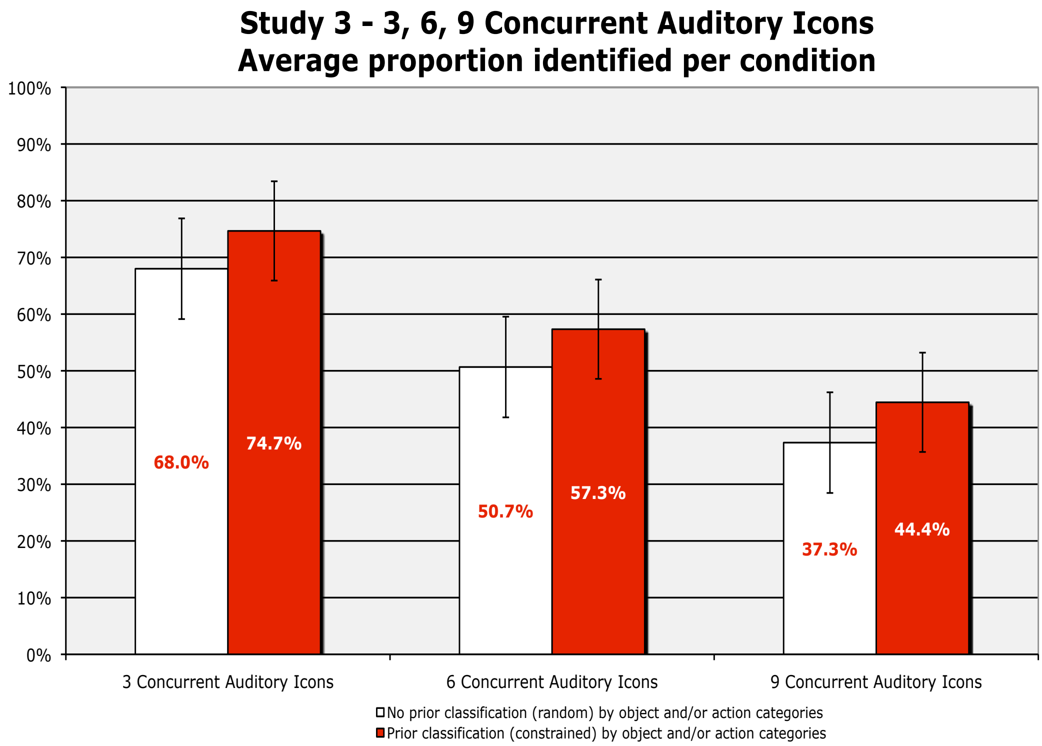


Fig. 5.20: *Graph showing the average proportion of Auditory Icons identified for the three, six, and nine Auditory Icon conditions. The number of Auditory Icons can be seen to have a similar effect on identification performance in both types of condition. Increasing the number of Auditory Icons leads to a monotonic trend of decreased identification in both types of condition, however prior classified conditions show a 7% improvement in identification over those with no prior (random) classification.*

Examining Figure 5.22 shows the identification per participant of the concurrent Auditory Icons conditions. The overall trend shows that for smaller numbers of sounds, listeners can easily segregate them. The trend 'A' (upper dotted line) shows the prior classification results and trend 'B' (lower dotted line) shows the no prior classification results. The results for type of conditions are similar at the lower numbers of concurrently presented Auditory Icons but as more and more concurrently presented Auditory Icons are added, an improve can be seen in the identification of Auditory Icons using prior classification.

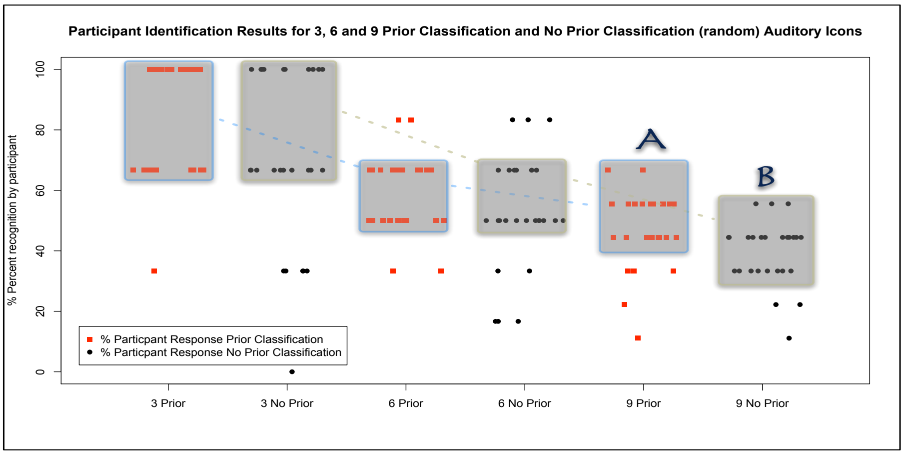


Fig. 5.21: *The identification of the Auditory Icons for the three, six, and nine Auditory Icon conditions by individual participants. Each dot represents a single participant's identification result for the particular condition. Trend 'A' (upper dotted line) indicates the reasonably good performance for participants in conditions with prior classification. Trend 'B' (lower dotted line) shows listeners identification in conditions with no prior classification (random selection).*

A possible explanation for the difference between the results in this study and from the two earlier pilot studies may be due to the lower numbers of stimuli that overlap in the conditions as shown in Table 5.10. The random selection of stimuli in the third study was taken from three larger pools than used in the earlier two studies and did not seek to increase the degree of overlap in either the action or object categories. The earlier two studies had much higher degrees of overlap of both action and object categories in the conditions without prior classification as shown in Table 5.11. Examining the three, six, and nine action and object categories from the first and the second studies, as shown in Table 5.11, a 100% overlap can be seen for both types of categories in all conditions. The overlap results for the third study, as shown in Table 5.10, varies from 38%–51%–42% for the actions and 20%–46%–57% for the objects in the respective, three, six, and nine conditions. Further large scale studies with highly overlapping action and object categories should be conducted to explore the applicability of this classification approach. The indications from Table 5.11 do suggest that the greater the overlap between the action and the object categories in conditions with no prior classification, the lower the identification rate for Auditory Icons in the condition.

Participant	Random Act 3	Random Obj 3	Random Act 6	Random Obj 6	Random Act 9	Random Obj 9	Average overlap for participant
1	0%	0%	33%	33%	22%	44%	22%
2	0%	0%	33%	33%	89%	67%	37%
3	0%	67%	50%	67%	0%	67%	42%
4	0%	0%	50%	50%	44%	44%	31%
5	0%	67%	50%	67%	0%	67%	42%
6	0%	0%	67%	67%	22%	44%	33%
7	67%	0%	0%	33%	44%	56%	33%
8	67%	0%	0%	33%	22%	44%	28%
9	67%	0%	33%	0%	33%	67%	33%
10	67%	0%	83%	33%	44%	56%	47%
11	100%	0%	67%	33%	22%	67%	48%
12	67%	67%	67%	83%	89%	67%	*73%*
13	0%	67%	67%	50%	78%	44%	51%
15	0%	0%	67%	67%	56%	78%	44%
17	67%	0%	33%	33%	22%	44%	33%
18	67%	67%	67%	83%	44%	78%	68%
19	0%	67%	67%	33%	67%	44%	46%
20	67%	0%	0%	0%	22%	22%	19%
21	0%	67%	83%	67%	22%	44%	47%
22	67%	0%	0%	33%	33%	56%	31%
23	100%	0%	83%	67%	56%	78%	64%
24	0%	0%	83%	33%	44%	22%	31%
25	0%	0%	50%	33%	67%	78%	38%
26	100%	0%	83%	67%	56%	78%	64%
Average overlap for condition	**38%**	**20%**	**51%**	**46%**	**42%**	**57%**	**42%**

Table 5.10*:* Percent of overlapping sound categories within the 3, 6, and 9 concurrent auditory presentation conditions in the third study. Participant's 14 and 16 were dropped from the study.

Number of concurrent Auditory Icons (AIs) presented	Overlap of Actions	Overlap of Objects	Average overlap for number of presented AIs
3	100%	100%	100%
4	75%	100%	88%
5	80%	100%	90%
6	100%	100%	100%
7	71%	100%	86%
8	62%	75%	*69%*
9	100%	100%	100%
10	80%	90%	85%
Average overlap for condition	**83.5%**	**95.6%**	

Table 5.11*:* Percent of overlapping sound categories within the concurrent auditory presentation conditions in the first and in the second pilot studies.

Average proportion of correctly identified Auditory Icons For each set of (three, six, and nine) concurrently presented Auditory Icons, the set of Auditory Icons presented and the set of participant responses to those Auditory Icons were compared. The individual participant results are shown in Table G-7 in Appendix G.1. The hypotheses made prior to this study were supported as shown in the two trends for sounds with and without prior classification. The remainder of the results section for the third study concentrates on a number of outliers and possible explanations for them. These results are explained in more detail in the object and action causal uncertainty analyses in the following sections. The results of this study highlighted that object and actionhood were used as salient criteria for identification by listeners.

The distribution was not normally distributed which required the use of Kruskal Wallis and Dunn multiple comparison tests (Crawley, 2005) to determine if any of the difference shown in Figure G-7 were statistically significant. The results matched with common sense finding with more sounds presented, each individual sound is more difficult to identify. This is shown in Table 5.12. Six differences were found, as expect both of the three simultaneous Auditory Icon conditions are found to be easier than the nine Auditory Icon condition with or without constraints (Prior Classification 9 or No Prior Classification 9). These results show that Auditory Icons identification grows worse as the number presented is increased.

Conditions	*Mean Rank Difference*	*P-value*	*Conditions*	*Mean Rank Difference*	*P-value*
Prior Classification 3 vs. No Prior Classification 3	14.188	ns $P>0.05$	Prior Classification 6 vs. No Prior Classification 6	14.500	ns $P>0.05$
Prior Classification 3 vs. Prior Classification 6	29.167	ns $P>0.05$	Prior Classification 6 vs. Prior Classification 9	30.917	ns $P>0.05$
Prior Classification 3 vs. No Prior Classification 6	43.667	** $P<0.01$	**Prior Classification 6 vs. No Prior Classification 9**	48.479	*** $P<0.001$
Prior Classification 3 vs. Prior Classification 9	60.083	*** $P<0.001$			
Prior Classification 3 vs. No Prior Classification 9	77.646	*** $P<0.001$	No Prior Classification 6 vs. Prior Classification 9	16.417	ns $P>0.05$
No Prior Classification 3 vs. Prior Classification 6	14.979	ns $P>0.05$	No Prior Classification 6 vs. No Prior Classification 9	33.979	ns $P>0.05$
No Prior Classification 3 vs. No Prior Classification 6	29.479	ns $P>0.05$			
No Prior Classification 3 vs. Prior Classification 9	45.896	** $P<0.01$	Prior Classification 9 vs. No Prior Classification 9	17.563	ns $P>0.05$
No Prior Classification 3 vs. No Prior Classification 9	63.458	*** $P<0.001$			

Table 5.12: Kruskal-Wallis and Dunn multiple comparison test results for the three, the six, and the nine concurrent Auditory Icons conditions.

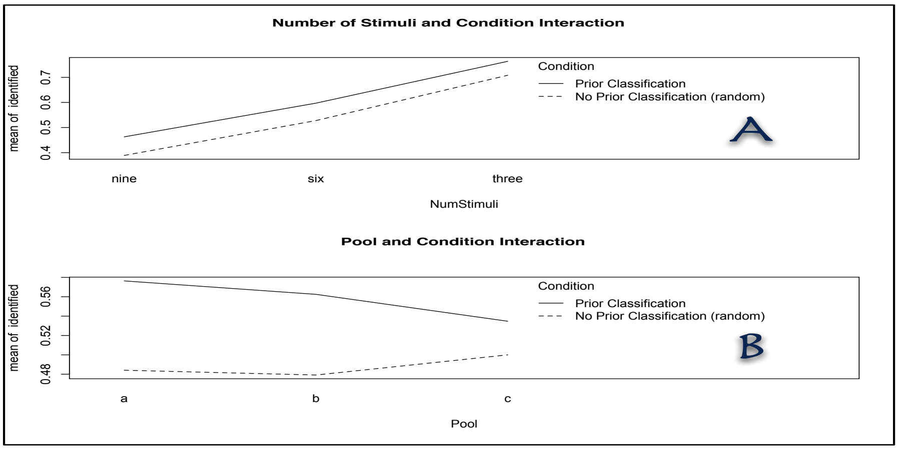


Fig. 5.22: *The interactions between the number of stimuli and presentation condition (top) and between the stimuli pool and presentation condition (bottom). Trend 'A' indicates the better identification of sounds in conditions with prior classification versus sounds in conditions without classification. Trend 'B' shows how the particular pool which a stimuli is drawn from has little effect on the identification rate.*

Causal Uncertainty Analysis of Free Text Descriptors The responses, classification, and details from each of the participants was extracted and categorized as shown in Table G-18 for Participant 1 to Table G-43 for Participant 26. Examining the results for correct identification with prior and with no prior classification shows that Auditory Icons identification grows more difficult as the number presented is increased when using prior classification however this is effect is not as severe as the difficulties encountered in the conditions where no classification of Auditory Icons was used. In order to provide greater detail about the difficulties, the next paragraph explores the analysis of the participant's textual descriptors using the causal uncertainty method. This provides details on which of the sounds were confused. The causal uncertainty results for each of the sounds used in this study is shown in Appendix G.1 in one of ten tables (Tables G-8, G-9, G-10, G-11, G-12, G-13, G-14, G-15, G-16, and G-17).

Example	Sound Name	3 Prior Action Hcu	3 Prior Object Hcu	6 Prior Action Hcu	6 Prior Object Hcu	9 Prior Action Hcu	9 Prior Object Hcu	Not Heard By	Heard By	Total Hcu
1	KnockingOnDoor.wav	0	0	-	-	-	-	0	3	0
2	Polsiren.wav	0	0	0	0	0	0	1	10	0
3	SawingWood.wav	-	-	-	-	-	-	7	0	0
Example	**Sound Name**	**3 Random Action Hcu**	**3 Random Object Hcu**	**6 Random Action Hcu**	**6 Random Object Hcu**	**9 Random Action Hcu**	**9 Random Object Hcu**	**Not Heard By**	**Heard By**	**Total Hcu**
1	KnockingOnDoor.wav	-	-	-	-	0	1	0	3	1
2	Polsiren.wav	0	0	-	-	0	0	1	10	0
3	SawingWood.wav	-	-	-	-	-	-	7	0	0

Table 5.13: Three exemplars of the different types of causal uncertainty results for the third study. A dash or minus symbol ('-') represents a condition where participants did not hear any of the sounds.

Three different types of result from the third study are shown in Table 5.13. In Example '1', the sound '*KnockingOnDoor.wav*' can be seen to be heard by three listeners in the three concurrent Auditory Icon condition with prior classification and in the nine concurrent Auditory Icon condition without classification. The causal uncertainty result of 1 for the '9 No Prior Hcu' condition would indicate that there were two distinct and different views as to what this sound was. In Example '2', the sound of '*Polsiren.wav*', we can see that there was no confusion about the identity of the sound and that 12 of 13 listeners heard the sound when presented to them. In the case of this sound, it was used in all the conditions except the '6 No Prior Hcu'. Example '3' is the sound, '*SawingWood.wav*', this sound was no heard by any of the 8 listeners who heard it. This example is one of the issues to be aware of with the causal uncertainty measure, as the result of the method will be zero even thought no listener heard the sound. This is why it is important to supplement causal uncertainty results with the number of listeners who were presented the sound and the actual number of the listeners who heard the sound. A convention is used both in Table 5.13 and in Appendix G.1 where blank spaces are left for the conditions which did not contain the particular sound. The four worst and best performing sounds are shown in Table 5.14 and Table 5.15. The sound with the highest number of different interpretations or the most confused sounds are shown in Table 5.16. An overview of the causal uncertainty results for all the conditions is shown in Figure 5.23 for actions and in Figure 5.24 for objects. Examining the best, the worst, and the most confused sounds show there is a range of results of causal uncertainty but in general the results showed more identification difficulties with hearing of the sounds rather than confusion with regard to their identification. The sounds with zero causal uncertainty are those most suitable for use, however as indicated in the earlier studies and in results of this probe. Designers should explore each particular sequence of sounds using a similar approach to ensure that a listener can clearly identify each sound in a concurrent sequence.

Sound Name	Not Heard By	Heard By	Total Hcu
43745_gelo_papas_Lighter_Ignition.wav	7	0	0
BRUSHTEE.wav	5	0	0
SawingWood.wav	7	0	0
tearing paper 02.wav	5	0	0

Table 5.14*:* The four worst identified sounds in the third study.

Example	Sound Name	3 Prior Action Hcu	3 Prior Object Hcu	6 Prior Action Hcu	6 Prior Object Hcu	9 Prior Action Hcu	9 Prior Object Hcu	Not Heard By	Heard By	Total Hcu
1	7803_hanstimm_dieselB.wav	-	-	-	-	0	0	0	8	0
2	Knocking.wav	-	-	0	0	0	0	1	8	0
3	Polsiren.wav	0	0	0	0	0	0	1	10	0
4	TELEPHON.wav	0	0	-	-	0	0	0	8	0
Example	Sound Name	3 Random Action Hcu	3 Random Object Hcu	6 Random Action Hcu	6 Random Object Hcu	9 Random Action Hcu	9 Random Object Hcu	Not Heard By	Heard By	Total Hcu
1	7803_hanstimm_dieselB.wav	-	-	0	0	-	-	0	8	0
2	Knocking.wav	-	-	0	0	-	-	1	8	0
3	Polsiren.wav	0	0	-	-	0	0	1	10	0
4	TELEPHON.wav	-	-	0	0	-	-	0	8	0

Table 5.15*:* The four best performing sounds in the third study. A dash or minus symbol ('-') represents a condition where participants did not hear any of the sounds.

Example	Sound Name	3 Prior Action Hcu	3 Prior Object Hcu	6 Prior Action Hcu	6 Prior Object Hcu	9 Prior Action Hcu	9 Prior Object Hcu	Not Heard By	Heard By	Total Hcu
1	25819_FreqMan_Splash_1_short.wav	-	-	-	-	-	-	1	0	0
2	50092_sunupi_stone_falling_water_short.wav	0	0	-	-	0	0	0	3	0
3	Cutpaper.wav	-	-	-	-	-	-	0	0	0
4	Sawing.wav	-	-	0	-	-	0	1	4	0
Example	Sound Name	3 Random Action Hcu	3 Random Object Hcu	6 Random Action Hcu	6 Random Object Hcu	9 Random Action Hcu	9 Random Object Hcu	Not Heard By	Heard By	Total Hcu
1	25819_FreqMan_Splash_1_short.wav	-	-	0	0	1	1	2	4	2
2	50092_sunupi_stone_falling_water_short.wav	1	1	-	-	-	-	2	5	1
3	Cutpaper.wav	-	-	0	0	1	1	4	3	2
4	Sawing.wav	-	0	-	1	1	0	0	4	2

Table 5.16*:* The four most confused sounds in the third study. A dash or minus symbol ('-') represents a condition where participants did not hear any of the sounds.

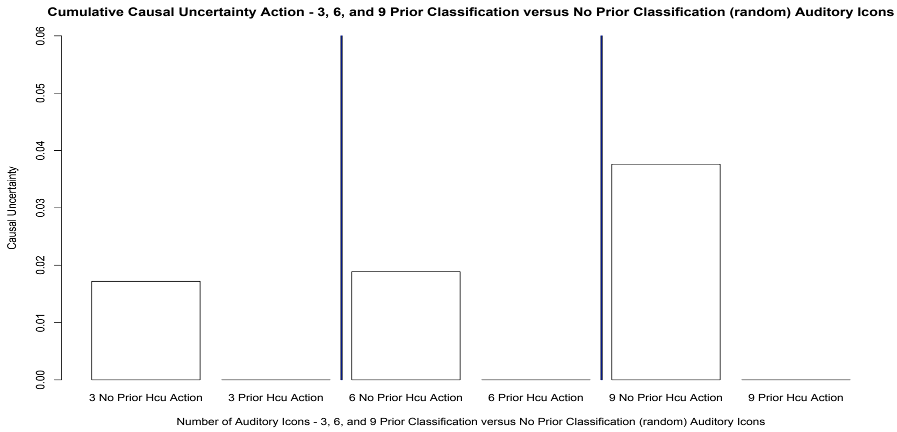


Fig. 5.23: *The average or cumulative causal uncertainty results of action categories in the third study*

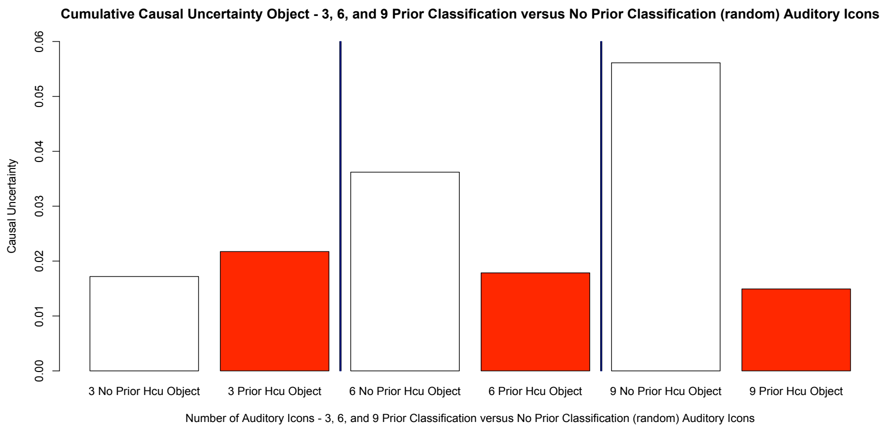


Fig. 5.24: *The average or cumulative causal uncertainty results of object categories in the third study*

The different approach to presentation and randomisation of the stimuli used in the third study provides an overview, rather than the focus of causal uncertainty results as shown in the first and second studies in this chapter. This approach can help gather a large selection of stimuli in the hundreds (this study used 189) and rapidly determine the most suitable set for use. This set can be further examined using the methods similar to those used the first and second study such as sonograms which can help identify additional issues such as masking. This two stage study where a large number of stimuli then a smaller focused study of the best performing stimuli can help identify appropriate combinations of sounds for dense concurrent sound conditions.

Post study participant debriefing As part of this study, a post study debriefing was used to collect data from the participants. Participants highlighted the particular difficulties they found in the more dense 6 and 9 Auditory Icon conditions. One participant stated that they were "*starting to question their sanity*" due to the odd combinations of sounds. This highlights the need to further consider the ecological combinations of Auditory Icons, in the case of the particular participant a number of unusual sounds were combined that would rarely if ever occur in the real world. The majority of participants felt similar to comments like "*some had a lot going on but they weren't really distinguishable, certain ones had a lot but couldn't describe them all*" or "*much harder in some, did I hear it or not, was it in the set ?*". Several felt that "*I know there are way more sounds, I'm just not getting them!*". In a similar fashion to the comments from the previous study participants did feel that upon hearing a "*certain related sound, creating a script where each feed into the other*". This again indicates how the participants find it easier to associate sounds to a particular sequence of events. In the case of conditions where unusual or non-natural sequences of sounds occurred, the difficulty of identification increased. Participants comments helped indicate in this case how they listened to the auditory scene and how their approaches for how they organised the sounds in their head. However, the comments do not give a detailed insight into this mappings and categorisation. This result was one of the rationales for carrying out the studies in the next chapter as a means for providing a more detailed understanding into the mappings and categorisations.

Results and implications from the third study for 3, 6, and 9 concurrent Auditory Icons The results from this study show a trend where increasing the number of Auditory Icons increases the difficulty in identifying them. In particular, it shows that there is a distinct performance advantage between conditions with similar numbers of Auditory Icons but which have been classified rather than those which were not classified. In particular, for 9 concurrent Auditory Icons without classification the identification rate was approximately 37%. The results for the 9 (44%), 6 (57%), and 3 (74%) concurrent Auditory Icons indicates that prior classification is effective for sound selection when designing Auditory Displays with many concurrent sounds. These results indicate that designers should use 3 concurrent Auditory Icons or less to convey their messages. Ballas's method of causal uncertainty (1986) can be used to further refine the sound selections by determining which sounds are most confused. These sounds can be removed and replaced with less confused and more perceptually distinct sounds. The results of this study show a smaller difference when compared to the identification results of the earlier two studies, however the number of participants (26) and the wider selection of stimuli from three distinct pools give this result a much firmer grounding and shows that prior classification can help the identification of Auditory Icons in concurrent presentation. The results indicate that further studies with larger numbers of sounds will not be beneficial in a practical sense. The condition with 6 concurrent Auditory Icons shows that on average listeners would not correctly identify 2 of the sounds. This is not acceptable for use in a practical Auditory Display. The third study indicates that concurrent presentation of Auditory Icons is less damaging when prior classification is used as it improves the accuracy of identification by approximately 7%. A side note from this result is that the degree of overlap between either object or action categories in conditions with no prior classification was lower in the third study and may have contributed to improvement in identification results when compared to the earlier studies. This is due to the changes in procedure with regard to how the stimuli were selected from the pools. The rest of all three studies would suggest that with a higher overlap of action and object categories, the third study would have show similar results to the first and second studies. This is an question which needs to be addressed by future research.

The results from the studies described here indicate that if Auditory Icons are carefully designed they can be used as an effective means of communicating multiple messages or bits of information using sound. The work described has demonstrated that it is possible to recognise and identify complex concurrent sounds. The results described here used minimal training and single testing sessions and would fit into a 'worst case'. The effectiveness of the listeners in sound identification would suggest that more training and longer familiarity with the sounds will improve identification results.

5.2 Applying the results in practice to three hypothetical domains

The three domains previously introduced in Section 2.5 and in the previous chapter will be again used to present a review of how this research helps those wishing to use both the results and methods from this thesis in practise to create their own Auditory Displays. This is illustrated in the three following examples.

Mobile device Auditory Display for messaging/notifications As previously stated, the idea for this theoretical application draws upon the work of Williamson et al. (2007) and the Shoogle application. The causal uncertainty method is useful for determining which sounds are confused and to what extent. In this scenario, it could be possible that several different messages or notification occur at once and multiple auditory notifications are required. The method in this chapter provides the auditory designer with a method that helps them to select the sounds that are identifiable and distinguishable to listeners for situations where multiple auditory notifications are required. Auditory Icons may have additional difficulties when used with mobile technologies outside of the laboratory as the same or similar everyday sounds may occur in the real world. This is an interesting area for future work and real world studies.

Network or processing monitoring using an Auditory Display Monitoring using Auditory Displays whether for stocks (Mauney and Walker, 2004) or networks (Gilfix and Crouch, 2000) often need to deal with large numbers of events or changes that may be occurring simultaneously. This type of information situation can often cause issues for Auditory Displays where sounds can mask or confuse one another. The method in this chapter provides one approach to help select concurrently presented sounds that are clearly distinguishable and easily identifiable to listeners.

Interactive table surface and its related Auditory Display Interactive surfaces such as the Reactable (Jordà et al., 2005) show how interfaces are becoming not only multi-touch but also multi-user. Sound is one natural way of providing feedback for these surfaces, where each action can be designed to be identifiable and distinguishable using the approach in this thesis. Concurrent sound is one possible means for the communication of concurrent multiple user interactions on such surfaces.

These three domains have shown a small sample of possibilities using this technique with regard to situations where concurrent sounds are required. The next section details the specific contributions from this chapter to Auditory Display.

5.2.1 Conclusions

The work in this chapter has shown that multiple Auditory Icons can be played in parallel while maintaining high rates of recognition. This allows for complex information in sound to be presented without slowing down the interactions with the interface. This research will allow a wider application of concurrent Auditory Icons within Auditory Displays. This chapter presented an approach for investigating the identification of Auditory Icons. It has shown that when concurrently presented without prior classification of the Auditory Icons can interact with each other in such a manner as to increase the difficulty in identifying individual Auditory Icons. The relationship between Auditory Icon identification and the proportion of those Auditory Icons that can be successfully identified was studied as well as the relationship between the identification of a Auditory Icon's actionhood and objecthood as described by the participant. It was found that by reducing the number of Auditory Icons concurrently presented, the proportion of those Auditory Icons that can be successfully identified is increased (see Figure 5.4, Figure 5.13, and Figure 5.20). It was also found that by ensuring that the objecthood and actionhood properties of Auditory Icons did not overlap in concurrent presentation that the proportion of those Auditory Icons that can be successfully identified is increased.

These studies presented an approach for determining what sounds were best identified and what properties of each of the particular sounds (its actionhood or objecthood) were better identified where many simultaneous sounds were presented. This information is very important in helping to predict the particulars sounds suitable for use in the development of metaphors and mappings for design. Further investigation into both the identification of sounds in concurrent Auditory Icon presentation and into their action/object identification is required. This work presented an approach for carrying out this type of study but given the difficulty of predicting the interactions occurring in concurrent audio presentation, it is difficult to broaden the results into principles beyond that of presenting the methods and approach used.

The results of this chapter offer the following suggestions to researchers and designers interested in using the presentation of concurrent Auditory Icons in non-spatialised environments:

Suggestion 1 The studies showed that as the number of Auditory Icons increases so does the difficult for listener identification. The six Auditory Icons with prior classification in the third study shows approximately 57% identification, more than this number of Auditory Icons and performance significantly decreases. Classification of Auditory Icons using object and action properties and the further selection of sounds that do not overlap in either of the properties can improve identification in concurrent conditions.

Suggestion 2 Prior classification can help even in concurrent conditions with three or less Auditory Icons as show in these studies, particular the third study where the identification improvement in accuracy was approximately 7% (68% versus 75%).

Suggestion 3 Ballas's method of causal uncertainty (1986) can be used to determine the confusion metric for sounds and combined with analysis of participant descriptors to select distinct and identifiable sounds for use.

Suggestion 4 Auditory Icons with the same onset time should be avoided as previous work by McGookin and Brewster (2004) suggested for Earcons. This also applies to the concurrent presentation of discrete non-continuous Auditory Icons. Ensuring at least a 300 ms onset-to-onset gap between the starts of concurrently presented Auditory Icons can improve user identification of the Auditory Icons.

The studies in this chapter widen the research in the field and help in providing a better understanding as put by Walker and Kramer (2004, p. 167), of "*how individual sounds will blend in or stand out from the growing acoustic crowd*". In particular, the studies looked at individual Auditory Icons in situations where multiple Auditory Icons were presented concurrently. Sound classification using action and object descriptors was examined as a method for selection sounds for use in concurrent auditory presentation. The exploratory results using sounds and classifying them had approximately a 7% improvement in identification accuracy as opposed to those presented without classification. However, poor identification rates occur when more than 6 (57%) classified sounds were presented concurrently. The results have shown that results are good for 3 classified sounds (75%) meaning that it is possible to have a three fold bandwidth improvement over non-concurrent / single sound Auditory Displays. These results do not take into account long term learning, which could result in further improvements in identification. Identification results were considered in relation to sound confusion as measured by Ballas's method of causal uncertainty (1986). This pointed out which of the sounds were identified but with multiple interpretations. This information can help an Auditory Display designer in choosing sounds with the clearest interpretation and identifiability. Subjective workload is another issue that should be considered when designing Auditory Displays. The NASA Task Load Index (TLX) scales (Hart and Staveland, 1988, Gawron, 2000) provides a method to rate subjective workload.

The work in this chapter supports Mynatt's (1995, p. 71) second guideline about evaluating the identifiability of auditory cues with free-form descriptors and expands upon this using Ballas's (1986) method of causal uncertainty to determine the confusion of sounds. In particular, this work combines the approaches and applies them for the first time to the domain of concurrent Auditory Icon presentation. A hypothetical example of a chat client application that supports multiple simultaneously conversations with Auditory Icons signifying different alerts could use the methods to select suitable sounds that are well identified and unambiguous in the ears of listeners.

The studies in this chapter provide a rich source of information in the form of the written descriptions from participants. These descriptors were meaningful to participants and exploited their own tacit knowledge of the world. This lead to a search for methods, which could explore the descriptors more deeply, in order to provide mappings, classifications, and metaphors. The method investigated in the next chapter, the repertory grid method (Kelly, 1955) was explored as a means of gaining an insight into listeners understanding of a set of presented everyday sounds. This understanding includes not just their immediate responses but also the listeners reasoning about the set of sounds under investigation. The next chapter uses both the repertory grid method (Kelly, 1955) and Ballas's causal uncertainty measure (1986) to see if they are better suited to "*uncovering mental models*" (Walker and Kramer, 2004, p. 169) than traditional psychoacoustic studies.

Chapter 6

Investigating People's Tacit Knowledge of Auditory Icons using Kelly's Repertory Grid Technique

> "Constructs are used for predictions of things to come, and the world keeps on rolling on and revealing these predictions to be either correct or misleading. This fact provides the basis for the revision of constructs and, eventually, of whole construct systems."
> *Principles of Personal Construct Psychology (p.14)*, George Kelly

This chapter focuses on the question of tacit knowledge and how listeners classify everyday sounds using this type of knowledge. The studies in this chapter investigated how this knowledge can be elicited and analysed. The aim of the study was to explore the classifications of everyday sounds to provide an understanding of listeners' perceptual spaces, to see how the attributes relate to one another, and to provide an approach for gathering this data to provide a vocabulary and metaphors based upon the listeners' spaces and upon the descriptions from participants. The summary in Table 6.1 discusses the key contributions to the field. The approach presented in this chapter gathers subjective information from participants but it is focused on analysing this information rather than on analysing the participants as would happen in psychology. As Auditory Displays become increasingly more complex, new methods are required to explore new facets of auditory interface design. The approach provides value to researchers specifically by adding:

- an understanding of the multidimensional structure of the listener's perceptual space using salient perceptual attributes.
- a knowledge of how attributes relate to other judgements such as quality, understanding, context or to the listener's personal preferences.

- a vocabulary based on the participant's free text descriptions and potential metaphors using their descriptions.

Informational	Principal component analysis of constructs and elements by individual participants. Multidimensional scaling results for constructs and elements by individual participants. Cluster analysis results for constructs and elements by individual participants. Confusion metric for everyday sounds as part of RGT analysis.
Inspirational	Metaphors, interpretations, and descriptors in the listener's own words of individual sounds. Detailed design spaces based on analysis of constructs and elements.
Difficulties	Implementation of statistical scripts for analysis. Interpretation of subjective individual results is complex.
Contributions	Method that allows for both informational and inspirational design relevant data. Expanded RGT method to Auditory Display domain. Combined RGT method with causal uncertainty technique. Expansion of existing sound taxonomy by CLOSED project.

Table 6.1*:* Summary of informational and inspirational aspects of the methods and techniques from this chapter and also the difficulties and contributions from this chapter.

Walker and Kramer (2004, p. 168) have pointed that it is important to know "*users at a more cognitive level*" than has previously addressed in traditional studies which they highlighted as not being suited for "*uncovering mental models*"[1]. This approach has not been explored in Auditory Display and offers an opportunity to overcome one problem of problem of traditional methods, which have "*diminished the important of learning and experience*" for interpreting sounds. The studies in this chapter concentrate on everyday sounds and their classification. In particular, this chapter focuses on the development of individual responses within an approach that aims to characterize such responses in a quantitative manner that could potentially lead to some consensus perspective without the requirement for a group discussion (Zacharov and Lorho, 2005). This approach is a based on techniques from modern perceptual evaluation where the testing of participants is used to discover their personal constructs they used to evaluate the detailed perceptual characteristics of a set of stimuli. These constructs and the derived perceptual characteristics of the sound provided by the participants' contain useful design relevant information. The quantitative and qualitative (informational or

[1]The use of the term *mental model* is taken in the context of its meaning in the domain of Auditory Display, rather than the definition of this term in cognitive psychology.

inspirational) data from this technique can be analysed at many levels to answer different kinds of questions with regard to the participant's subjective experience of the sounds. The participants provide the views using an unbiased procedure without intervention from the researcher, which allows rich subjective information to be derived and provides more reliable information due to this method. Statistical methods are used to provide projections to show the broader levels of abstraction that exist between the network of categories and ideas as expressed by the participants. These sensory re-constructions are useful in helping to break down abstract or ill-defined concepts from the participants' experiences and help create sensorial interaction metaphors. These metaphors allow the designer and the participants to create a mental image of what impressions are evoked by the sounds. This information helps the understanding of what and how participants experience the particular sounds. This is useful for determining the sonic interaction characteristics of the Auditory Display during its conceptual design phase as well as providing a means for communicating between the designer and potential users. The basis for this approach is taken from the theory of personal constructs by George Kelly (1955).

Kelly's approach, in particular the repertory grid technique, has been used in the wider acoustical domain to explore participant's subjective views of spatial attribute identification and scaling of loudspeakers (Berg and Rumsey, 1999), to explore individual's timbre space of guitar timbres (Atsushi and Martens, 2005), and to explore perceptual differences with multichannel microphone recording techniques (Martens and Sungyoung, 2007). The repertory grid when used to gather subjective experience information is not used to analyse the subject as would occur in psychology but it is focused on the constructs and elements created by the subject. Analysing the artefacts from different participants can produce design relevant information as the differences between elements and constructs assist in bringing life to a design space. The studies in this chapter are the first exploration that applies this technique within the Auditory Display domain to explore the tacit knowledge of listeners' identification and scaling of everyday sounds. The results from the investigations provide information on the categories and classifications used by listeners. This information provides a means of determining what a person thinks of when hearing a sound, in terms of the identity of the sound and its source. This chapter further explores which of the sounds are most likely to be confused by using the previously discussed causal uncertainty method. The results are analysed within the frame of the categories and conceptual framework used by the CLOSED project's sound taxonomy (Houix et al., 2007b,a). The results from this thesis suggest a number of modifications to the CLOSED project's taxonomy to accommodate a wider range of classes and sonic interactions while keeping within its overall conceptual framework.

The techniques used in this chapter generate relatively large amounts of data for analysis. Statistical techniques where used to assist in data reduction and to provide an overview of the data to facilitate trend or pattern detection. In the first study, statistical techniques examined the perceptual relationships between the auditory stimuli as rated by the participants using a repertory grid approach (Bech and Zacharov, 2006). Statistical techniques were similarly used. In the second study were the repertory grid approach was supplemented with Ballas's method of causal uncertainty (Ballas et al., 1986). In an effort to gather co-operative evidence, a mixed method approach (Greene et al., 1989) was used, with the purpose of giving complementary assessments of different aspects of the same phenomena. The reasoning behind the use of this approach was to increase the confidence of the results presented by identifying confirming or contradictory evidence and to discover findings that could not be made using a single method. Both the repertory grid approach (Bech and Zacharov, 2006) and Ballas's causal uncertainty (1986) examine similar textual descriptors. The results in this chapter show how these methods can be used to analyse the same data to provide complementary assessments. A further advantage for this complementary assessment is that it does not require participants in the study to perform any additional tasks. The next section clarifies the definition of tacit classification with regard to Auditory Icons as used in this thesis. It then follows with an introduction to personal construct theory and the methods used in this thesis to analyse the data resulting from this approach.

6.1 Tacit classification of Auditory Icons

This study is focused on eliciting the tacit knowledge of a participant. Polanyi (1966) has described how individuals develop and use tacit knowledge in a process based action focused manner where people "*know more than we can tell*" (Polanyi, 1966). A definition closer to the domain of Auditory Display has been defined by McAdam et al. (2007) as:

> "Tacit knowledge – knowledge-in-practice developed from direct experience and action; highly pragmatic and situation specific; subconsciously understood and applied; difficult to articulate; usually shared through interactive conversation and shared experience." (McAdam et al., 2007, p. 46)

One approach to exploring tacit knowledge is to use Personal Construct Theory and the Repertory Grid Technique.

6.1.1 Personal Construct Theory

George Kelly's (1955) Personal Construct Theory (PCT) was based on the view of "*man-the-scientist*". This theory holds that people construct a model of the world, act on the basis of their model, and where their model is continuously updated based on feedback from their actions in the world. Kelly's repertory grid technique (1955) was developed to enable individuals to construct a model of a particular domain of knowledge by verbalising how they perceive certain factors from within that domain. The verbalisations are *constructs*, the factors are *elements*, where the construct is a bipolar semantic dimension, where each pole represents the extreme of a particular view, in a similar fashion to Osgood's (1964) idea of bipolar semantic scales. These scales can be used with this theory and allow auditory attributes to be related to sensory dimensions, activities, or even sound sources. In this chapter, the scales used are taken from the participants' own constructs which ensures that the labels are perceptually relevant for the judgement made by the particular participant. This differs from experimenter selected scales which can impose prior knowledge or judgements rather than using the labels that the listeners find most perceptually relevant for the particular judgement. Personal constructs are described by Kelly (1955) as an individual's way of constructing certain items as similar while different from other items. The stimulus context in a given experiment provides for a restricted range of applicability, wherein the individual construct provides the basis for how the elements are understood. The repertory grid technique (Fransella et al., 2004, Bech and Zacharov, 2006), a modern descriptive analysis technique uses a scientific methodology and an associated mathematical construction of an individual's psychological space to derive the personal constructs. This technique is used in the studies in this chapter to derive participant's psychological spaces and how this is achieved is introduced briefly in the next section.

Deriving personal constructs using the repertory grid technique

The Repertory Grid Technique (RGT) (Fransella et al., 2004, Bech and Zacharov, 2006) is a method used to elicit and structure information from a participant. This technique typically consists of two phases, an elicitation phase and a rating phase. The process for this technique as related to the studies in this chapter is illustrated in Figure 6.1. The method can be used to reveal the structure of a person's classification of experiences in a manner that encourages personal reflection upon the qualities of the stimuli under examination. These stimuli, or elements as they are also known, are derived along with the definition of a personal set of constructs that differentiate between the elements, or sounds in the case of this thesis. Triadic comparison where three stimuli are presented, also known as the method of triads, was developed by Richardson (1938) and was the method used for sound presentation in the studies presented in this chapter. The studies in this chapter presented triads (triplets) of sound stimuli to subjects. Participants were instructed to describe how two of the stimuli were alike and how they differed from a third stimuli. The next triad is then presented and the same question asked. The result of this method is a set of bipolar constructs (elicited descriptors).

The constructs are created out of opposing pairs of terms, such as loud – soft or animal – mechanical. These constructs build a similarity structure for the particular listener. This structure can be analysed using statistical techniques such as multidimensional scaling, this is further discussed in Section 6.1.2. This method is more suited to the discovery of a participant's underlying *semantic constructs* than the selection of adjectives from a 'pre-set list' given by experimenters to participants. The method allows for the gathering of a participant's natural semantic responses to the sounds presented. The final stage used a rating method for each stimulus where the participant rated the stimuli along bipolar scales created from the elicited listener's descriptors. This approach does not require training or expertise with regard to the participant, nor does it bias the listeners' descriptions as the descriptions are based on their language. This approach in this chapter has an implicit acceptance that there is some form of close correspondence between the (latent) auditory sensations and their descriptions. It accepts the assumption in perceptual evaluation of everyday sound reproduction that listeners are able to analyse their complex auditory precepts in terms of separable attributes. While this assumption is not always well supported by experimental, it is in fact generally accepted.

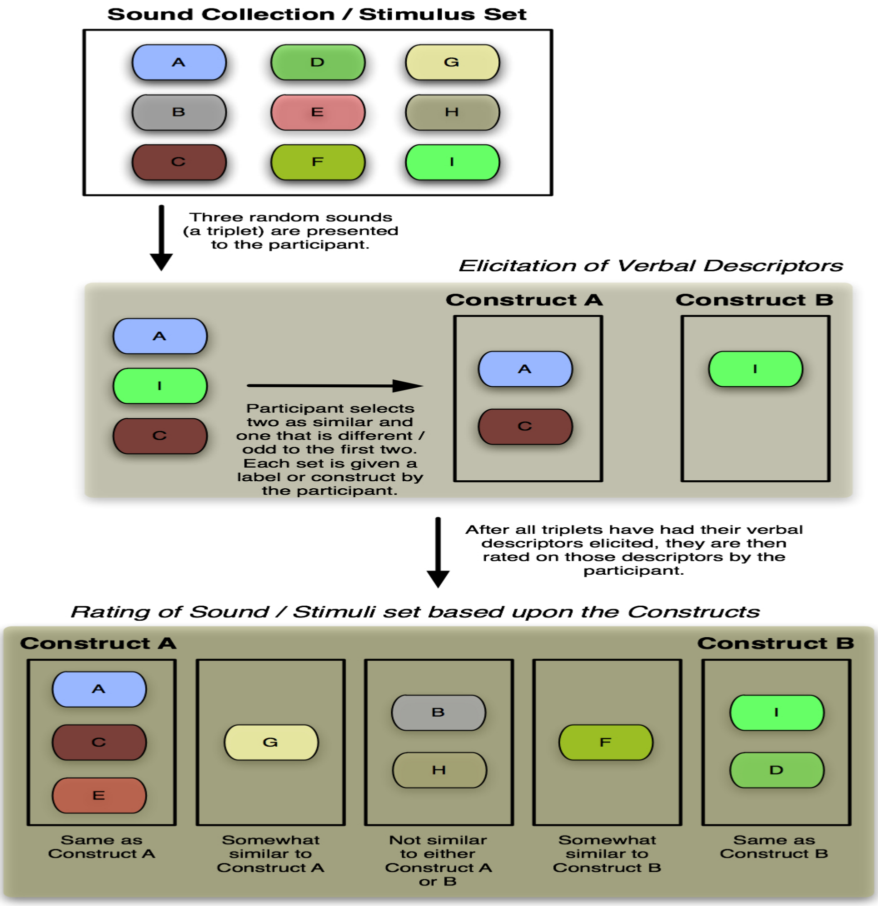


Fig. 6.1: *An illustration of the RGT technique as performed by participants during the studies in this chapter.*

6.1.2 Statistical techniques for analysis of repertory grid data

In this chapter, statistics were used to create an evaluation structure from listeners' results. This analysis shows the relationships between sounds and the major dimensions used by participants and it returned the criteria used by the participants taken from their own tacit knowledge. It further shows how the criteria were used for evaluating the sounds and the related dimensional structure of the wider sound set by the participants. The results in this study highlight three salient dimensions that were used by listeners in this study. The dimensions are listener preference, activity, and sense of daily life. The work in this thesis is similar to that of Kawai et al. (2004) who used a slightly different experimental approach and was focused in the field of noise evaluation studies. They used similar analyses to explore the psychological evaluation of environmental sounds but these were focused on group rather than the individual measures approach used in this chapter. The subjects in the study by Kawai et al. (2004) were primed to think about the sounds in either one of two locations, at home or outside a theme park. The goal of this study was to explore a new technique for determine the real life evaluation structures that people used to judge their sonic environments as part of gaining a group consensus with the aim of contributing to town planning. The work in this thesis did not prime subjects and its goal was similar but the desired goal was to gain a deeper insight and potentially a group consensus about everyday sounds with the aim of contributing to Auditory Icon research.

Data from descriptive analysis and from the RGT can be subjected to many types of group statistics such as cluster analysis methods, principal component analysis, multidimensional scaling, and a range of other measures. The grid data from an individual examined using the RGT is very rich and offers an insight into the underlying structure and content construing that formed the individual's grid responses. Using group statistics within a given population of responses for an individual enables us to determine the meaningfulness of a single grid. This approach does not require collapsing across the participants to generate a single group summary rather the work in this thesis has focused on each individual's responses to reveal their personal constructs. The associations within the participant's responses are *meaningful* in statistical terms, but it may be quite difficult to interpret their psychological meaning (Draffan, 1973). The difficulty in interpreting their meaning is due to the uniqueness of each person's experiences and this is reflected in their personal constructs as such they may require some teasing out to fully understand the meaning. This does not rule out the use of this approach, only that further exploration of the data by the researcher or by the researcher in conjunction with the particular individual may be required.

The RGT method used in the thesis combines the elicitation of the participant's descriptors and the rating of the descriptors to produce a large number of descriptors. Fransella et al. (2004) discussed several methods for analysis of the results of a repertory grid study including principal-component analysis (PCA) (Jolliffe, 2002), multidimensional scaling (MDS) (Borg and Groenen, 1996), and cluster analysis (CA) (Everitt and Hothorn, 2006). These methods are used to analyse the data from participants repertory grids. These methods as they apply to the RGT method are discussed in greater detail in Berg (2005) and in Choisel and Wickelmaier (2005). These studies used the R (Team, 2007) statistical package for data analysis. The calculations and graphical plotting in the data analysis was facilitated by the following packages MASS (Venables and Ripley, 2002), ape (Paradis et al., 2004), cluster (Maechler et al., 2005), plotrix (Lemon et al., 2007), mclust (Fraley and Raftery, 2007), and xtable (Dahl, 2007).

These methods are used in this thesis to provide for data reduction and to present visual representations of the data. These visual representations can highlight patterns, trends, or similarities between participants more easily than sets of tables. Patterns or trends are supplemented with statistical reporting to verify a hypothesis or a pattern. MDS is used to visual the pairwise distinction of participants similarity structures, created during the rating stage of the study. The MDS analysis was supplemented by minimum spanning tree (MST) analyses to identify any inaccuracies present in a MDS representation or influential observations (Jolliffe, 2002). CA is used to group the objects that are represented by participants' similarity structures. This highlights potential groupings of objects using a tree structure. PCA is used to find trends or variables of interest from the data, which are visualised using graphs. MDS and CA represent deep structures in the data using easily interpretable visual graphs. A more detailed introduction to the mechanics of the techniques with relation to how they were used to analyse the repertory grid data is given in Appendix H.1 in Sections H.1.1, H.1.2, and H.1.3 for the MDS, CA, and PCA techniques respectively.

6.2 The studies

The aim of the two studies was to understand, examine and explore in a systematic way the underlying structures of the sounds and to explore the suitability of the repertory grid for use in Auditory Icon research. The resulting information from the structures can help in formulating hypotheses about the participants and their constructs. The study was divided into two parts, the first study explored a limited subset of the stimuli and a second study that explored the full set of stimuli. The first study allowed for an observation of the procedures and the results. It helped in determining what types of sounds were most suitable and should be focused on for the second study. The second study combined the use of the causal uncertainty method as previously used in Chapter 5 as co-operative evidence sources within a mixed method approach (Greene et al., 1989). The combination of methods gives a better understanding of the data by providing different viewpoints to analyse it.

6.2.1 The second study

This research investigates the perceptual qualities of auditory events using an identification task methodology as suggested by Bonebright et al. (2005). The studies in this chapter examined the associations among auditory events in everyday sounds using a descriptive analysis approach as a complement to ongoing research (Fernström et al., 2005). Statistical techniques were used to examine the perceptual relationships between the auditory stimuli as rated by the participants using a repertory grid approach (Bech and Zacharov, 2006).

This study investigated participant's personal constructs using triad comparison and the repertory grid technique. In this particular study, participants were presented four conditions, each consisted of a single triad of sound stimuli for comparison. The listeners were asked to listen and compare three randomly selected stimuli. The three stimuli were presented sequentially. The listeners were asked to select one of the triad as being the most perceptually different from the others. There were no explicit terms or criteria provided to the listener to make this judgement but they were asked to generate a description to describe how the chosen stimuli was "odd" or different from the other two stimuli. As indicated earlier, the listener used a triadic comparison method before indicating how the stimulus was different. The listeners was asked to provide a free text description in their own terms of each of the stimuli present in the particular triad. The five subjects rated the twelve stimuli on their individually created four bipolar free text descriptor scales using a 5-point scale rating system. The focus of this study is the descriptors and the ratings produced by the subjects rather than the subjects themselves. A total of 240 ratings (5 participants x 12 stimuli x 4 presentation conditions) using a 5-point scale were collected. In some cases listeners gave multiple responses, but they were only used to clarify the intended meaning of the first description given. Listeners were instructed to use their own words and as such descriptors included descriptive, emotive, and attitudinal adjectives. As this study was interested in the terms and descriptions used by listeners, sentiment and judgement were important responses and they were included rather than excluded (Martin and Soren, 2005). The following hypotheses were made with regard to the results of this study, which examined participants personal constructs using triad comparison and the repertory grid technique. It was based on the results of other researchers applying this approach in related fields and from the last study on concurrent Auditory Icons where object and actionhood were found to be useful criteria for identification of everyday sounds and therefore likely to be used as part of a person's tacit classifications.

Hypothesis 1. *Participants use object and actionhood of sounds as salient criteria in their tacit classifications.*

Hypothesis 2. *Participants use tacit criteria to create a common attribute space which can be identified using the repertory grid technique.*

Hypothesis 3. *The repertory grid technique provides a structured process that ensures that subjects of varying ability can produce consistent attributes and descriptors.*

Participants

5 participants were recruited from the postgraduates at the University of Limerick. All participants reported normal hearing and had normal or corrected to normal vision. Written consent was obtained prior to the study from all participants. The participants had not taken part in any of the other studies.

Stimuli

18 high-quality sounds (44.1 Kilohertz 16-bit) everyday sounds (durations between 1.2 and 18.4 seconds, the majority were between 6 and 9) were used in the first study. These sounds were selected from the Freesound online sound effects collection of creative commons licensed audio (www.freesound.org, 2007). In Table 6.2, the sounds used for the first study training are shown and in Table 6.3 the sounds used in the study are shown.

ID	*Description*
p1	bee buzzing
p2	bird call
p3	rooster crowing
p4	cat meowing
p5	sheep bleating
p6	ceramics being hit or dropped

Table 6.2: The sounds used in the training phase of the first study with descriptions.

ID	*Description*	*ID*	*Description*
s1	gas stove	s7	turning paper
s2	bottling machinery	s8	rubbing and writing
s3	cutting machinery	s9	rubbing sandpaper
s4	electronic alarm clock	s10	stream, water flowing
s5	gas expelling	s11	water dripping
s6	knocking on door	s12	water pouring, bath

Table 6.3: The sounds used in the first study with descriptions.

Procedure

The participants listened to the recorded sounds (mono) in random order using headphones. They responded in free-text format to what each sound was and used the interface shown in Figure 6.2. Each participant was presented four counter-balanced conditions, the stimuli within each condition was composed of three everyday sounds to create a triad. Each triad was generated randomly per individual and no sound was presented in more than one condition. The participants listened to the sounds using the interface shown in Figure 6.2 and they recorded their descriptions on paper as shown in Figure 6.3. Paper cards were chosen for the descriptor elicitation task due to their simplicity, tangibility as well as being suitable for reuse in the later rating task. The study took between 35-45 minutes per participant.

Elicitation of descriptors The descriptors were elicited from participants in the following manner, each participant was presented triples of stimuli, and asked which of the three sounds differed the most from the other two sounds. They were then asked to describe the way the particular sound differed from the other two sounds. These descriptors where used to create the bipolar constructs, these words or phrases where later used as the poles of a rating scale. Participants were allowed to re-use existing descriptors and there was no limits on the number of times a sound could be replayed by a participant. Participants also provided free text descriptions in their own terms of each of the stimuli present in the triads. This approach seeks to *implicitly* elicit descriptors from participants. In order to prevent salient differences being found between two sounds if they were always presented together with a more dissimilar sound, each participant was presented with a randomised set of triples from the stimuli set being evaluated for the study.

Rating The participants carried out the rating process of the stimuli after all the triples had been presented to them. The aim of the rating process was to indicate the degree each construct was stimulated or excited by each stimulus (sound), and to generate numerical data for pattern matching between the constructs. This was accomplished by instructing the participants to rate each of their own personal constructs on a five-point scale for every stimulus in the rating sequence. The end points of the scale where the bipolar constructs (elicited descriptors) given by the particular participant in the previous stage of the study. Each set of bipolar constructs was used to rate the entire set of sounds excluding the three sounds, which had been used in the previous stage to generate the bipolar constructs.

Experimental Platform - Technical Details The experimental procedure used Apple's iTunes © as the presentation application. In addition, several scripts were developed using Ruby programming language on the Macintosh OS X platform. These scripts used RubyOSA[2] to provide a bridge from Ruby to the Apple Event Manager. This allows Ruby programs to automate or script Mac OS X applications in a similar fashion to using AppleScript. Using iTunes presented a familiar interface to many of the users. This familiarity reduced learning time and allowed the participants to focus more on the experimental tasks

[2]http://rubyosa.rubyforge.org/

than on learning how to use the application. A copy of the scripts and experimental results from the participants for this study can be found in accompanying DVD for this thesis in Appendix L.1. In Figure 6.2, the triad comparison task is shown on the top and the rating task is shown on the bottom within the iTunes ©interface. A picture of the actual experimental set-up with a participant performing the study is shown in Figure 6.3, with the top representing the first or description task and the bottom representing the second or rating task.

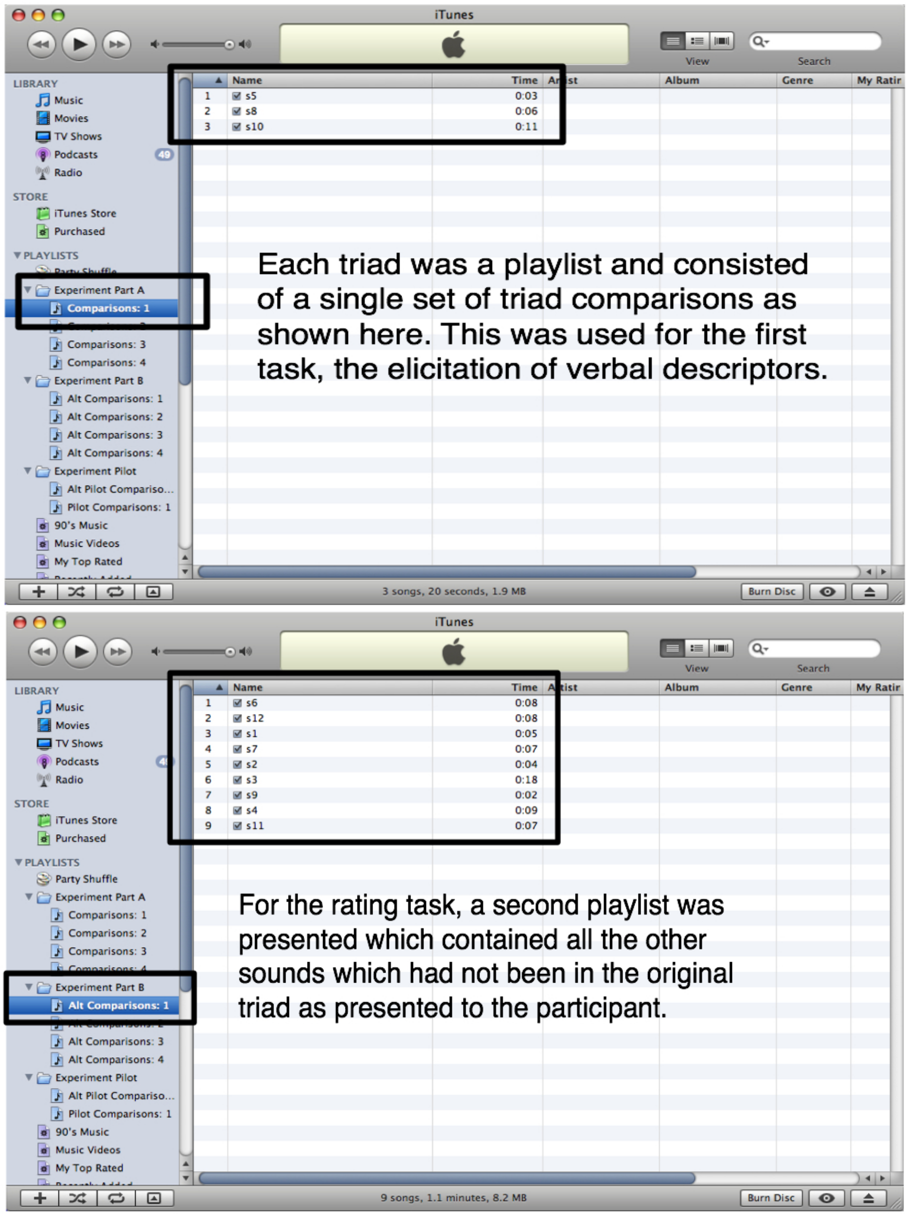


Fig. 6.2: *The interface (top) presenting a triad to the participant for the description task and the interface (bottom) presenting a set of sounds for rating with respect to an earlier triad in the verbal description task.*

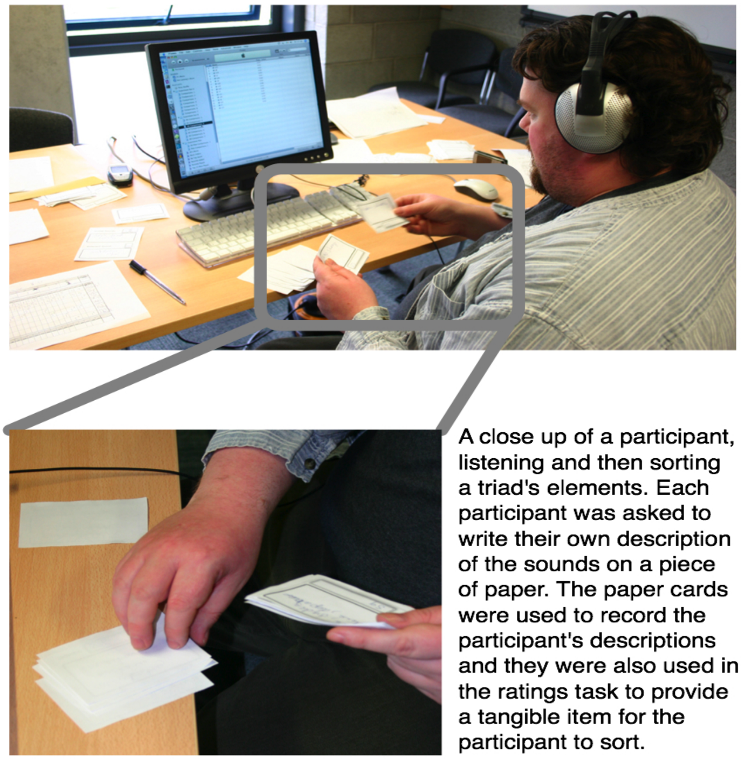


Fig. 6.3: *A picture of a participant performing the tasks during the study.*

Training

The focus of this training phase was to familiarise the participants with the presentation of the stimuli. A training set of stimuli was selected to allow participants to become familiar with the process of playback using the interface and of recording their descriptions on paper. The training interface was the same interface as used for the study and was shown in Figure 6.2. The stimuli used for the training phase were contained in the sub-playlists in the '*Experiment Pilot*' playlist and are shown in Table 6.2. These stimuli were not used in any of the later experimental conditions. The participants spent approximately 10 minutes using the interface after a short introduction on its operations. Users had headphones to listen to the sounds (in mono) while interacting with the system.

Results

The repertory grid method was used to collect the data, which was then analysed using the methods of principal component analysis, multidimensional scaling, and cluster analysis to provide the results for this study. The results presented in the next section use a less formal method of interpretation where visual analysis of patterns is favoured over the reporting of statistical matching of correspondences. This approach is more in the spirit of exploratory data analysis, which is the broader category for this type of statistical exploration. This is a useful approach as pattern matching can easily show relative similarities between participants. It offers three advantages as it can visually represent the relations, provide a map of this relation space, and offer an abstraction, which promotes a deeper understanding of the underlying knowledge. This is supplemented with statistical reporting in the relevant appendices to verify any hypotheses or patterns found.

An example of how to visually interpret the graphs of the PCA and the CA plots are shown in Figure 6.4 and in Figure 6.5 respectively. Figure 6.4 highlights which constructs were found to be similar or dissimilar by a participant. The visual analysis needs to be taken in conjunction with the more details statistical findings made for each participant and available in the appendix (in this case, Appendix H.1.3). In *A* & *B* of Figure 6.4, the construct "*water sounds – machine sounds*" (construct 2) was not associated with the construct "*mechanical/electrical sounds – paper sounds*" (construct 3, -0.64). The graphs help provide a visual overview of the detail found through statistical analysis. The CA plots in the studies can be interpreted in a similar overview fashion as shown in Figure 6.5. *A* & *B* in Figure 6.5 represent a single cluster and a cut off level where two clusters join. The choice of cut off level, represented on the y-axis, determines the number of clusters. Figure 6.5 shows four distinct clusters of elements. This clustering when considered with the component analysis can allow for trends to be seen across the sounds, such as the use of the naturalness of a sound source or the type of interaction within a sound (discrete or continuous) as tacit criteria by the participant.

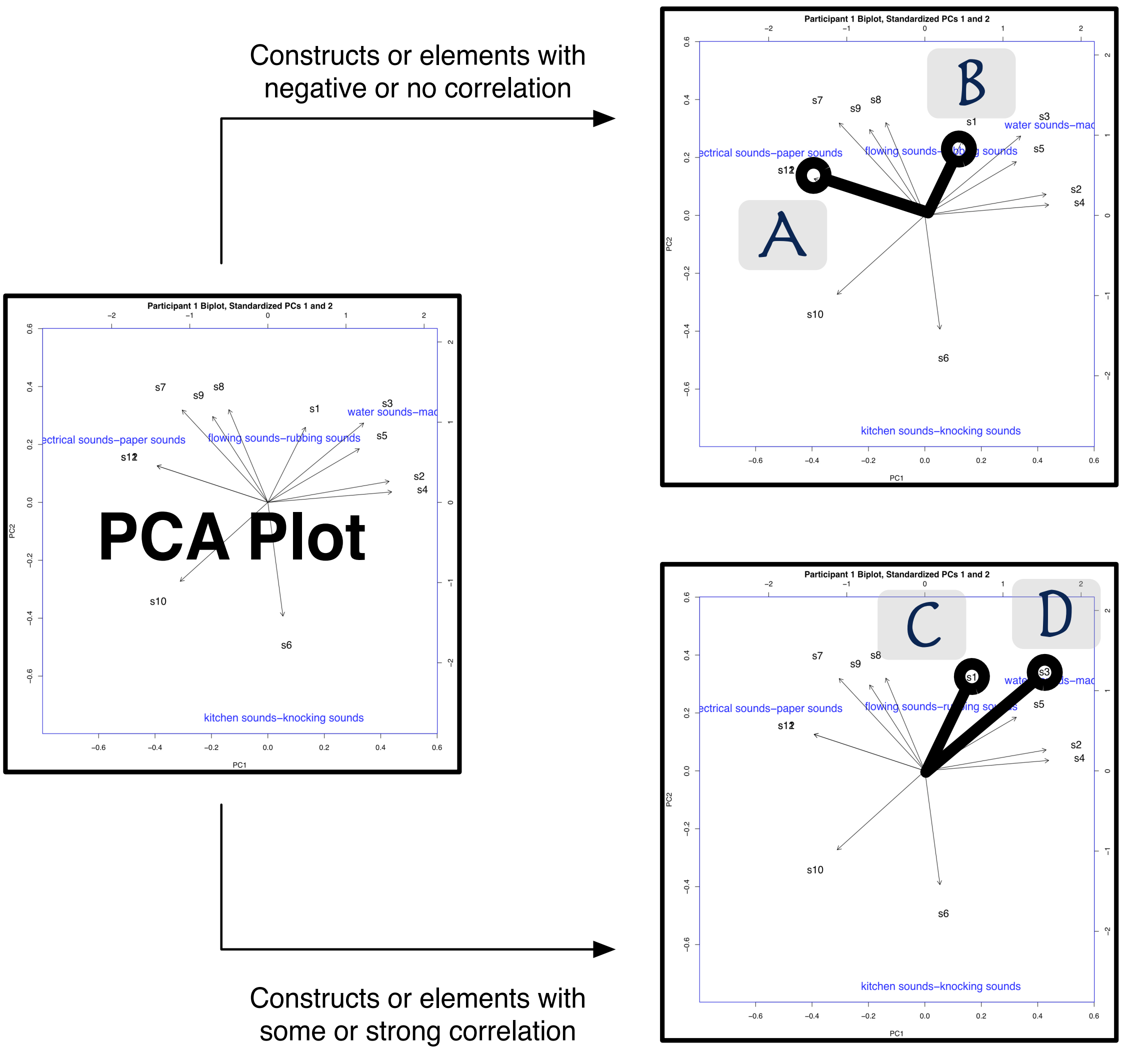


Fig. 6.4: *Interpreting the PCA plots for the studies. A & B reflect two constructs which have a negative correlation. C & D show two constructs with a positive or stronger correlation.*

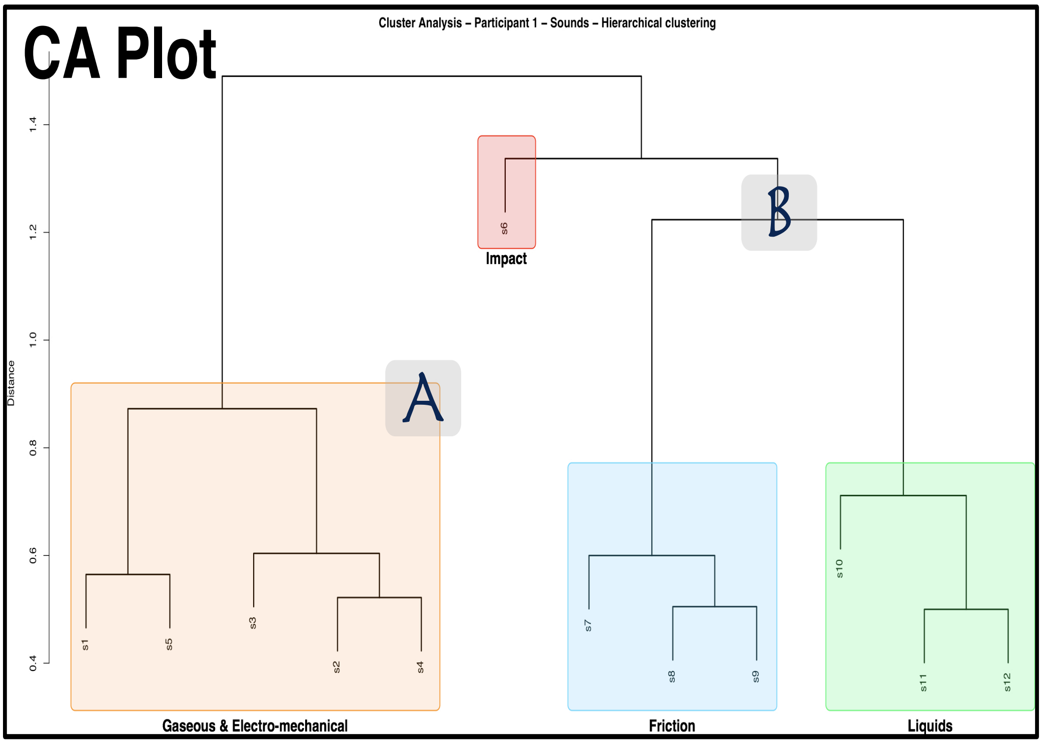


Fig. 6.5: *Interpreting the CA plots for the studies. A shows a cut off level which has highlighted a single cluster and B shows a cut off level where two clusters join.*

The first and second principal component analysis (PCA) of the participants constructs are shown in Figure 6.6 and in Table 6.4. The results from the participants show that the naturalness of a sound source, the type of interaction (discrete or continuous), a continuum between friction and impact, and the number of events occurring within a sound were all used as classification criteria by participants for constructs. Examining the analysis of the first and second principal component analysis of the participants elements as shown in Figure 6.7 and in Table 6.5 shows similar results. The results of the PCA can be supplemented by looking at the underlying data using cluster analysis (CA). Figure 6.8 and Table 6.6 show the results from a CA of the participants constructs. The results show a number of distinct and different clusters, a number of which are closely related to the suggested continua from the PCA analysis of participants constructs. There were also a number of individual clusters which may indicate other possible continua used for classification by participants. It needs to be remembered that the constructs from participants were construed individually with particular terms and knowledge relevant to each of the participants. The constructs can appear to be similar but the context or viewpoint that a Participant referred to for a particular construct may be different. Finding the common clusters and principal components between participants can help in avoiding the selection of unique or individual categories. The cluster analysis from the analysis of participants elements is shown in Figure 6.9 and in Table 6.7. The results of this clustering shows that liquid, friction, impact/deformation, gaseous, and the type of event (continuous or discrete) categories were common themes across the participants. These potential scales are further strengthened by mapping the free text descriptors onto the classification from the CLOSED project as discussed in greater detail in Section 6.2.1.

Participant	*First principal component*	*Second principal component*	*Details*
1	naturalness of sound sources	discrete interactions versus continuous interactions	Appendix H.1.3
2	discrete interactions versus continuous interactions	friction interactions versus impact type interactions	Appendix H.1.4
3	natural versus manmade	discrete interactions versus continuous interactions	Appendix H.1.5
4	real world versus imaginary nature	more events versus less events	Appendix H.1.6
5	dangerous with mechanical nature versus safer naturalistic	type of impact interactions versus view of safety	Appendix H.1.7

Table 6.4*:* The results of the principal components for the participants *constructs* in the pilot study.

Participant	*First principal component*	*Second principal component*	*Details*
1	naturalness of sound sources	friction interactions versus impact type interactions	Appendix H.1.3
2	friction interactions versus impact type interactions	natural versus mechanical	Appendix H.1.4
3	natural versus manmade	discrete interactions versus continuous interactions	Appendix H.1.5
4	real world versus imaginary nature	more events versus less events	Appendix H.1.6
5	dangerous with mechanical nature versus safer naturalistic	type of impact interactions versus view of safety	Appendix H.1.7

Table 6.5*:* The results of the principal components for the participants *elements* in the pilot study.

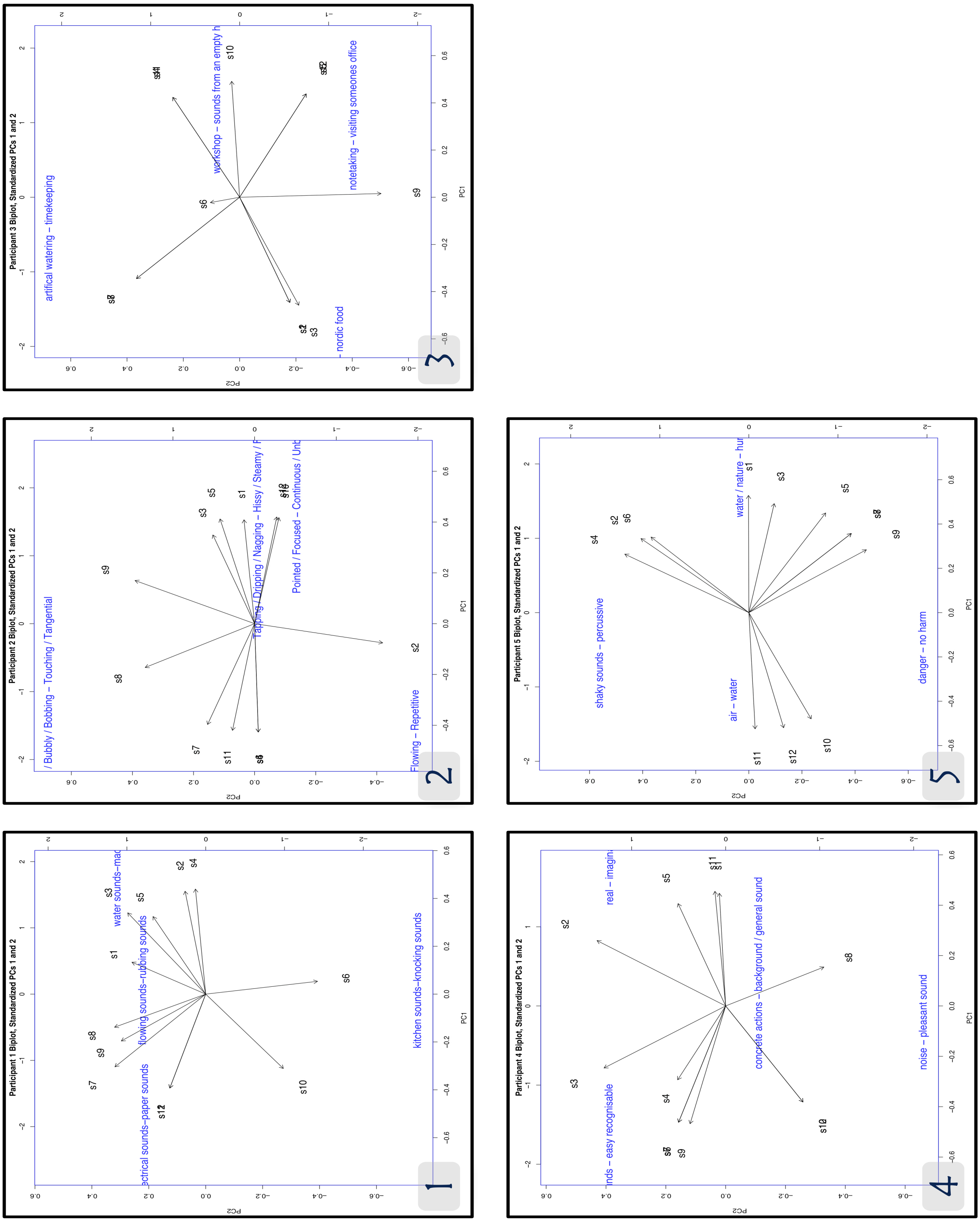


Fig. 6.6: *The participant’s constructs PCA results for the first and second principal components. From left to right, top to bottom ’1’ represents Participant one, ’2’ represents Participant two, and so on.*

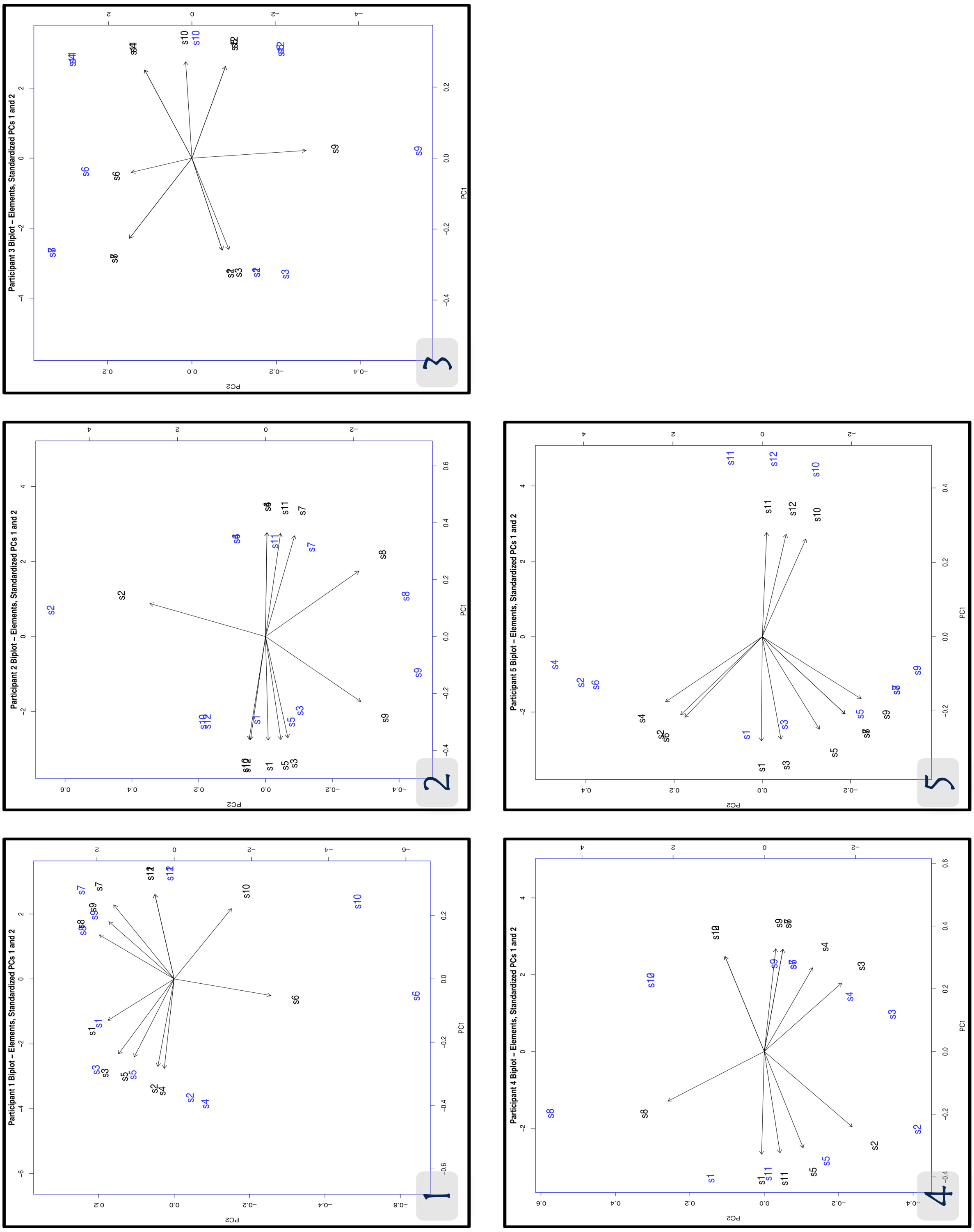


Fig. 6.7: *The participant's elements PCA results for the first and second principal components.*

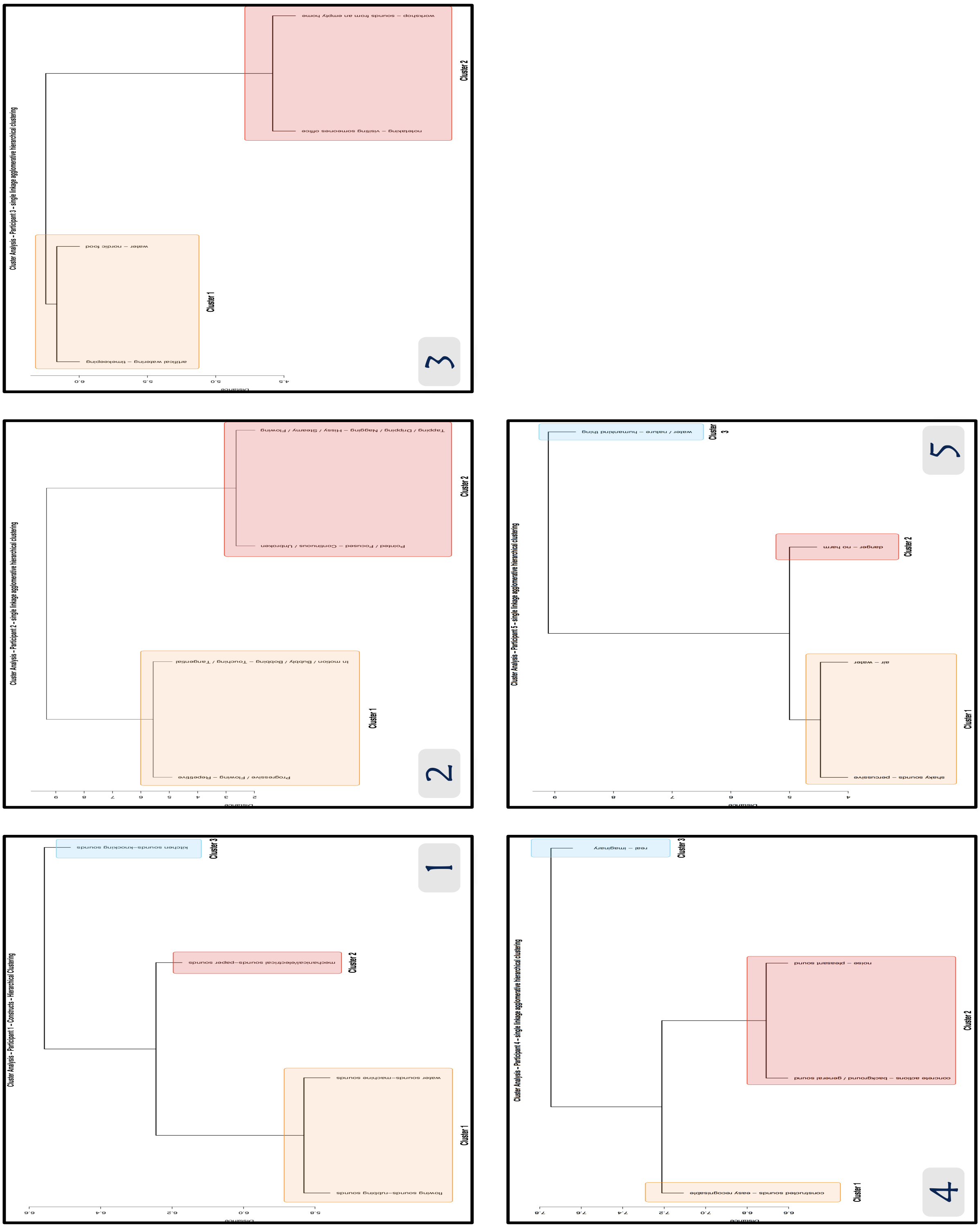


Fig. 6.8: *The participant's constructs CA results.*

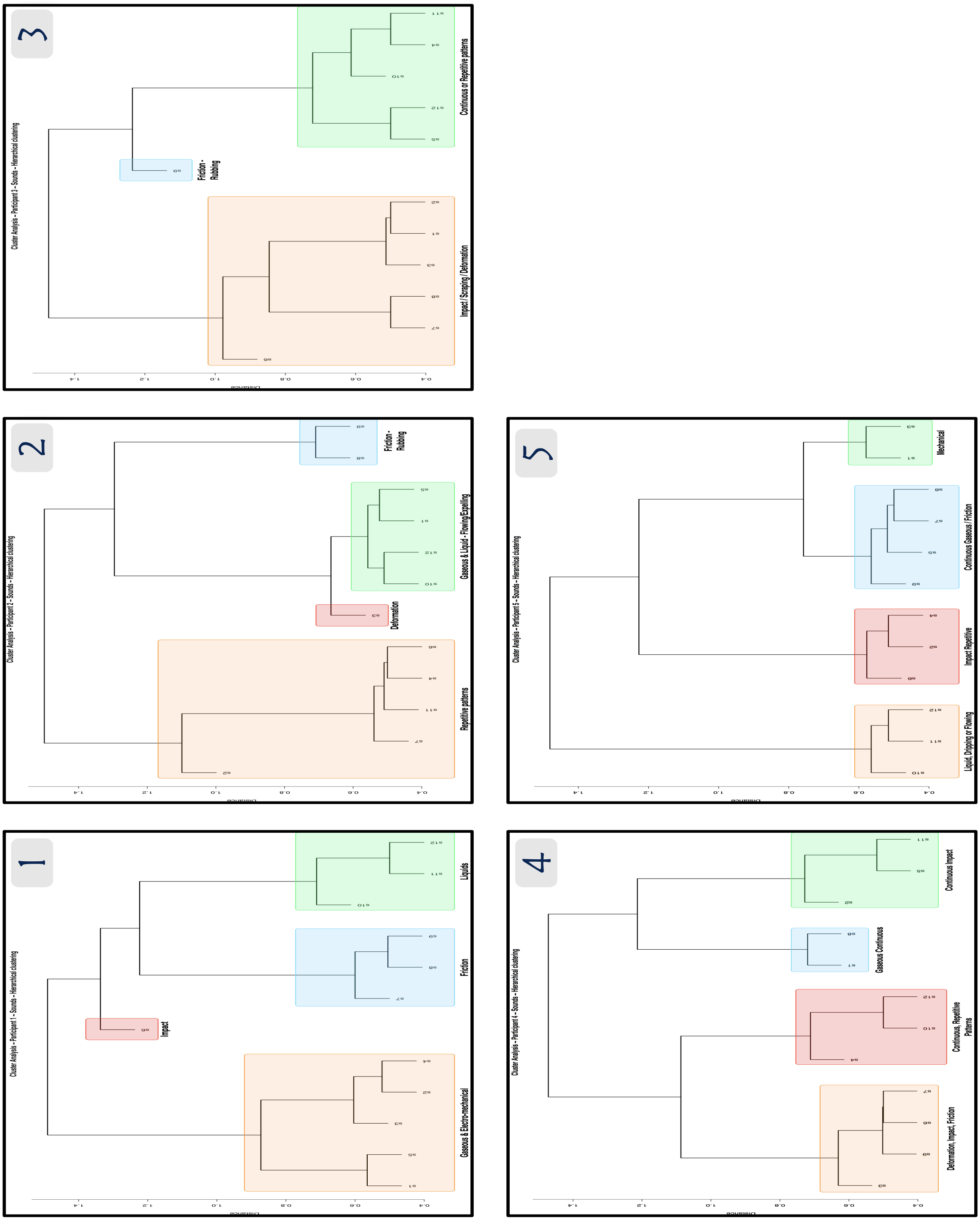


Fig. 6.9: *The participant's elements CA results.*

Participant	*Cluster 1 (Emergent Pole-Implicit Pole)*	*Cluster 2 (Emergent Pole-Implicit Pole)*	*Cluster 3 (Emergent Pole-Implicit Pole)*	*Details*
1	flowing sounds—rubbing sounds water sounds—machine sounds	mechanical/electrical sounds—paper sounds	kitchen sounds—knocking sounds	Appendix H.1.3
2	progressive/flowing—repetitive in motion/bubbly/bobbing—touching/tangential	pointed/focused—continuous/unbroken tapping/dripping/nagging—hissy/steamy/flowing		Appendix H.1.4
3	artificial watering—timekeeping water—nordic food	notetaking—visiting someones office workshop—sounds from an empty home		Appendix H.1.5
4	constructed sounds—easy recognisable	concrete actions—background / general sound noise—pleasant sound	real—imaginary	Appendix H.1.6
5	shaky sounds—percussive air—water	danger—no harm	water/nature—humankind thing	Appendix H.1.7

Table 6.6: The results of the hierarchical clustering for the participants *constructs* in the pilot study.

Participant	*Cluster 1*	*Cluster 2*	*Cluster 3*	*Cluster 4*	*Details*
1	Gaseous & Electro-mechanical s1, s5, s3, s2, s4	Impact s6	Friction s7, s8, s9	Liquids s10, s11, s12	Appendix H.1.3
2	Repetitive s2, s7, s11, s4, s6	Deformation s3	Gaseous/Liquid—Flowing/Expelling s10, s12, s1, s5	Friction—Rubbing s8, s9	Appendix H.1.4
3	Impact/Scraping/Deformation s6, s7, s8, s3, s1, s2	Friction/Rubbing s9	Continuous/Repetitive patterns s5, s12, s10, s4, s11		Appendix H.1.5
4	Deformation, Impact, Friction s3, s9, s6, s7	Continuous/Repetitive s4, s10, s12	Gaseous Continuous s1, s8	Continuous Impact s2, s5, s11	Appendix H.1.6
5	Liquid, Dripping or Flowing s10, s11, s12	Impact Repetitive s6, s2, s4	Continuous Gaseous / Friction s9, s5, s7, s8	Mechanical s1, s3	Appendix H.1.7

Table 6.7: The results of the hierarchical clustering for the participants *elements* in the pilot study.

Observations based on Participant Results

In Chapter 3, a number of theories and frameworks for the classification of everyday sounds were explored. The CLOSED project's taxonomy (Houix et al., 2007b,a) was chosen as the most appropriate theory for this thesis. It provides a well defined core and presents a general taxonomy for the classification of everyday sounds. This theory highlights the relations between sound events within a particular class, as well as highlighting the relationships between the categories, and in addition it provides a set of basic elements for sound interaction. The results from the participants in this study indicate that the CLOSED project's sound classification categories and the results from the participants could easily be mapped onto one another without any issues or difficulties. This suggests that the CLOSED project's categorisation scheme can be used to encapsulate how listeners classify everyday sounds. The descriptors from each of the participants were mapped onto this classification as shown in Figure 6.10. These results support the previous scales as potential factors used by the participants. These results and the ease of mapping to the CLOSED project's classification categories indicates the usefulness and potential for classification schemes. These categories provide a classification framework, which can help in the structuring of sound mappings. There was a strong degree of common mapping by participants with the categories A (Liquids 1), D (Electric & Electronic), E (Impact), G (Friction 2), and J (Friction 4) being used by all participants. Four of the five participants included C (Gas) as part of their mapping. The categories F (Friction 1) and H (Deformation) were used once by different participants. The individual mapping of participant results to category is shown in Figure 6.10. The results of mapping the participants descriptors in the first study were as follows:

Participant 1 Descriptors were mapped onto classes A, C, D, E, G, and J.

Participant 2 Descriptors were mapped onto classes A, C, D, E, F, G, and J.

Participant 3 Descriptors were mapped onto classes A, D, E, G, H, and J.

Participant 4 Descriptors were mapped onto classes A, C, D, E, G, I, and J.

Participant 5 Descriptors were mapped onto classes A, C, D, E, G, and J.

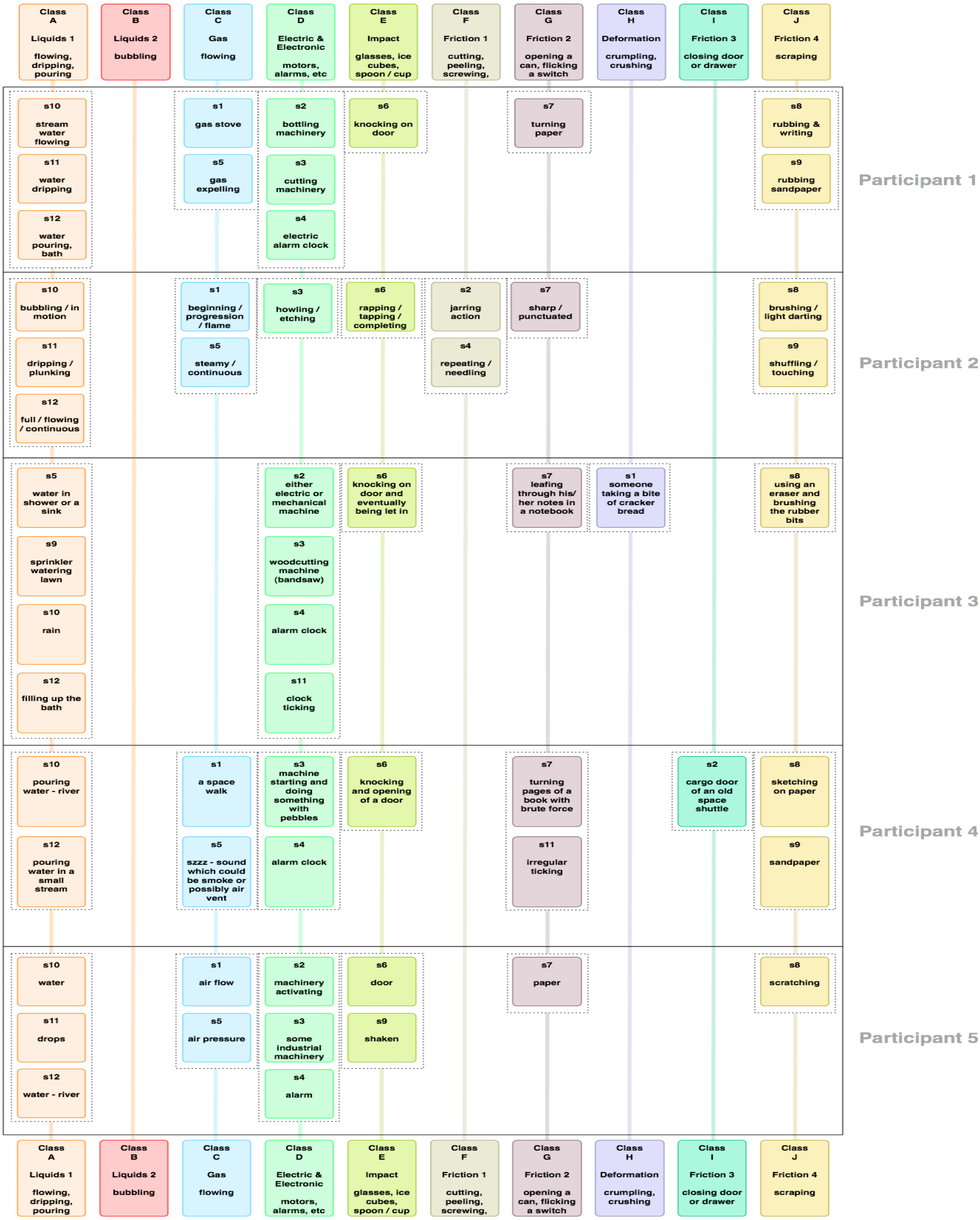


Fig. 6.10: *Projecting the element (sounds) cluster analysis from the five participant's in this study to CLOSED project's (Houix et al., 2007b,a) sound classification.*

Interpreting the statistical results of the first study

The repertory grid technique produces a large number of descriptors (up to 14 per person) and in order to provide a degree of reduction in the data, CA, MDS and PCA techniques were used. In many instances, several semantically related words or phrases will group together in the same cluster (i.e. the stimuli were rated in a similar fashion on the same scaling), this may not always be the case, and in such instances an interpreted 'label' was applied to the heterogeneous clusters. Taking an overview of the first study and looking at the results of the five participants and their constructs MDS spanning trees, the participants in this first study show a *simplex* or horseshoe like pattern (Buja and Swayne, 2002, Carreria-Perpinan, 2001). This suggests that a single curvilinear dimension can potentially give a description of the data and its classification. A review of the visual analyses from all of the participants and their respective construct and element MDS plots, shows how all the sounds were found to have this simple classification dimension as illustrated by a *simplex* pattern. The sounds were clustered on a variety of participant defined dimensions including a continuous–discrete event dimension, an activity–object/location dimension, and a real–artificial sound continuum. These are all perceptually relevant dimensions for the particular listeners and highlight that the stimuli set were a homogenous grouping, as the sounds were classifiable under a single dimension. Examining Table 6.3 shows the actual list of sounds used in the study and their descriptions. On an initial view this would seem to be a disparate selection of sounds. The results of this study highlight that even for such an apparently distinct selection of sounds; a single perceptually relevant classification dimension can exist in the mind of a listener. These results from the participant's repertory grids can help in providing a reasonable explanation of how the participants perceive the sounds within their worldview. The approach used in this study allowed a new understanding of each participant and their aspects of construing the participant stimuli. This is a new approach for Auditory Icon research and has shown a new technique for gaining a deeper insight into the views of people with regard to the stimuli or concepts being examined. The results of this research supported the complementary research by the CLOSED project on sound classification and this is examined in greater detail in the next paragraph.

Conclusions from the first study

The results from the first study and a review of the stimuli and procedures suggested a number of minor changes should be implemented. One important result was that the participant's descriptors matched categories in the CLOSED project's view of classification (Houix et al., 2007b,a) that we are situating this research within. This study showed that the method and approach were applicable to Auditory Icons. The descriptors produced by participants were similar to those analysed by Ballas's causal uncertainty measure (1986) and for the larger study it was decided to use this method in addition to the repertory grid method. This follows the earlier suggestion to use a mixed method approach (Greene et al., 1989) where complementary assessments can uncover different aspects of the same phenomena under investigation. A review of the procedure used in the first study found that the card sorting and free text description methodology successfully provided a structured process that was easily learnt by participants of all abilities.

A number of issues arose from the results of the first study. These issues prompted a number of changes to be considered for the follow-up study. The similarity of classification shown in the MDS plots of the sounds suggested a homogeneous collection of sounds. This lead to the hypothesis that a different collection of stimuli which might be less homogeneous should be examined to ensure that the results of the method was adequately differentiating between everyday sounds. This idea involved checking a different set of sounds to see if they also displayed a similar simplex pattern. This idea occurred at the time when the author was developing a prototype Auditory Display awareness system. The design of this system called for Auditory Icons to be used to distinguish different users. The idea for this Auditory Display was inspired by previous research (Huang and Mynatt, 2003, Cohen, 1994b) and aimed at providing information about the *presence* and the *availability* of co-located colleagues. This system is discussed in greater detail in Appendix I.1. The results from first study highlighted how the repertory grid results could be mapped to a common attribute space. This prompted the third reason for the change of stimuli, which was to explore sounds which did not immediately fit within the existing classification suggested by the CLOSED project (Houix et al., 2007b,a). These three reasons influenced the sound selection for the larger study, which covered a selection of everyday sounds including animal, bird, activity, and object sounds.

6.2.2 The second study, using the repertory grid and the causal uncertainty methods together

This study used two techniques; firstly it investigated participants' personal constructs using triadic comparison and the repertory grid technique. Furthermore, it examined the constructs using the causal uncertainty method. It is similar to the previous study, but presented a larger different selection of sounds but still using the triad comparison method. Listeners were asked to provide a free text description in their own terms of each of the stimuli present in the particular triad. These free text descriptors were analysed using both the repertory grid technique and the causal uncertainty method. The subjects rated the nineteen stimuli on their individually created seven bipolar free text descriptor scales using a 5-point scale rating system. The focus of this study is the descriptors and the ratings produced by the subjects rather than the subjects themselves. A total of 665 ratings (5 participants x 19 stimuli x 7 presentation conditions) using a 5-point scale were collected. In some cases listeners gave multiple responses for the descriptors, but they were only used to clarify the intended meaning of the first description given. Listeners were asked to use their own words and descriptions. The participant's descriptors did include many descriptive, emotive, and attitudinal adjectives and in a similar fashion to the last study they were included. As both sentiment and judgement are important responses and kept for analysis as suggested by Martin and Soren (2005). The following hypotheses were made with regard to the results of this study which examined participants personal constructs using the repertory grid technique and Ballas's (1986) method of causal uncertainty.

Hypothesis 1. *Participants use object and actionhood of sounds as salient criteria in their tacit classifications.*

Hypothesis 2. *Participants use tacit criteria to create a common attribute space which can be identified using the repertory grid technique.*

Hypothesis 3. *The repertory grid technique provides a structured process that ensures that subjects of varying ability can produce consistent attributes and descriptors.*

Hypothesis 4. *The causal uncertainty method could be used in to analyse the descriptors from the RGT method to determine confused or poorly identified sounds.*

Participants

There were 5 participants (4 males, 1 females) in the second study. They were either post-graduate students or employees at the University of Limerick, Computer Science Department and had not taken part in the earlier study. In pre screening for the study, all reported to having no hearing or sight problems. Written consent was obtained prior to the study from all participants. Two of the participants required glasses for reading; none of the participants reported any hearing problems.

Stimuli

The same training sounds as used in the first study were used. None of the stimuli used in the first study were included in the second study. 25 high-quality sounds (44.1 Kilohertz 16-bit) everyday sounds (durations between 1.2 and 10 seconds, the majority were between 6 and 9) were used in these investigations. These sounds were selected from the Freesound online sound effects collection of creative commons licensed audio (www.freesound.org, 2007). In Table 6.8, the sounds used for training are shown and in Table 6.9 the sounds used in the study are shown.

ID	*Description*
p1	bee buzzing
p2	bird call
p3	rooster crowing
p4	cat meowing
p5	sheep bleating
p6	ceramics being hit or dropped

Table 6.8*:* The sounds used in the training stage for the second study.

ID	*Description*	*ID*	*Description*
s1	cat	s13	glass breaking
s4	owl	s14	church bell ringing
s5	bird song 1	s18	seagull
s6	bird song 2	s19	seal
s7	bird song 3	s20	horse
s8	rooster	s22	lion roaring
s9	donkey	s23	power saw
s10	horse	s26	coins counting
s11	goat	s28	heavy ball bouncing
s12	sheep		

Table 6.9*:* The sounds used in the second study with descriptions.

Procedure

The procedure for the second study was the same as the first study. Participants listened to the recorded sounds (mono) in random order using headphones, responding in free-text format to what each sound was, using the same interface as shown previously in Figure 6.2. Each participant was presented seven counter-balanced conditions, the stimuli within each condition comprised of three everyday sounds to create a triad. Each triad was generated randomly per individual and two random sounds were presented twice otherwise each sound only occurred in a single condition. This was due to the number of stimuli used in the study being 19 and as such was not a multiple of three, hence the need for repeating sounds. The participants listened to the sounds using the interface as shown previously in Figure 6.2 and they recorded their descriptions on paper in the same manner as in the first study. Paper cards were chosen for the descriptor elicitation task due to their simplicity, tangibility as well as being suitable for reuse in the later rating task. The study took between 50-60 minutes per participant.

Elicitation of descriptors The procedure for the elicitation of descriptors was the same as the procedure used for the first study and was discussed in Section 6.2.1. The participants were presented triples of stimuli, asked which of the three stimuli differed most from the other two sounds and asked to describe this in their own terms. These terms were used to create the bipolar constructs; the constructs were used as the poles of the rating scale.

Rating The rating process of the stimuli was the same as in the first study and carried out by the participants after all the triples had been presented to the participants. The participants rated each of their own personal constructs on a five-point scale for every stimulus, were each scale used end points as described by the set of the participant's earlier bipolar constructs. The entire set of sounds was rated against each of the bipolar constructs to create the grid data for analysis.

Experimental Platform - Technical Details The same experimental platform as used in the first study was used and was shown previously in Figure 6.2. The same procedure was followed as in the first study with Figure 6.2, showing the triad comparison task on the top of the image and the rating task is shown on the bottom.

Training

The training phase was similar to the first study and familiarise participants with the presentation of the stimuli. The same training set of stimuli was used as the first study. The training allowed participants to become familiar with the process of playback using the interface and of recording their descriptions on paper. The training interface was the same interface as used for the study and was shown previously in Figure 6.2. The stimuli used for the training phase were contained in the sub-playlists in the '*Experiment Pilot*' playlist and the stimuli were not used in any of the later experimental conditions. The participants spent approximately 10 minutes using the interface after a short introduction on its operations. Users had headphones to listen to the sounds (in mono) while interacting with the system.

Results

Three methods, repertory grid analysis, causal uncertainty analysis, and post study questionnaires were used to provide the results for this study. The results use the same visual analysis approach as used in the first study. This approach is supplemented with the reporting of statistical results in the relevant appendices.

The first and second principal component analysis of the participants constructs are shown in Figure 6.11 and in Table 6.10. The results from the participants show that the naturalness of a sound source, the type of wild life (bird or animal), various continua such as familiar to unfamiliar, complete to incomplete, and relaxing to annoying were all used as classification criteria by participants for constructs. Examining the analysis of the first and second principal component analysis of the participants elements as shown in Figure 6.12 and in Table 6.5 showed that participants used criteria varying from the type of wild life (bird or normal), the distance for the sound source from the listener, to whether the sound source would typically be found indoors or outdoors. The results of the PCA can be supplemented by looking at the underlying data using cluster analysis. Figure 6.14 and Table 6.12 show the results from a CA of the participants constructs. The results show a number of distinct and different clusters, a number of which are closely related to the suggested continua from the PCA analysis of participants constructs. There were also a number of individual clusters which may indicate other possible continua used for classification by participants. It needs to be remembered that the constructs from participants were construed individually with particular terms and knowledge relevant to each of the participants. The constructs can appear to be similar but the context or viewpoint that a participant referred to for a particular construct may be different. Finding the common clusters and principal components between participants can avoid selecting unique or individual categories. The cluster analysis from the analysis of participants elements is shown in Figure 6.9 and in Table 6.13. The results of this clustering shows that objects, birds and their associate habitats, and animals were common themes across the participants.

Participant	*First principal component*	*Second principal component*	*Details*
1	naturalness of sound sources	animal sounds versus everyday sounds	Appendix J.1
2	naturalness of sound sources	animal sounds versus bird sounds	Appendix J.1.1
3	animal–alerting versus man-made–unfamiliar	familiarity versus completeness of the sound	Appendix J.1.2
4	relaxing–identifiable versus annoying–closed	naturalness of sound sources	Appendix J.1.3
5	open–welcoming versus short–mechanical	complete versus incompleteness	Appendix J.1.4

Table 6.10*:* The results of the principal components for the participants *constructs* in the larger study.

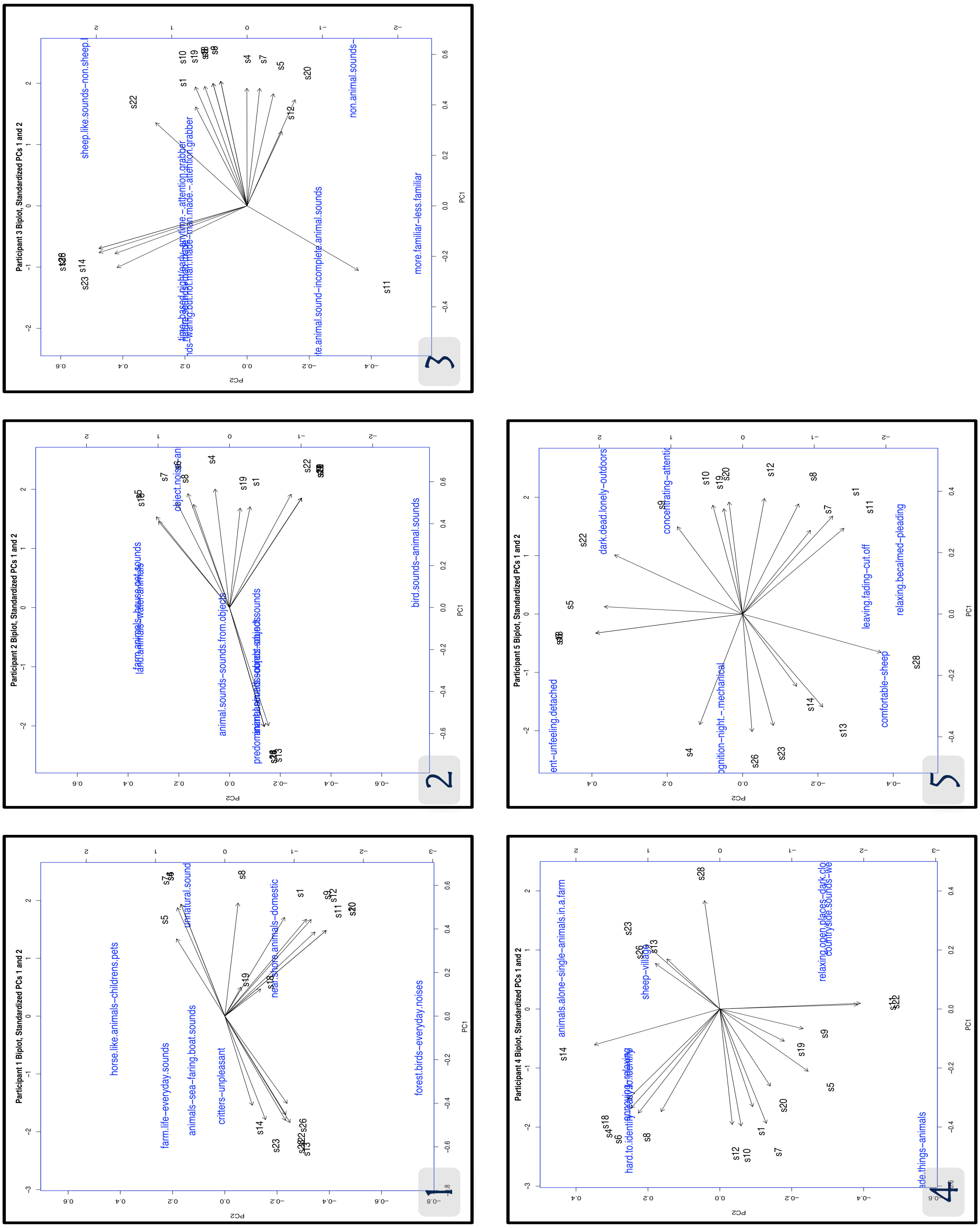


Fig. 6.11: *The participant's constructs PCA results for the first and second principal components. From left to right, top to bottom '1' represents Participant 1, '2' represents Participant 2, and so on.*

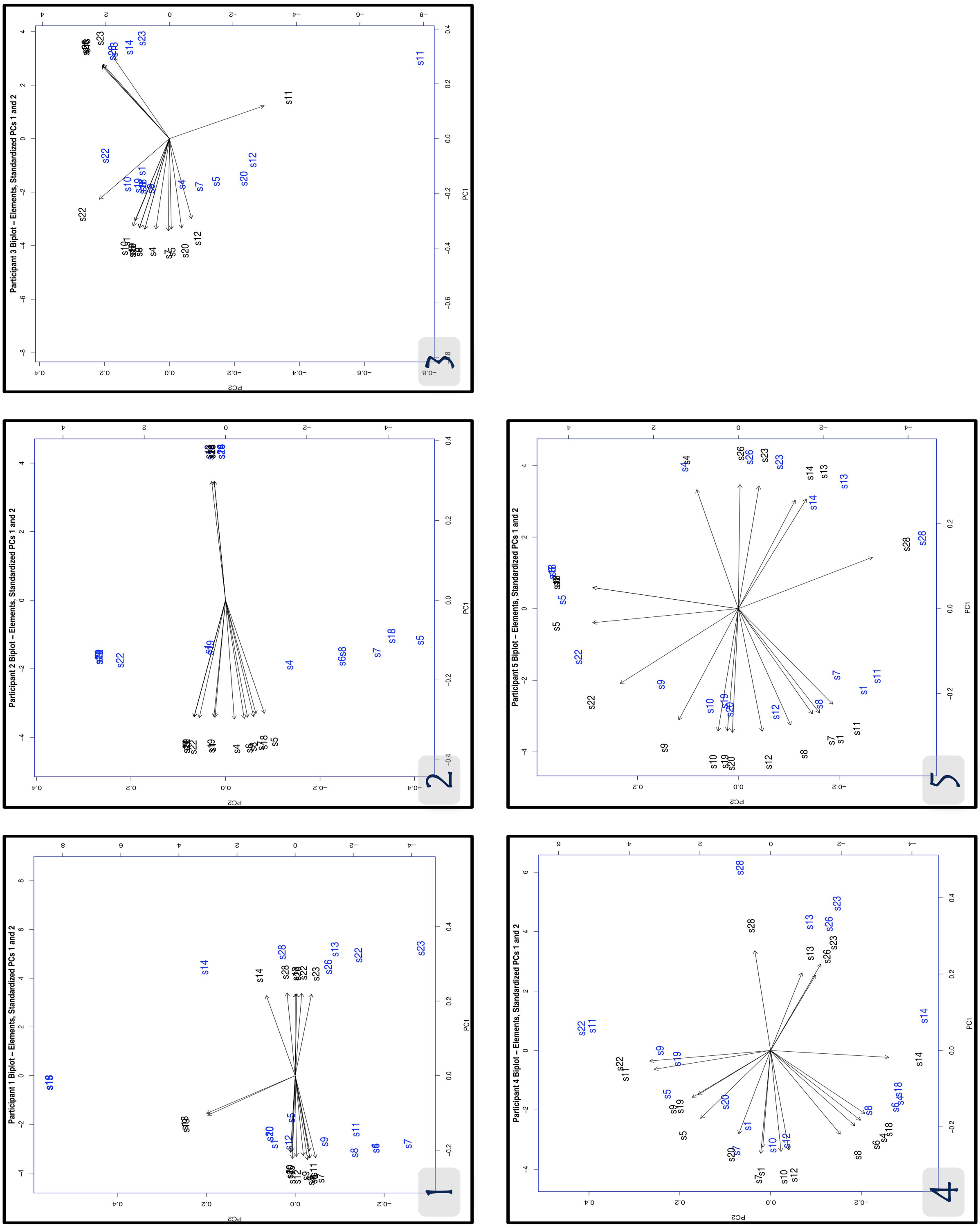


Fig. 6.12: *The participant's elements PCA results for the first and second principal components.*

Participant	*First principal component*	*Second principal component*	*Details*
1	animal–bird sources versus object sources	bird versus animal	Appendix J.1
2	animal–bird sources versus object sources	bird versus animal	Appendix J.1.1
3	animal–bird sources versus object sources	indoor versus outdoor sources	Appendix J.1.2
4	animal–bird sources versus object sources	indoor versus outdoor sources	Appendix J.1.3
5	animal–bird sources versus object sources	near sources versus distant sources	Appendix J.1.4

Table 6.11*:* The results of the principal components for the participants *elements* in the larger study.

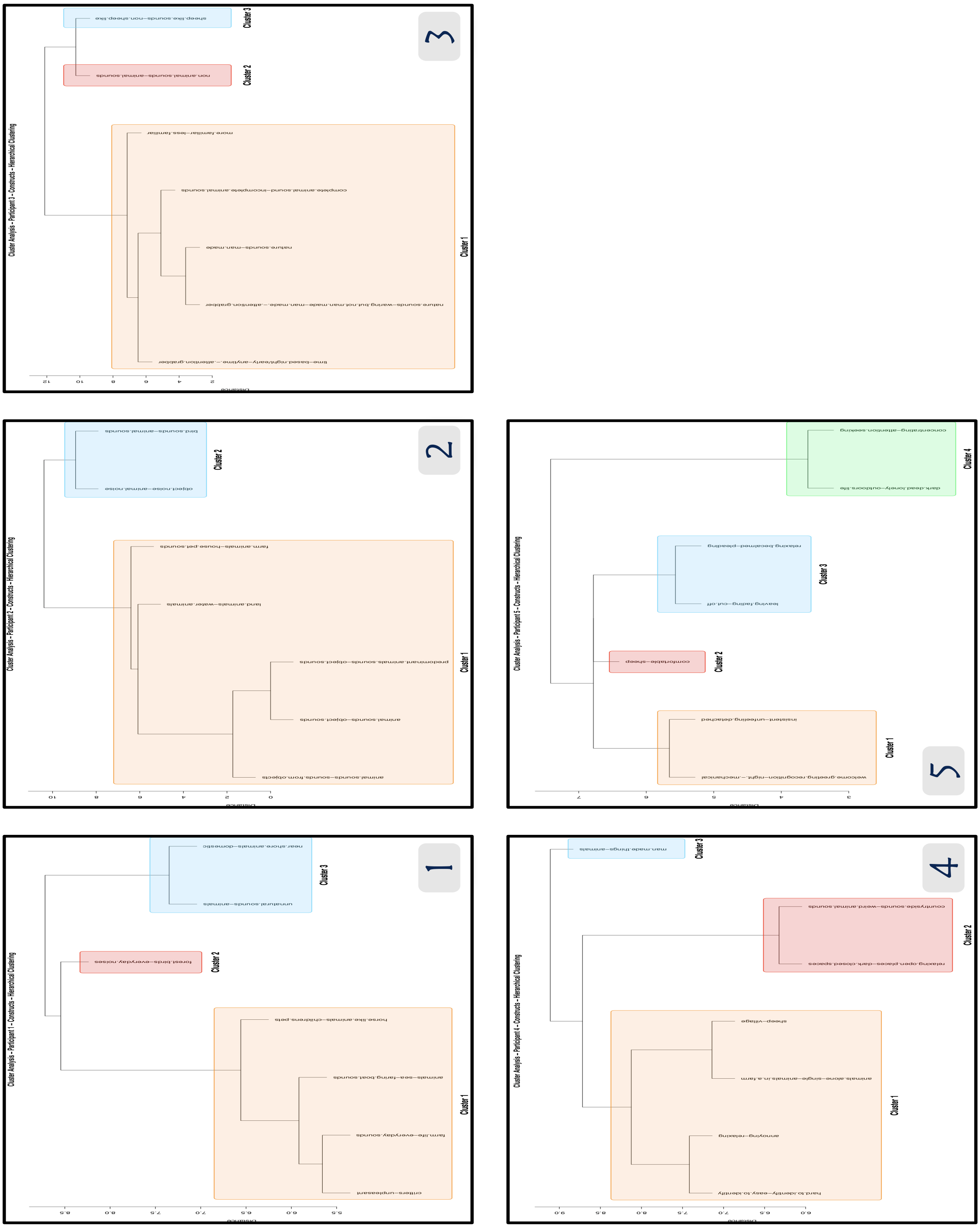


Fig. 6.13: *The participant's constructs CA results.*

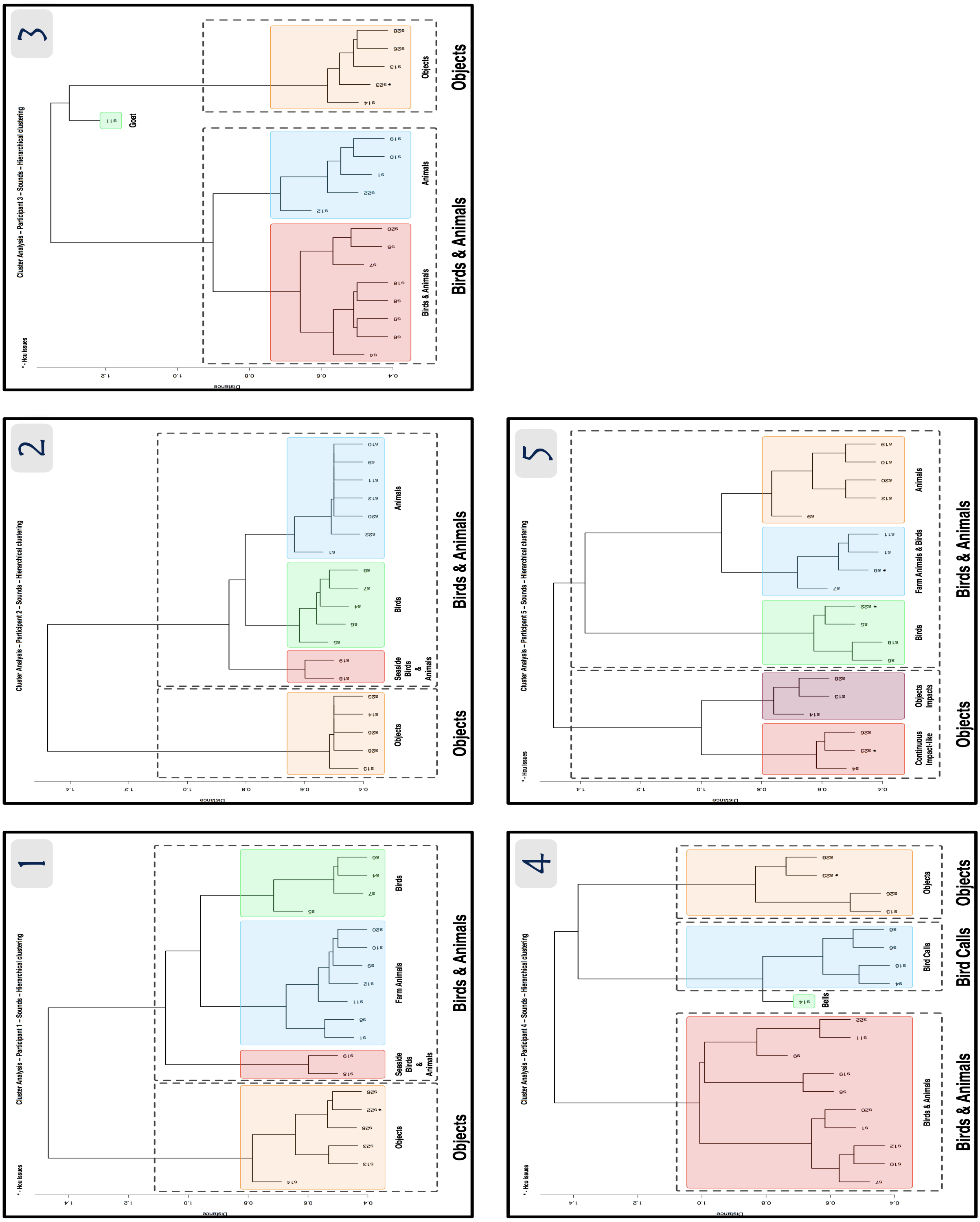


Fig. 6.14: *The participant’s elements CA results.*

Participant	*Cluster 1 (Emergent Pole-Implicit Pole)*	*Cluster 2 (Emergent Pole-Implicit Pole)*	*Cluster 3 (Emergent Pole-Implicit Pole)*	*Cluster 4 (Emergent Pole-Implicit Pole)*	*Details*
1	critters–unpleasant farm life–everyday sounds animals/sea–faring boat sounds horse like animals–childrens pets	forest birds–everyday noises	unnatural sounds–animals near shore animals–domestic		Appendix J.1
2	animal sounds— –object sounds predominant animals sounds–object sounds land animals–water animals farm animals–house pet sounds	object noise– –animal noise	bird sounds–animal sounds		Appendix J.1.1
3	time based night/early/anytime –attention grabber nature sounds/waring but not manmade –manmade/attention grabber nature sounds–man made complete animal sound–incomplete animal sounds more familiar–less familiar	non animal sounds –animal sounds	sheep like sounds –non sheep like		Appendix J.1.2
4	hard to identify –easy to identify annoying–relaxing animals alone–single animals in a farm sheep–village	relaxing open places –dark closed spaces countryside sounds– weird animal sounds	man made things–animals		Appendix J.1.3
5	welcome greeting recognition –night/mechanical insistent– unfeeling/detached	comfortable–sheep	leaving/fading–cut off relaxing/becalmed –pleading	dark/dead/lonely –outdoors life concentrating –attention seeking	Appendix J.1.4

Table 6.12*:* The results of the hierarchical clustering for the participants *constructs* in the larger study.

Participant	*Cluster 1*	*Cluster 2*	*Cluster 3*	*Cluster 4*	*Cluster 5*	*Details*
1	Objects s14, s13, s23, s28, s22*, s26	Seaside Birds & Animals s18, s19	Farm Animals s1, s8, s11, s12, s9, s10, s20	Birds s5, s7, s4, s6		Appendix J.1
2	Objects s13, s28, s26, s14, s23	Seaside Birds & Animals s18, s19	Birds s5, s6, s4, s7, s8	Animals s1, s22, s20, s12, s11, s9, s10		Appendix J.1.1
3	Birds & Animals s4, s6, s9, s8, s18, s7, s5, s20	Animals s12, s22, s1, s10, s19	Goat s11	Objects s14, s23*, s13, s26, s28		Appendix J.1.2
4	Birds & Animals s7, s10, s12, s1, s20, s5, s19, s9, s11, s22	Bells s14	Bird Calls s4, s18, s6, s8	Objects s13, s26, s23*, s28		Appendix J.1.3
5	Continuous Impact-like s4, s23*, s26	Object Impacts s14, s13, s28	Birds s6, s18, s5, s22*	Farm Animals & Birds s7, s8*, s1, s11	Animals s9, s12, s20, s10, s19	Appendix J.1.4

Table 6.13*:* The results of the hierarchical clustering for the participants *elements* in the larger study. A '*' indicates a sound that had a high confusion as measured by Ballas's causal uncertainty metric (Ballas et al., 1986).

Observations based on Participant Results

The results for this study showed that both object and action categories were used for tacit classification. It further illustrated that objects, birds and their associate habitats, and animals were common themes across the participants. These themes raise an interesting contrast for the CLOSED project's taxonomy (Houix et al., 2007b,a) as the analysis of the participants descriptors did not match with the sound classification categories. The CLOSED taxonomy was tested mainly on two types of everyday sounds, physically based sound models and consumer product sounds. The sounds in this study, shown in Table 6.9 used somewhat different kinds of everyday sounds including bird and animal sounds. This type of sound was not well suited to the existing categories including the CLOSED project's taxonomy. Investigating the existing sound classifications schemes proposed by other researchers in Chapter 3 has shown that the schemes of Gérard (2004) and of Marcell et al. (2000) explicitly included animal sounds. This meant the CLOSED project's sound classification was not suitable in its current format as a taxonomy to represent the common attribute space. However if this attribute space is extended to include the sounds used in this study, then the addition of five additional categories could provide a common attribute space. These participant derived categories were animals farmyard / pet, animals wild, animals seaside, bird songs, and bird calls as shown in Figure 6.15. The labels for the categories were created based on an analysis of the results of the themes and descriptors used by participants. The additional categories merely extend this classification scheme and helps address a problem found through this research.

This result points out that many of the existing models and classification schemes for everyday sounds need further research to capture the wide variety of the sonic world and related it back to the particular model or classification scheme. Researchers in biology have noted that interspecies communication occurs e.g, between avians and mammals (Nakagawa and Waas, 2004) or between bats and primates (Kanwal and Rauschecker, 2007). Research by Zuberbühler (2000) investigated how different non-human primate species could use the acoustic signals of a different species as labels for their underlying mental representations. This earlier research and the results in this study has shown that it is clear that animal and bird vocalisations should be considered within the context of everyday sound taxonomies. The method in this chapter has shown how to produce a set of participant categories for a set of sounds. This information is useful for extending the knowledge of Auditory Icons but can also be of use to designers when structuring their sound mappings. In a similar result to the first study, it was found that the methodology used in this study allowed participants of any ability to easily learn and use the structured process it provided.

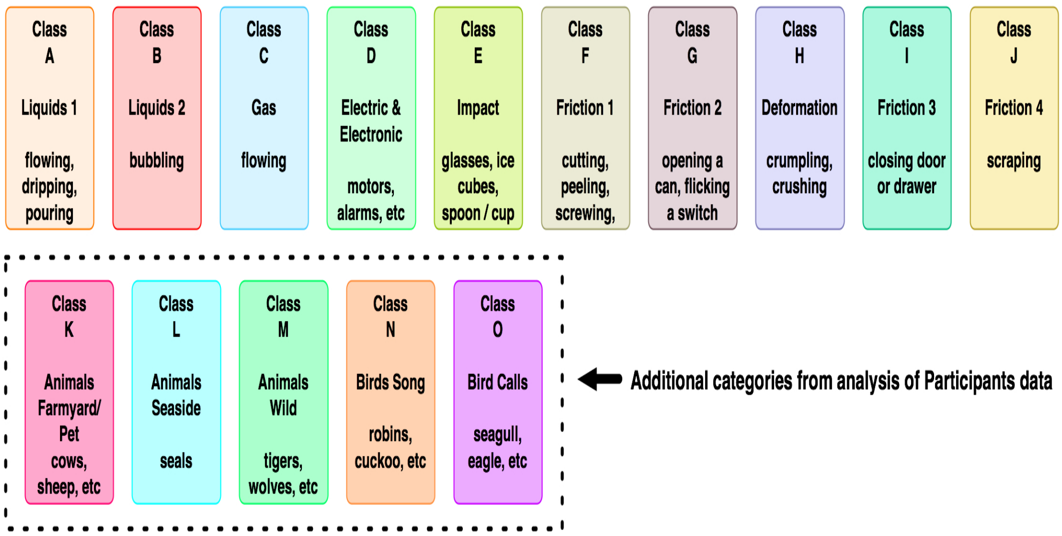


Fig. 6.15: *The CLOSED project (Houix et al., 2007b,a) sound classifications with five additional categories as suggest by this study.*

Interpreting the statistical results of the second study

The repertory grid technique produces a large number of descriptors (up to 14 per person) and in order to provide a degree of reduction in the data, CA, MDS and PCA techniques were used. In many instances, several semantically related words or phrases will group together in the same cluster (i.e. the stimuli were rated in a similar fashion on the same scaling), this may not always be the case, and in such instances an interpreted 'label' was applied to the heterogeneous clusters. Taking an overview for this study and looking at its results for the five participants and for their constructs MDS spanning trees highlights some interesting patterns. The results of participant's two and four in this study (shown in Figures J-17 and J-37) and to a lesser degree in the other participants, show a *simplex* or horseshoe like pattern (Buja and Swayne, 2002, Carreria-Perpinan, 2001). This suggest that a single curvilinear dimension as the potential description of the data. In Table 6.14 and Table 6.15, a summary of the visual analyses from the participants construct and element MDS plots are presented respectively. In the cases of participants two and four, the results would seem to support a clustering of the sounds into two major categories on a single scale from object to animal. This was further supported by the results of their constructs and element principal component analyses. The findings in Table 6.15 help to highlight the stimuli set were a diverse selection of sounds, if a participant's elements MDS plot was found to have a *simplex* pattern, it would indicate that the sounds were classifiable under a single curvilinear dimension and that the sounds were a homogenous grouping. The complexity of the visual analyses confirms the decision to select a different sound set of investigation in this study. The results highlight the sound list as shown in Table 6.9 which could be initially be interpreted as being fairly homogenous based when compared to the more diverse sound list in the first study as shown in Table 6.3. However, the studies in this chapter have shown that tacit classifications are not readily apparent given that the diverse sound list in the first study turned out to be a homogenous group while the seemingly similar set used in this study was in fact quite distinct.

The results of the participant's repertory grids and the visual representations of the results can help in providing a reasonable explanation of the way the participants see and hear the world, or in the case of this study how the heard sounds fitted into their world view. Kelly's idea was to use the grid in "*getting beyond the words*". This approach allows us to understand the person and their aspects of construing with regard to the stimuli presented. The person may be totally or partially unaware of their construing and this method helps a designer get a deeper insight into the views of the person with regard to the stimuli or concepts under examination.

Participant	*Constructs - Pattern*	*Interpretation of Pattern*
1	circumplex	Two or more dimensions required to account for constructs
2	simplex	Single dimension can account for constructs, similarity of other constructs to the construct "*animal sounds – sounds from objects*"
3	circumplex	Two or more dimensions required to account for constructs
4	simplex	Single dimension can account for constructs, similarity of other constructs to the construct "*relaxing open places – dark closed spaces*"
5	circumplex	Two or more dimensions required to account for constructs

Table 6.14*:* The patterns from visual analysis of the participants constructs MDS plots.

Participant	*Elements - Pattern*	*Interpretation of Pattern*
1	circumplex	Two or more dimensions required to account for elements
2	circumplex	Two or more dimensions required to account for elements
3	radex	Three or more dimensions required to account for elements
4	circumplex	Two or more dimensions required to account for elements
5	circumplex	Two or more dimensions required to account for elements

Table 6.15: The patterns from visual analysis of the participants elements MDS plots.

Questionnaire results from the second study

Response data from participants' was gathered at the end of their sessions using questionnaires. It explored participants' childhood living environments, current living environments, and their musical training. There were no findings of interest from the analysis of participants' response data. The results of this analysis are shown in Appendix K.1.

Causal Uncertainty as a Complementary Method to the Repertory Grid Technique

The approach used by the RGT method asked participants to provide descriptors in their own language for the everyday sounds presented to them. These descriptors were similar to those elicited from participants in Chapter 5 and this meant it was possible to further analyse them using same approach taken in Chapter 5. Combining both causal uncertainty and the repertory grid technique can provide deeper insights while only requiring a single listening test per participant. This study used a complementary design approach (Greene et al., 1989) with descriptors obtained the RGT method being further assessed using Ballas's method of causal uncertainty (1986). This method measures how many different identifications a set of listeners make for a given sound. Recapping on the method, it uses a listening test approach as used by other researchers (Ballas, 1993, Gaver, 1988, Vanderveer, 1979) and produces a measure of how a single sound may be produced by different causes. The causal uncertainty approach has an advantage as it can easily illustrate unity, degree of split, or skewed responses.

Auditory Icons Causal Uncertainty using Action Descriptors Analysis

From the participants responses, the action segments of the texts were extracted and categorised. The use of a radar plot (Saary, 2008) allows the multivariate results from the causal uncertainty analysis to be easily visualised. The action descriptors causal uncertainty analysis is shown in Figure 6.16, where only the sounds with a solid blue line on the plot were confused, all other sounds were identified correctly. Examination of this plot highlights five confused sounds, the details of which can be further examined in Table 6.16. Examining Table 6.16 which presents the results of the casual uncertainty and Table 6.9 which lists the Auditory Icons and their descriptions, we can see that of particular interest are two sounds with high causal uncertainty. The first sound is identified as sound *s22*, which was a sound of a lion roaring. The second sound is identified as sound *s23*, which was a sound of a power saw in use. The results show that both sounds were heard by all listeners but that they had different interpretations, the extent of this confusion is measured using causal uncertainty and shown in Table 6.16. This highlights the important of using methods such as causal uncertainty to ensure the interpretation of everyday sounds does not differ between listeners. Designers can use this method as part of a design process to exchange the confused sounds for semantically similar but clearly identified sounds.

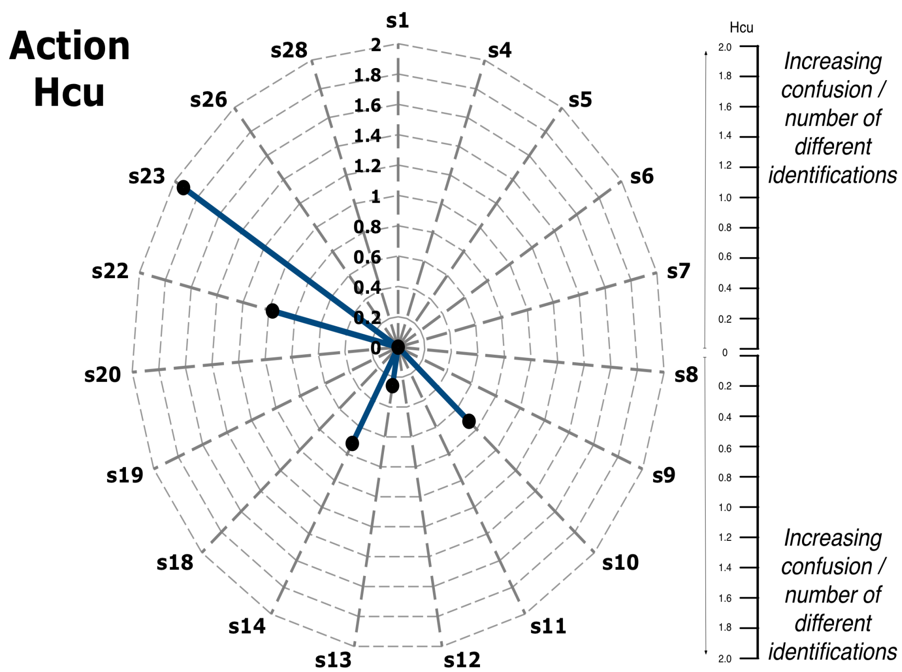


Fig. 6.16: *The Causal Uncertainty of the stimuli action-hood interpretation based on the analysis of the participants free-text responses.*

Stimuli	*Number of interpretations in*				*Hcu*
	category A	*category B*	*category C*	*category D*	
s1	5	0	0	0	0
s4	5	0	0	0	0
s5	5	0	0	0	0
s6	5	0	0	0	0
s7	5	0	0	0	0
s8	5	0	0	0	0
s9	5	0	0	0	0
s10	4	1	0	0	***0.72193***
s11	4	0	0	0	0
s12	4	0	0	0	0
s13	4	0	0	0	***0.25754***
s14	4	1	0	0	***0.72193***
s18	5	0	0	0	0
s19	5	0	0	0	0
s20	5	0	0	0	0
s22	3	2	0	0	***0.97095***
s23	2	1	1	1	***1.92193***
s26	5	0	0	0	0
s28	5	0	0	0	0

Table 6.16*:* The Causal Uncertainty of the stimuli action-hood data based on the analysis of the participants free-text responses.

Auditory Icons Causal Uncertainty using Object Descriptors Analysis

In a similar fashion as described in Section 5.1.1 in Chapter 5, the texts were extracted and categorized by object segments, such as what objects/materials were involved in the interaction. A radar plot was again used to facilitate the visualisation, with the object descriptors causal uncertainty analysis shown in Figure 6.17. As previously described only the sounds with a solid blue line on the plot were confused, all other sounds were identified correctly. Examination of this plot highlights six confused sounds, the details of which can be further examined in Table 6.17. Examining Table 6.17 and the results of the causal uncertainty and Table 6.9 for the Auditory Icons and their descriptions, four sounds are of particular interest due to their high causal uncertainty.

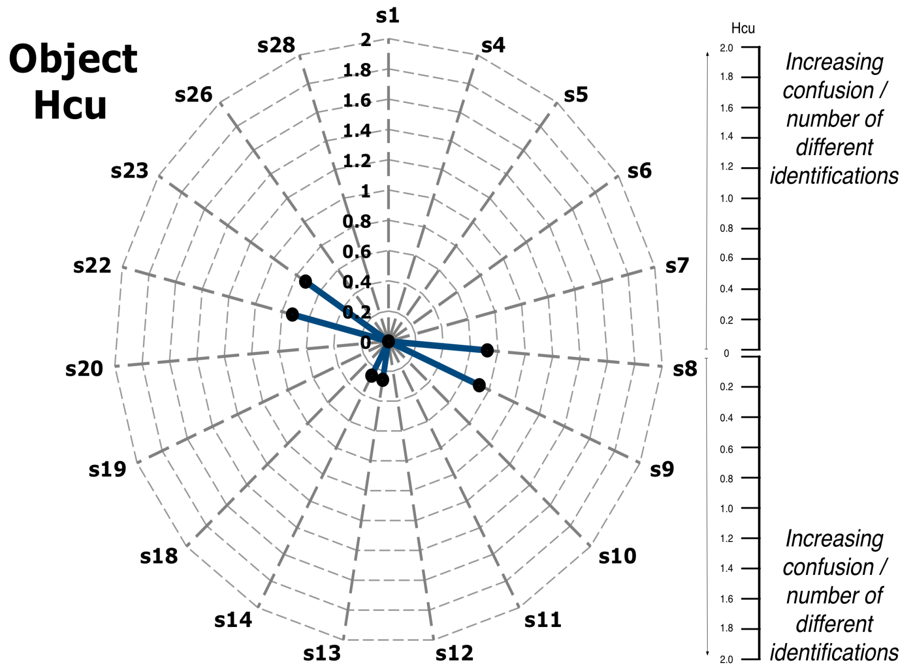


Fig. 6.17: *The causal uncertainty of the stimuli object-hood interpretation based on the analysis of the participants free-text responses.*

The first sound is identified as sound *s8*, a rooster crowing in the morning. The second sound is identified as sound *s9*, a sound of a donkey braying. The third sound is identified as sound *s22*, a sound of a lion roaring. The fourth sound is identified as sound *s23*, a sound of a power saw in use. The confusion in identification of these sounds would affect an Auditory Display design using them, if the correct interpretation of the everyday sounds in part of the metaphors or mapping used. The results of both causal uncertainty analyses would suggest, where correct interpretations are necessary that the sounds *s8*, *s9*, *s22*, *s23* be replaced by more distinct sounds with clear interpretations. The other two sounds *s13* and *s14* where not confused but rather were not described by one of the participant's. In this instance, the participant used action type descriptions without any reference to objects. This method can help determine the suitability of any new sounds by rerunning a causal uncertainty analysis with listening test descriptors on these replacement sounds.

Stimuli	*Number of interpretations in category A*	*category B*	*Hcu*
s1	5	0	0
s4	5	0	0
s5	5	0	0
s6	5	0	0
s7	5	0	0
s8	4	1	***0.72193***
s9	4	1	***0.72193***
s10	5	0	0
s11	5	0	0
s12	5	0	0
s13	4	0	***0.25754***
s14	4	0	***0.25754***
s18	5	0	0
s19	5	0	0
s20	5	0	0
s22	4	1	***0.72193***
s23	4	1	***0.72193***
s26	5	0	0
s28	5	0	0

Table 6.17*:* The causal uncertainty of the stimuli object-hood data based on the analysis of the participants free-text responses using Ballas's metric (1986).

Commentary on Causal Uncertainty analysis

The results found that two or three sounds, depending on either the action or object identification, were responsible for the majority of the identification confusion. Using this method in an iterative development cycle, allows for the opportunity to replace any sounds found to be confusing with regard to their identification, the new sounds can be subjected to similar listening tests and causal uncertainty analysis to ensure they convey the same or similar meaning while being more identifiable to the participants. The majority of the sounds (16 or so sounds of the 19 sounds) where clearly identified by the participants. These results show that the causal uncertainty method can be used in conjunction with the RGT to assist in the identification of confused or poorly identified sounds. The combination of both methods does not increase the data collection required or add any additional experimental tasks for participants while offering deeper insights by triangulation of the methods. This thesis has extended the field of Auditory Icon research by adapting the repertory grid for use within it and has additionally shown how this method can be complemented by causal uncertainty analysis.

6.2.3 Discussion

The studies in this chapter used the repertory grid technique, the causal uncertainty method, and questionnaires. This combination of methods offers insights and directions by providing multiple viewpoints from which to analyse the results. The work in this chapter has provided new information about the listeners' auditory perceptual space and the factors or concepts, which influenced the two sets of everyday sounds presented. The collection of methods and varied sounds helped to more systematically examine a wide range of everyday sounds. In the case of the second study, it points towards evaluating the results of the statistical study with the knowledge that four of the sounds *s8*, *s9*, *s22*, and *s23* had multiple interpretations by the participants. Kelly's idea was to use the grid as one method for "*getting beyond the words*" of people to gain deeper insights about the situation or objects being explored. The repertory grid approach allows us to understand the person's classifications and taxonomies used with regard to the stimuli presented. The person may be totally or partially unaware of their constructs and the method helps get a deeper insight into the views of the person with regard to the stimuli or concepts under examination.

The second study used a set of everyday sounds, which had been selected to be less homogeneous. The second set was found to display more 'circumplex' patterns conforming it was indeed less homogeneous. The sound in the second study were selected as potential Auditory Icons to distinguish different users. These Auditory Icons are similar to the hybrid versions used in the Audio Aura (Mynatt et al., 1998) and indicate the applicability of this method to the wider type of Auditory Icons. This suggests that the method could also be used for Earcons.

The repertory grid is a suitable approach for externalising the view of a person with regard to the sounds presented. It requires larger studies to determine whether such views are idiosyncratic or whether parts are shared with others. Commercial studies using this method such as the work by Zacharov and Lorho (2005), typically use five or more participants in two or more groups to provide such information. The causal uncertainty measures allow us to determine the sounds with particularly high causal uncertainties as was shown in the previous chapter and can be used to analyse the data provided from repertory grid technique without any additional experimental requirements. Any confused sounds found using this approach can then be potentially replaced by finding a similar but alternative more easily identifiable sound. The questionnaires provided useful feedback about the sounds and the listeners themselves. Future work is required to gain a better understanding of how a person and how people develop their semantic understanding of everyday sounds.

6.2.4 Limitations of the Study

There are a couple of caveats which should be noted that apply to the studies presented in the chapter. These highlight areas for consideration prior to using the methods. The issues and limitations with the causal uncertainty method were covered previously in Section 5.1.5 in Chapter 5. The issues and limitations with regard to questionnaires were covered in Section 4.4.2 in Chapter 4. As such this section will briefly highlight some of the limitations with the repertory grid technique. Training and awareness of the potential issues can ensure the correct use of the methods and approaches suggested in this chapter.

Repertory grid limitations A limitation with this method is related to how it is procedurally presented as changes to its presentation can produce substantial differences in grid outcome. Another limitation with this method is the use of bipolar constructs. The simplest error can occur where the researcher misinterprets the elicited scale, due to the fact that the user had given a particular idiosyncratic metric to the scale. This can be mitigated somewhat by exploring the mid-point of the scale with the participant to gain an insight into how they are operationalising the scale. A summary of the problematic issues with bipolar constructs is given by Yorke (1983) and included:

- the extent the meaning of one pole is itself construed as a negation of that of the other
- the ways how the gradation between the two polar extremes are constituted
- the meaning(s) inherent in the mid-point of the scale
- the relationship between the scale and implicit evaluation of goodness or badness
- the linguistic character of the bipolar opposition

In Yorke (2001) several approaches that can deal with these problematic issues are discussed. Yorke's key piece of advice is to "*go back to the original data for meaning and do not rely entirely on meta-data from statistical analyses.*" (2001, p. 182). This is a useful piece of advice for Auditory Display researchers and means that one should never blindly assume the results of the statistical analysis are true without exploring the original data.

Element selection is an important consideration when using the repertory grid technique, as such it is possible that certain types of sounds may lead to greater differentiation of grid results between participants. Previous repertory grid research on careers found that negative or 'disliked' occupations are views with a greater differentiation than positive or 'liked' occupations (Bodden and Klein, 1972). The study was later replicated by Parr and Neimeyer (1994), who verified the earlier results. This would suggest that the particular elements or sounds used may exert a significant impact on the construing of constructs and of the grid's structure. As such careful consideration should be given to the selection of elements to ensure that when using sounds they are similar in length, unambiguous, representative, and not a subset of any other sound within the proposed set of elements. The effect of one sound or element that is an outlier or 'weird' would skew the results by compacting the remaining constructs and elements. The use of visual plots can help in highlighting any skewing of the constructs or elements.

6.3 Applying the results in practise to three hypothetical domains

In a similar fashion to the previous two chapters, the three hypothetical domains will be presented here to help showcase how the methods from this chapter can help in addressing specific challenges for Auditory Displays in the specific domains.

Mobile device Auditory Display for messaging/notifications As previously stated, the idea for this theoretical scenario is similar to the idea of Williamson et al. (2007) for their Shoogle application. At the earliest stages of design, it is important to use the same vocabulary as the intended users and to understand how those users perceive the sounds with regard to one another and to their perceptual attributes. The repertory grid method (Kelly, 1955) presented is one method which can provide the designer with new insights. In the case of this scenario, this method could allow for further customisation of the Auditory Display in a similar fashion to customised ring-tones. This customisation would take advantage of the vocabulary generated to help tailor sound schemes using sounds that could be selected from wider collections using searching based on keywords or descriptions linked to the participant's vocabulary. One interesting idea could be that different but semantically similar sounds could implicitly be determined by the mobile device rather than requiring direct user interaction based on criteria such as time, location or other contextual detail that the device could establish.

Network or processing monitoring using an Auditory Display Trends and sequences in domain specific monitoring situations can be difficult to determine as tacit knowledge is used by operators in such environments. This could be a petrochemical refinery, a stock trader's station, a manufacturing plant, or an intensive care ward in hospital. This method, when used as part of a participatory design approach including domain experts can help to design a soundscape for monitoring the particular environment. A sound scheme based on the operators' tacit knowledge could be created using this approach. It would be interesting to see if there is a stronger group census from operators' when considering the sounds' vocabulary due to the specialised monitoring type Auditory Display being considered.

Interactive table surface and its related Auditory Display Imagining a Reactable-like (Jordà et al., 2005) interactive surface and an interface with has scratching, dragging, or other friction type interaction sounds from a range of different materials and objects as the hypothetical application. The method presented in this chapter allows a designer to investigate both a participant's vocabulary and how they are mapped with regard to perceptual aspects such as size, context, urgency, or any other relevant aspect. These results could be incorporated with the results of the scaling method to find the sounds that are the most expressive and best convey the desired perceptual aspect to the users for each particular mapping. For example, it might be that materiality of a sound might be better than a sound's physical size property as a mapping to reflect file size. This type of information would be easily determined when combining the results of the scaling and of the repertory grid methods. The use of the causal uncertainty method can further ensure that the sounds chosen are distinct and identifiable.

The combination of methods can help in addressing a set of issues that face many Auditory Display researchers. The next section details the specific contributions from this chapter to Auditory Display. This section has shown a number of concrete examples where the techniques can be applied and as this chapter has shown that the techniques can often be combined to provide a number of viewpoints on the sounds. These multiple viewpoints can help in triangulating to find the best Auditory Display design given the available sounds and design space.

6.3.1 Conclusions

This chapter presented the repertory grid technique for investigating a listener's tacit knowledge and their classifications of everyday sounds. In overview the contributions in this chapter are:

- applying the repertory grid technique within the domain of Auditory Display to provide an understanding of the multidimensional structure of the listener's perceptual space using salient perceptual attributes.
- combined the causal uncertainty method with the data collected by the repertory grid technique which provided multiple viewpoints on the data at no additional experimental cost.
- the participant's free text descriptions provided as part of the elements and constructs in the repertory grid can be used to create a vocabulary and metaphors.
- provided more insight into how listener's perceptual attributes related to other judgements such as context or their personal preferences.

This chapter provides an example in use of how to elicit, structure, and analyse a listener's perceptual space and shows how the listener's derived attributes for this perceptual space relate to one another. This approach is exploratory, listener-focused and helps in understanding what are the sound attributes that are salient to listeners. The results of this approach provide a vocabulary and can help generate metaphors based upon from this data. The constructs created by participants hold a meaning for them for the particular sound. They additionally created associations between the sounds. This is a new type of insight for Auditory Icon research and helps deepen the knowledge in the field.

The results of the grid data were classified using the existing taxonomy from the CLOSED project (2007a). The results presented in this chapter extend the taxonomy to include animal and bird sound aspects. This chapter also used the causal uncertainty method (Ballas et al., 1986), covered in Chapter 5, to determine the degree of a sound's identification and what properties of a sound were identified as having one or more interpretations by the participants. The multiple viewpoints offered from the different methods used provides additional insights not available through the use of either method in isolation, and without requiring any additional experimental tasks. In conclusion, this chapter presented an empirical study investigating the classification of everyday sounds, in particular how individual participants create constructs and the associations between the Auditory Icons presented to them. This provides an understanding of listeners perceptual spaces, showed how attributes related to one another, and provides an approach for gathering the data which allowed for a vocabulary and for metaphors to be created from the spaces and from the descriptions of participants.

The final chapter of this thesis provides a summary of the work undertaken, as well as its limitations. In addition, future directions in the study of everyday sounds and Auditory Icons are discussed.

Part III

The Conclusions

The third part of this dissertation provides a summary of the work undertaken, as well as its limitations. Future directions in the study of everyday sounds and Auditory Icons are discussed.

Chapter 7

Conclusions

> "Never laugh at live dragons, Bilbo you fool!"
> *The Hobbit*, JRR Tolkien

This final chapter gives a summary of the work undertaken in this thesis and highlights its contribution to the field. The chapter generalises the results of this thesis and explores a number of future directions for Auditory Icon research based upon the research presented in this thesis. From the earlier research and literature presented in this thesis, it is clear that Auditory Icons in Auditory Displays requires additional research. The combination of the methods explored in this thesis offers new insights and a greater depth of knowledge when combined. The results of the synthetic realism judgement, the scaling of the synthetic sounds, the causal uncertainty analysis, and the analysis of participant's repertory grids contribute to opening new avenues of research for Auditory Displays. The results in this thesis can help in providing a reasonable explanation of the way the participants understand sounds and the relationships between them, or in the case of the experiments in this thesis, how the heard sounds fitted into their world view. A classification framework can help in structuring sound mappings. The combination of a classification framework and of methods can ensure that researchers or designers creating new Auditory Displays can make more informed design decisions.

7.1 Research Problems

The initial motivation for this thesis was to explore three broad issues affecting Auditory Icons and everyday sounds. These issues were:

1. What kind of auditory image comes to a person's mind when they are listening to a particular everyday sound or everyday sounds presented to the listener through an Auditory Display

2. How do people identify objects or events in the associated everyday sounds when presented using an Auditory Display
3. How do people confuse certain sounds and how can this be avoided in the context of Auditory Display presentation of everyday sounds

These issues were focused into five narrower research questions:

RQ1 Does the subjective realism of a sound affect the response of a listener to the sound ?

RQ2 Do listeners subjectively hear the same physical properties of objects when both synthesised and sampled versions of the same sounds are used ?

RQ3 Do listeners subjectively use the action and object categories of multiple sounds for sound identification ?

RQ4 Is it possible to reduce the subjective confusion of the sounds using action and object categories ?

RQ5 Can a listener's tacit criteria about how they attribute meaning to everyday sounds be elicited ?

The next section addresses these questions and shows where the research contributions from this thesis help in answering the questions.

7.2 Research Contributions

This thesis presented several investigations of Auditory Icons focused on issues of concurrent presentation, of identifiability of presented sounds, of perceptual scaling of sounds, and of the associations and meanings created by listeners. A summary of the contributions and the results of its explorations are shown in Table 7.1. These explorations contribute to forming the foundations of an empirically inspired framework for the design of Auditory Displays. The framework incorporates empirical methods and approaches but only to the extent that they are accessible and practical for use by Auditory Display designers. This is similar to the approach taken by discount HCI (Nielsen, 1989) and by discount ethnography techniques, such as cultural probes (Gaver et al., 1999) or contextual inquiry (Beyer and Holtzblatt, 1997). The rationale behind these approaches is to have methods that are efficient and practical for addressing their particular goals within limited time constraints. The approach lacks the full coupling of analytical and methodological concerns of laboratory studies but this coupling is not required by designers when time is a major consideration and where this does not produce design relevant material. There is an existing body of work addressing this type of laboratory study within psychoacoustics (Zwicker and Fastl, 1990) and within ecological acoustics (Neuhoff, 2004). The work is this framework can be seen as bearing a broad resemblance to these approaches, in terms of its use of quantitative methods and analytical approaches. However, these approaches can fail in capturing design relevant material in a fashion that is accessible to designers (Tomico, 2007). This framework is a legitimate alternative approach where the primary emphasis is placed on accessibility and implications for design within Auditory Display.

Chapter 4	New 2D method for scaling and comparing sounds (real or synthetic). First exploration into perceptual scaling of parameter-based synthesis models.
Chapter 5	Method that allows for identification of appropriate combinations of concurrent sounds. Verification that overlapping categories of action or object cause identification difficulties. First exploration of concurrent everyday sounds for use in Auditory Display.
Chapter 6	Method that allows for both informational and inspirational design relevant data. Expanded RGT method to Auditory Display domain. Combined RGT method with causal uncertainty technique. Expansion of existing sound taxonomy by CLOSED project.

Table 7.1: Summary of contributions from this thesis.

This thesis presented a new method for understanding the perceptual scaling of synthesised sounds, and the impact of the realism of such sounds on their scaling and interpretation. Previous work in this area used similar approaches of multidimensional scaling (Bonebright, 2001, Lakatos et al., 2000), however a novel approach is presented in this work with regarding to this issue. This work is the first exploration into the perceptual scaling of parameter-based synthesis models. This thesis presented the first investigation into the extent that concurrently presented Auditory Icons interfere with each other, and the impact of Auditory Icon identification on such interference. This provides a new approach that can provide for the identification of appropriate combinations of concurrent sounds. This thesis evaluated the effectiveness of selecting sounds based on their action or object properties for use in concurrent presentation as a means of preventing inference between Auditory Icons. Previous studies (McGookin, 2004, Papp, 1997) have concentrated on Earcons for concurrent presentation, this thesis presents work on the related topic of concurrent presentation of Auditory Icons. McGookin (2004) found that increasing the number of Earcons decreased identification performance. The numbers of concurrent Earcons explored were between one to four. This suggests that there is more potential in using Auditory Icons for concurrent presentation. The previous studies concentrated mainly on mapping issues for Earcons, many of which are not directly applicable to Auditory Icons. Additionally, this thesis presents the first exploration of the repertory grid method as a method to elicit and analyse a listener's tacit knowledge with a focus on the individual's classifications of everyday sounds within the domain of Auditory Display. This method make use of individual responses within an approach that characterises the responses in a quantitative manner and allows for a consensus perspective without the requirement for a group discussion. This thesis showed that the repertory grid method and the causal uncertainty techniques could provide complementary data analysis. The sound taxonomy suggested by the CLOSED project was expanded by the work in this thesis to deal with animal and bird sounds. Previous studies that have used this approach or derivatives include studies on spatial audio (Berg and Rumsey, 1999), on individual timbre space construction for particular musical instruments (Atsushi and Martens, 2005), and for evaluating the perceptual differences in multichannel microphone techniques (Martens and Sungyoung, 2007).

RQ1 — Does the subjective realism of a sound affect the response of a listener to the sound ? RQ1 explored the issue of realism and if this affected the mapping a listener uses when considering sampled versus synthetic sounds. Chapter 4 investigated a new method for understanding how the synthesised sounds used were scaled by listeners in comparison with their physical dimensions and how realistic the participants found the synthetic sounds. Excluding a single outlier in the pilot stage, even while the synthesised sounds were found to be unrealistic by participants, it did not affect their ability to perceptually scale them along the 2-dimensional perceptual dimensions of the height of drop and of the size of object. These results highlighted that the perceived realism of a sound did not affect the ability of listeners to extract meaningful information from the sound, however in the case of the synthesised sounds less meaningful information was extracted. The history of previous Auditory Displays has focused on the use of real sampled (recorded) sounds to provide the realism and convey the particular mapping. The motivation for this research question was due to work by Rath (2004) and by Fernström, Brazil and Bannon (2005) on synthetic "cartoon-like" sounds were had certain perceptual parameters exaggerated. This idea is similar to the use of graphical icons as representations instead of photo-realistic images. It was the motivation for this research question and for the work in Chapter 4. The results from the exploration show that more work is need to fully understand these synthetic sounds as the current models were not yet suitable to replace real sounds in Auditory Displays. Novel approaches in parametrically controllable real time sound models from projects such as the Sound Object (Rocchesso and Fontana, 2003) offer new approaches for synthetic sounds and in particular for Auditory Displays have shown a growing need for synthetic sounds to be studied. This study extents the field of

Auditory Display by providing one of the first studies that specifically addresses synthetic parametrically modelled Auditory Icons.

RQ2 — Do listeners subjectively hear the same physical properties of objects when both synthesised and sampled versions of the same sounds are used ? Chapter 4 investigated RQ2 and provided a new method within the context of scaling of real and of synthetic sounds by listeners in a multidimensional space. This allowed for comparisons between the sounds and their physical dimensions. The results showed a large difference in scaling between real and synthetic sounds. The synthetic sounds with the best scaling results of their perceived physical properties were those that were limited to a single material, wood in particular. The synthetic sounds did convey perceptual information to a listener about particular dimensions such as height or size of an object but it was with a wider range of interpretation in results than the real sounds. This suggests that more acoustical richness (Carello et al., 2003) may be needed in the sounds to help improve listener results. These investigations showed that an Auditory Display can provide quantifiable information in multiple dimensions using synthesised sounds but that real sounds do so in a better fashion and should be favoured over synthesised sounds. By verifying the perceptual parameter-based Auditory Icons (of both synthesis models and of real sounds), it is possible to ensure the creation of Auditory Icons possessing acoustic commonalities with the objects they are designed to represent and that are able to successfully communicate quantifiable information to users. It demonstrates that it is possible to verify if Auditory Icons maintain their acoustic commonalities with the objects they are designed to represent.

Many Auditory Displays are limited to using samples or recordings of everyday sounds for their Auditory Icons. This results in large numbers of samples or recordings where the designer wishes to provide sounds that vary and can deal with various mappings. A single synthetic sound based on the type of physically modelled sound could replace these large sound libraries and be both size and computationally efficient. This was the motivation for this research as these types of synthetic sounds allow for the mapping to be represented dynamically and in a perceptually relevant manner. This thesis provided a new method that can be used to verify either real or synthetic sounds and their mappings. The advantages for this type of synthetic sound suggests that more work is required as the results in this thesis are not favourable towards low complexity synthetic models. New models with three, four, or more modes should be developed to address the issue of acoustical richness (Carello et al., 2003). The difference with earlier work by Gaver (1989) on the SonicFinder is that this approach requires less disk space by using a single synthetic model which can represent multiple sounds where the mapping can be dynamically represent a number of different ranges of values for the item being mapped such as file size.

RQ3 — Do listeners subjectively use the action and object categories of multiple sounds for sound identification ? A particular issue identified in the literature review was that varying the number of concurrently presented auditory sources affected their identification. The area of concurrently presented everyday sounds has not received as much attention as related areas such as multi-talker speech (Brungart and Simpson, 2002) or concurrent Earcons (McGookin, 2004). The research in the related areas have found a linear relationship between the number of auditory sources presented and the listener's performance for identification of the sources. The work in this thesis focused on concurrently presented Auditory Icons and provides new detail in this area. A problem for this linear relationship was that the magnitude of it differed when explored by the various researchers with regard to the number of sources and the listener's identification rates. RQ3 explored this issue for concurrently presented everyday sounds and sought to determine if such a relationship existed for the sounds and how difficult it was to identify concurrently presented everyday sounds.

Chapter 5 investigated this question using an exploration of everyday sounds where between three to ten everyday sounds were presented concurrently, and participants were asked to identify the sounds. These investigations were split across two sub-studies and in addition, the everyday sounds were comprised of two distinct sets. The first had no prior classification of the sounds such that sounds with the same objects, actions, or agents could be present in a particular condition. The second set used classification to ensure that no overlap of objects, actions, or agents occurred in any of sounds in any of the conditions presented.

The studies show that there is a distinct performance advantage between conditions with similar numbers of Auditory Icons but which have been classified rather than those which were not classified. The first study highlighted the ability of people to easily identify (89%) and distinguish six concurrent sounds with prior classification. The second study also showed a major difference between those sounds which had prior classification and those that had none. Both studies were underpowered and a third larger study was run to confirm the earlier findings of the studies. The third study found that with nine simultaneous sounds showed identification rates of 44% versus 37% for the prior classified sounds versus those that were not. The results for the 9 (44%), 6 (57%), and 3 (74%) concurrent Auditory Icons in the third study indicate that prior classification is effective for sound selection when designing Auditory Displays with many concurrent sounds. In the third study, the condition with 6 concurrent Auditory Icons shows that listeners would 40% of the sounds in that condition, which is not really acceptable for use in a practical Auditory Display so it is suggested that Auditory Display designers use 3 concurrent Auditory Icons. The third study also indicated that concurrent presentation of Auditory Icons is less damaging when prior classification is used as it improves the accuracy of identification by approximately 7%. A caveat about the result for the third study was that it had some procedural changes when compared to the first and second study. These changes resulted in more sounds in the nine simultaneous condition with no prior classification having distinct object and action categories. This was due to the random selection of the sounds from a larger stimuli pool. It is hypothesised the results would be worse if there was a greater degree of overlap between action or object categories for this particular study. The experiment suggested that conditions using prior classification based on object and action properties had a higher rate of identification for the Auditory Icons. The use of prior classification has been found in this research as a method of achieving lower levels of interference between Auditory Icons. Ballas's method of causal uncertainty (1986) can be used to further refine the sound selections by determining which sounds are most confused. These sounds can be removed and replaced with less confused and more perceptually distinct sounds. The results from the studies described here indicate that if Auditory Icons are carefully designed they can be used as an effective means of communicating multiple messages or bits of information using sound. This research answers RQ3 and provides a method of determining Auditory Icon identification in concurrent presentation situations.

RQ4 — Is it possible to reduce the subjective confusion of the sounds using action and object categories ? This question is somewhat related to RQ3 but it is focused on understanding more about the approach of prior classification and if it can improve the number of sounds that can be successfully identified. In order to robustly design Auditory Icons, an understanding of the issues affecting their identification is required. Other studies (Howard and Ballas, 1980) indicate that one important factor for the identification of Auditory Icons is the number of alternative interpretations that listeners have for a particular sound. RQ4 uses the method of *causal uncertainty* (Ballas, 1993) to determine the number of alternative interpretations or causes for a particular sound from participants. The causal uncertainty measure analysed the free text response from the participants and was used in Chapter 5. The results of the causal uncertainty measures are not as clear cut as average identification rates in discussed RQ3, one reason for this is the results are affected by the identification rates as conditions with larger numbers of concurrent sounds where the sounds could be masked or unnoticed. This can adversely affect the causal uncertainty measure for that condition and needs to be considered when using the method for selecting sounds for concurrent presentation.

The results of the action and agent/object identification showed that both the actionhood and agent/objecthood of sounds are generally well identified even in cases with higher numbers of concurrently presented sounds. Across the concurrent presentation conditions in the first two studies, where the range spanned from 3 to 10 sounds it was actions rather than objects that were best identified. The actionhood of the Auditory Icons were better identified as shown in Figures 5.6 and 5.15 (actions) and in Figures 5.18 and 5.8 (objects) are compared. The third study explored a wider range of sound that similarly highlighted that sounds with less confusion were better identified. Future investigations could determine if applying acoustic comparisons of the sounds within a condition to determine potential masking of sounds would improve the action or object identification of the sounds within a condition. The results in Chapter 5 have shown that a single sound or a couple of sounds (in the conditions with 6+ Auditory Icons) were responsible for the majority of participant confusion in identifying objecthood or actionhood. This suggests that the use of masking analysis may not be as important as determining the sounds with particularly high causal uncertainties and finding a similar but alternative sound to replace it in the particular condition may be as effective as determining masking for improving sound identification. The best approach would be to ensure no masking occurred and to use sounds with lower causal uncertainty measures. This study has addressed this question and extended the method of causal uncertainty for use in concurrent Auditory Icon presentation.

RQ5 — Can a listener's tacit criteria about how they attribute meaning to everyday sounds be elicited ? The richness of the textual descriptors from the concurrent presented Auditory Icon studies motivated this question as the results hinted at the how tacit knowledge was being used by listeners when they classified the everyday sounds presented. A descriptive analysis technique based on the Repertory Grid Technique (Kelly, 1955) was used to assist in answering this question and provided a scientific methodology that helped in interpreting textual descriptors. Chapter 6 used the Repertory Grid Technique (RGT) to investigate RQ4 by analysing participant's personal constructs to gain a better insight into to their tacit knowledge using two different studies with a focus on the individual's classifications of everyday sounds. The constructs of participants were elicited, these constructs hold a specific individualised meaning of the sound for the participants. The chapter's second study combined the repertory grid method with the previously described method of causal uncertainty of free-text descriptions from Chapter 5 to provide deeper insights into the data. The data was classified using the existing taxonomy from the CLOSED project (2007a) and this taxonomy was extended to include animal and bird sound aspects. Whilst other studies have explored the Repertory Grid Technique (RGT) to investigate spatial audio (Berg and Rumsey, 1999), or individual timbre space construction (Atsushi and Martens, 2005), or for evaluating the perceptual differences in multichannel microphone techniques (Martens and Sungyoung, 2007), there is no existing work using this technique for exploring individual personal constructs of Auditory Icons.

The results of the RGT technique differed slightly in the two studies and this is attributable to two different sets of participants and to the two different sound collections used. The visualisations from the analysis of the RGT data of the individual participants provide an accessible way for representing the meaning structures or semantic constructs (Fransella et al., 2004) used by the participants. The work in the studies did not produce group summary data as occurred for the explorations for RQ1, RQ2, or RQ3. As participant's individual language and unique constructs were elicited there was no single measure or metric that could quantify the data as a group summary. This chapter interpreted the personal constructs and elements from each participant which were classified within the CLOSED project's taxonomy as discussed in Section 2.1.1. The results from the first study match available categories in this framework but the results from the second study did not fit within this framework. The results of this second study suggested a number of possible extensions to the CLOSED framework to include several animal and bird categories. The second study found that it was possible to complement the repertory grid method with the causal uncertainty technique. The results of this technique found that two or three sounds were responsible for the majority of confusion in the second study. Replacing the sounds with other similar sounds and repeating the testing could ensure the selection of sounds more identifiable to listeners.

A considered discussion on this issue is given in the next Section 7.3 and situates the views derived from participants in a broader classification framework that can contain individual idiosyncratic views while situating the views within an overall framework for the classification of sound.

7.3 Generalising the Results

This thesis has synthesised and organised existing research on Auditory Icons, and extended this work with the studies presented. The results of this thesis is generalised to create the foundations of a framework for designers. The framework provides a structure approach to the methods presented in this thesis. It is aimed at by designers and tailored for use outside of strict laboratory conditions. Designers need empirically based or inspired methods to guide their overall design process which do not suffer from the specificity of psychoacoustic studies or that require a relatively long time to conduct. A typical design problem is wider than those addressed by psychoacoustic studies and the approach of this thesis joins these disciplines in a manner that is accessible at a reasonable cost to designers. The framework is discussed in more detail in Section 7.3.1.

A new application and methodology was presented for exploring sounds in a multidimensional space to verify their mapping in a perceptually relevant manner. This approach helps those seeking to explore large sets of sounds interactively, allowing tagging and scaling information to be easily collected. Designers can use this approach to test potential mappings between everyday sounds and Auditory Icons. Concurrent presentation of audio can lead to confusion and problems with identification, the work in this thesis highlighted one approach for sound selection for simultaneous presentation using prior classification to ensure understandability and ease of identification. Designers who follow this approach can convey more information using several Auditory Icons at once. Vocabulary, metaphors, and tacit knowledge can be difficult to elicit in a manner that is useful for designers. The work on extending the Repertory Grid Technique (Kelly, 1955) and combining it with Ballas's causal uncertainty (Ballas, 1993) helps designers find this type of information. These types of experiments can potentially be applied to all types of sound but this thesis uses and only makes its claims for everyday sounds. This thesis does not give any prescriptive rules and guidelines with the exception of reiterating some well established psychoacoustics guidelines. The results of this thesis extended the avenues of research for Auditory Icons within Auditory Display.

The results from extending the CLOSED framework as shown in Figure 7.1 and from the explorations in Chapters 4, 5, and 6 have shown that participants were able to scale, identify, and had a tacit understanding of the sounds used in the experiments. The results from the experiments in this thesis underlie the importance of a sound's action and of a sound's source, whether it be an agent or an object, when designing and understanding Auditory Icons. This is shown in Figure 7.1 and helps in validating results such as the work of McGregor et al. (2006) in soundscape classification studies as discussed in Chapter 2 and of Gaver (1993a,c) as well as many others as discussed in Section 2.3. The work in this thesis is not limited to validating existing models for classification or categorisation. The results from this thesis provide a number of methods that may serve as tools for future research such as addressing the extension of the framework shown in Figure 7.1. This framework builds upon prior research as discussed in Chapter 3 and itself will need further studies to address its limitations.

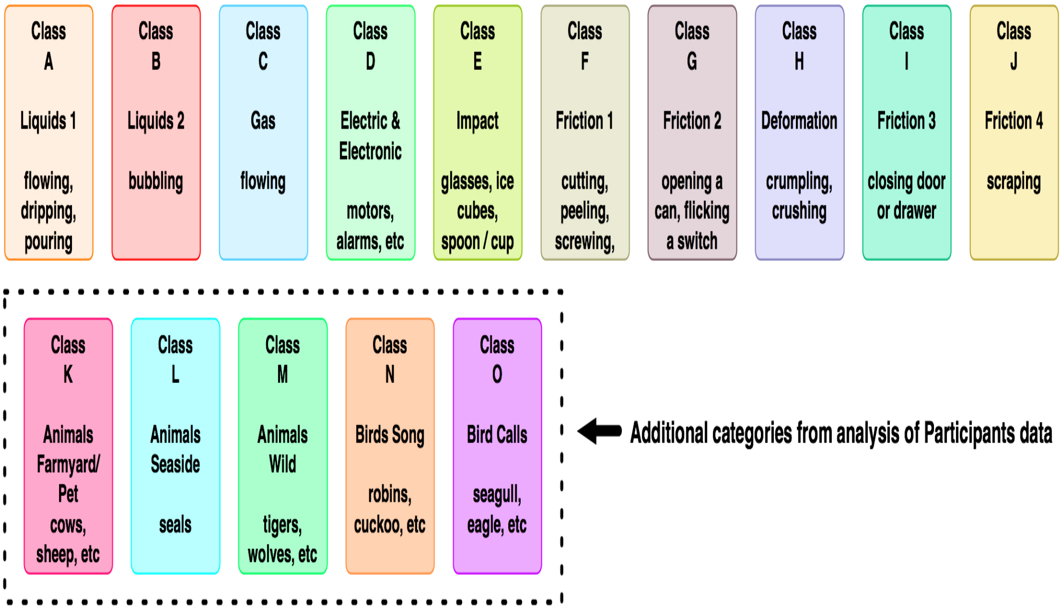


Fig. 7.1: *The extended framework derived from the CLOSED project's (2007a) taxonomy with sources added from earlier experimental results as shown in Chapter 6.*

7.3.1 Foundations for an Empirically Inspired Design Framework for Auditory Displays

The work in this thesis implicitly conveyed the foundations of a framework for the selection and evaluation of Auditory Icons in the early conceptual stages of design. This section briefly describes each of the steps and the results. The approach is depicted in block diagram form in Figure 7.2. The aim of the explorations in this thesis were to ensure the accessibility and practicality of the methods for Auditory Display designers. This allowed the framework to be verified and provided a documented process for designers. It provides new methods for knowledge elicitation and refined the information for the evaluation and design of Auditory Icons. It highlighted the general applicability of this approach through the use of multiple different sets of Auditory Icons. These results suggest that other sound categories such as speech or Earcons could be assessed by the framework described. The framework is believed to be adaptive and expandable into other contexts such as other sound categories. The foundations of this framework are presented in this thesis as other types of sounds and methods as it is designed to be open to new additions and evolution. The hope is that future research will grow this foundation to address a wider range of issues within the field of Auditory Display. Studies are needed to gauge the value of this framework when applied by designers within their own work practises. This is an area of future work to be pursued.

The first step is the definition of what the Auditory Display will be used for and its context. The second step is to select the sounds that will populate the display, here the framework can guide designers regardless of whether they choose real or synthetic sounds. Auditory Icons typically have values or ranges mapped to their attributes (i.e. size, materiality, etc.) and it is important to understand if this mapping works, this is addressed in Step 3. The problem of clarity with regards to identification where listeners hear several sounds at once is addressed by the method presented in Step 4 which leads to Step 6. The problem of gaining deep insight into listeners categorisation is addressed in Steps 4 and Steps 6–10, which provide an approach to elicit and analyse responses on a per individual basis.

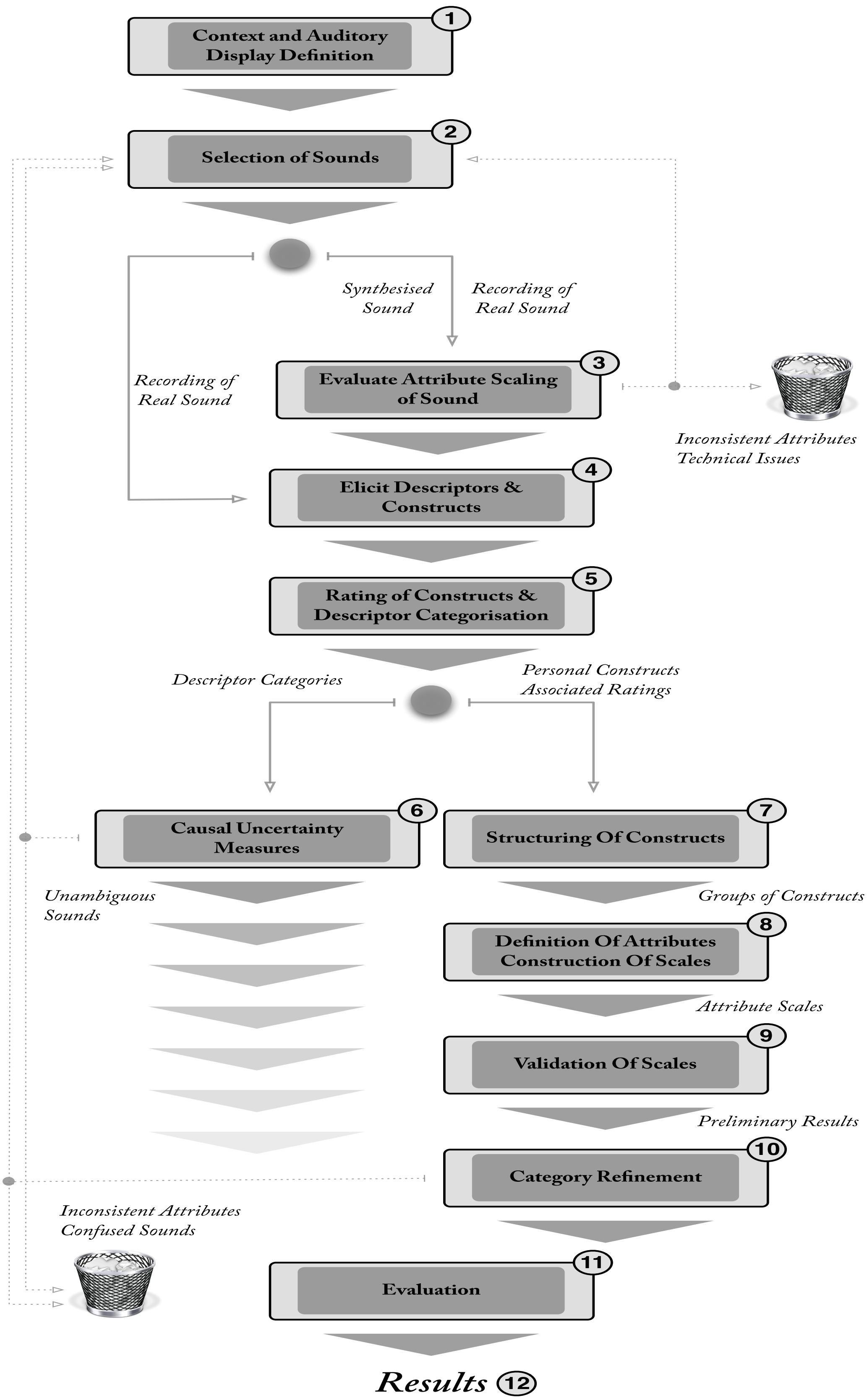


Fig. 7.2: *The empirically inspired design framework used for Auditory Icon selection at the early stages of design as proposed by this thesis.*

1. *Context and Auditory Display Definition*: The purpose of the Auditory Display is defined, investigation of Auditory Icons at the early stages of conceptual design and what listeners perceive in terms of attributes and tacit knowledge.
   Results: Sound requirements are defined, a set of desirable sound mappings are selected for evaluation. Listeners are recruited for listening tests.

2. *Selection of Sounds*: A pool of sounds that fit the selected mappings are gathered and organised for evaluation. These sounds can be real, synthetic or a mix of both.
   Results: A stimulus set of sounds.

3. *Evaluate Attribute Scaling of Synthesised Sound*: Participants are asked to listen and compare synthetic and real sounds by scaling them using a computer interface. The scaling investigated were derived using the sound mappings from the first stage. Each participant makes individual scaling judgements as well as determining realism for the sounds presented. Real, synthesised, or a mix of such sounds can be testing using this method. This can be seen in more detail in Chapter 4.
   Results: The pilot involved 4 people and 18 sounds while the study involved 5 people and 18 different sounds. Traditional approaches such as pairwise comparison would have required 1377 comparisons, this approach allowed listeners to appreciate the entire set while making the comparisons as opposed to pairwise comparison tasks.

4. *Elicit Descriptors & Constructs*: 11, 5, and 26 participants created descriptors for the 18, 34, and 18 sounds presented respectively, this is detailed in Chapter 5. Constructs were created from two groups of 5 listeners who produced written descriptors of two different sets of sounds, contain 18 and 25 sounds respectively. The method used to elicit constructors is presented in Chapter 6.
   Results: Approximately 600 descriptors and 215 elicited constructs.

5. *Rating of Constructs & Descriptor Categorisation*: Each subject rated the stimuli using these constructs created in the last stage. The rating methods are presented in Chapter 6. The descriptors were categorised by the experimenter into action and object categories. This categorisation is discussed in Chapter 5.
   Results: The rating data of 215 elicited constructs and the object / action categorisation of the written descriptors.

6. *Causal Uncertainty Measures*: The categorisation details were used to calculate the causal uncertainty of sounds. The method and details on its first use for concurrent audio are presented in Chapter 5. It was further shown how it could be combined as

another viewpoint when conducting construct elicitation. This is discussed in Chapter 6. *Results*: The causal uncertainty measures were taken from four studies with 11, 5, 26, and 5 participants respectively where 18, 34, 18, and 25 sounds presented respectively.

7. *Structuring of Constructs*: Cluster analysis, multidimensional scaling and principal component analysis were made on the ratings data to clarify attributes and to reduce dimensionality and remove redundancy. These methods are documented in detail in Appendix H.1 with the overview of results shown in Chapter 6.
   Results: Construct and element groups were identified from two groups of 5 listeners, the first group heard 8 sound triads and the second group heard 14 sound triads to create the constructs (5x8=40,5x14=70). The same two groups of listeners identified the elements, the first group heard 12 sounds and the second group heard 19 sounds to create the elements (5x12=60, 5x19=95). These studies are discussed in Chapter 6.

8. *Definition of Attributes, Construction of Scales*: The construct groups were analysed for their content. The appropriate descriptions for the participant identified attributes were formulated. The rating scales were defined.
   Results: A set of attribute scales, in the form of written descriptions were created.

9. *Validation of Scales*: The scales created were explored in terms of existing categorisations and taxonomies to test the appropriateness of the scales generated.
   Results: This resulted in modifications to the CLOSED sound taxonomy to include various animal sounds.

10. *Category Refinement*: The details from the earlier causal uncertainty measures and from the scales were used to suggest the removal of particular sounds as unsuitable for use in the particular sonic context.
    Results: A revised set of sounds for use were defined.

11. *Evaluation*: Three hypothetical domains were proposed throughout the thesis as exemplars of where these methods could be used in practical situations. A ambient Auditory Display for co-located colleagues is discussed in Appendix I.1 and the second study in Chapter 6 considered potential Auditory Icons for this system.
    Results: We elaborate upon the differences between domains in relation to how Auditory Icons might be used.

12. *Results*: The framework and its complementary methods provide detailed results on the set of Auditory Icons.

Simplification of an Empirically Inspired Design Framework for Auditory Displays

The presented framework consists of a number of steps, it is envisaged that in future when Auditory Icons and their subjective qualities are better known, some stages may be simplified or found to be redundant. The use of multiple techniques for triangulation in the method may also indicate that additional steps may be added when appropriate new techniques are found and validated with the evaluation method.

Generalisability of an Empirically Inspired Design Framework for Auditory Displays

The work in this thesis and its studies employed knowledge elicitation and refined the information for the evaluation and design of Auditory Icons. Descriptors and personal constructs were the results of the elicitation, which used Auditory Icons as sound stimuli. This thesis used different sets of Auditory Icons, which resulted in changes in the descriptors and personal constructs provided by participants. This highlights the applicability of this method across a wide range of everyday sounds. This observation indicates that similar sound stimuli such as speech or Earcons can be assessed by the framework described. The framework is adaptive and expandable to other contexts.

7.4 Limitations

There are a number of limitation with this work that should be noted. A major limitation of this work was the setting aside of spatialisation issues as the broader issue of spatialisation was set as beyond the scope of this thesis. Further studies are needed to explore the issue of how spatialisation effects Auditory Icons with regard to the results presented in this thesis. Another limitation is due to the different sets of Auditory Icons used in experiments lead to another argument about the lack of consistency between the stimuli used in the studies presented in Chapters 4, 5, and 6. The reason for the broad choice of stimuli used in the experiments was to explore different types of sounds and to attempt to prevent any effects arising from the use a single stimuli set. These effects could affect the whole set or any individual stimuli. The effects could vary from a particular recording level being muted to a particular type of impact sound or a particular type of animal sound having an individualised contexts for listeners. The use of similar but different stimuli can help in minimising any negative effects.

An additional concern was related to the experimenter's codification of the participant's descriptors into action and object categories. The coding of the object and action descriptors has two limitations; it was performed by a single rater, and the categorisation of the participant's free text responses represents a simplification of the complex descriptions given. The effect of prior experience with training and the correlation results from earlier studies would suggest that using a single codifier or sorter should not be a major issue but further studies are required to clarify this issue. Important nuances might have been lost and some descriptions may have been “simplified”. This approach has been used in research by Fernström et al. (2005), where this author participated as one of the three description codifiers, by Houix et al. (2007a) in the EU Closed Project, and also by Tardieu et al. (2008). The works by Fernström et al. (2005) and by Houix et al. (2007a) used three different people who functioned as human classifiers / sorters and codified the participants responses. The work by Houix et al. (2007a) used a single stage classification which took approximately three days per classifier and where the classifiers were instructed to keep their criteria consistent. Fernström et al. (2005) used a two stage classification which took approximately four days where classifiers did not have to keep the criteria consistent in the first stage. In the second stage, the criteria were grouped by classifiers to reduce the number of criteria while maintaining consistency to the internal measures that the classifiers had used. The work in this thesis used a similar two stage classification process which was carried out by this author only for all the responses from Chapter 5 and from Chapter 6. Houix et al. (2007a) analysed the correlations between classifiers. This showed similar results for object classifications and a degree of dissimilarity for the results of action classification but they tended to be related to a small number of sounds which had a large degree of dissimilarity between classifiers. Unpublished data related to the earlier study by Fernström et al. (2005) found strong correlations between both object and action classifiers. These studies indicate that the use of a single rater is acceptable in the theory and method exploration approach taken by this thesis. In situations where these methods are being applied for real world testing and Auditory Displays; additional raters would help ensure that descriptions were not overly “simplified” when categorised. However literature (Bech and Zacharov, 2006, Atsushi and Martens, 2005, Martens and Sungyoung, 2007) is sparse on detailing how many additional raters would be required. The clearest advice is from Houix et al. (2007a) suggesting that with three or more raters the categorisations and related measures could be assessed for reliability and reproducibility.

The final concern or limitation of the research presented in this thesis was that it used only a single taxonomy to express the results. This may not adequately answer the question of how listeners organise their individual categories of sounds. A single hierarchy may not be suitable for representing how a listener organises sounds internally. Conceptual representation research has explored alternative organisations or cross-classifications (Ross and Murphy, 1999) where items can potentially be part of several categories simultaneously. The research by Ross and Murphy (1999) found that for food items there exists a non-hierarchical network of category relations where the items are connected to each of the categories they represent. Items are related to other items where they share a category membership or where they have shared properties. This shows for everyday sound perception, a categorical structure for sounds does not have to be limited to a single hierarchical structure. The CLOSED project's taxonomy introduces relationships between items at the same level but also between items at different levels of abstraction.

7.5 Future Work

This section provides a road map for ways this thesis can be further developed to improve the understanding of Auditory Icon presentation and identification.

7.5.1 Spatial separation and identification

Attention competition between auditory objects is also another problem for listeners (Shinn-Cunningham et al., 2005). The work by Shinn-Cunningham et al. (2007) has indicated that the final interpretation of an auditory scene may be dependant upon the object that is the focus of attention for the listener. They further suggest that the auditory scene and the understanding by the listener of its organisation is related to how well they can reduce interference from unwanted objects and focus on a particular object of interest. Understanding the effect of spatial cues can help in the understanding of ambiguous auditory scenes. Further investigations with everyday sounds and Auditory Icons and what effects spatial cues have on these types of sound would be a useful area of study.

7.5.2 Exploring richer acoustic synthesised sounds

The synthesised sounds in Chapter 4 has a number of issues with them including acoustic richness (Carello et al., 2003). Newer synthesised models need to be developed and tested. The advantages for synthesised parametrically controllable models suggest that despite the poor results in this thesis that the approach merits further study. The simplified or "*cartoonified*" (Fernström, Brazil and Bannon, 2005) approach is an area that needs further work. Work on "room acoustics" also need explorations, however new developments such as the development of parallel algorithms for estimating sound radiation (Zheng and James, 2009) could provide solutions to these issues and the wider question of acoustic richness.

7.5.3 Exploiting more Auditory Icon parameters

The work on parametrically controlled synthesised sounds in Chapter 4 has shown that even unrealistic synthetic sounds can convey perceptual dimensions to listeners. The synthetic sounds were not as meaningful to listeners and have a wider variance than real sounds. This does not completely negate this approach but suggests its use needs to be carefully considered. The technique used in Chapter 4 was equally applicable to both real and synthesised sounds and offers a means for testing parameter mapping in Auditory Icons. In Chapter 4, two distinct perceptual dimensions were used for the scaling task, it is likely that increasing the number of parameters will require mapping to additional perceptual dimensions. An increase in parameters mapped to a sound could increase its interpretation difficulty and its identifiability however they would offer the opportunity to convey more information to listeners.

7.5.4 Presenting more Auditory Icons concurrently

In this thesis concurrent Auditory Icon presentation was explored as part of one of the studies. The next step in this process is to determine the limits of identification and of confusion for concurrently presented Auditory Icons. The issue of cacophony may seem a minor nuisance in many circumstances but in medical or control room situations, understanding and preventing it is crucial for successful operations.

7.5.5 Identification in mixed auditory environments

This thesis has focused on Auditory Icons but it is unlikely that an auditory interface use just Auditory Icons. Hybrid auditory interfaces can use Earcons, speech, or Auditory Icons where most appropriate. Earcons and speech are alternative ways for encoding data in sound. They offer different advantages and disadvantages. They may be more suitable for the data depending on the data itself and the context. Researchers and designers should have an understanding of the interactions occurring between Auditory Icons, sonification, speech, and Earcons but such an endeavour is vast and complex, requiring many experiments to explore. The work in this thesis can provide a foundation for understanding Auditory Icons, just as McGookin (2004) provides an understanding for concurrent Earcons and as both have noted, the space between these two types of sound offers a rich area for exploration.

7.6 Final Remarks

The concluding remarks of this thesis are that Auditory Icons and everyday sounds can offer several advantages, however there are several issues that need to be considered when designing and presenting them. This thesis presented a body of work exploring the identification of Auditory Icons, the meaning of Auditory Icons as construed by listeners, the effect of the realism of Auditory Icons, the scaling of synthetic Auditory Icons, as well as providing suggestions for how to improve their design. The body of work presented in this thesis has added to the understanding of Auditory Icons and opens new avenues for research.

References

Afonso, A., Katz, B. F. G., Blum, A., Jacquemin, C., and Denis, M. (2005). A study of spatial cognition in an immersive virtual audio environment: Comparing blind and blindfolded individuals. In Brazil, E., editor, *International Conference on Auditory Display (ICAD-05)*, pages 228–235, Limerick, Ireland.

Agresti, A. (1996). *An Introduction to Categorical Data Analysis*. Wiley Series In Probability and Statistics. John Wiley and Sons.

Albers, M. (1994). The varese system, hybrid auditory interfaces, and satellite-ground control: Using auditory icons and sonification in a complex, supervisory control system. In Kramer, G. and Smith, S., editors, *Second International Conference on Auditory Display*, pages 3 – 14, Santa Fe, N.M.

Aramaki, M. and Kronland-Martinet, R. (2006). Analysis-synthesis of impact sounds by real-time dynamic filtering. *IEEE Transactions on Speech and Audio Processing*, 14(2):695–705.

Atsushi, M. and Martens, W. L. (2005). Constructing individual and group timbre spaces for sharpness matched distorted guitar timbres. In *Proceedings of the 119th Convention of the Audio Engineering Society*, pages 1–15, New York, NY, USA.

Ballas, J. A. (1989). Acoustic and perceptual-cognitive factors in the identification of 41 environmental sounds. Tech Report ONR-89-2, Center for Behavioral and Cognitive Studies, George Mason University, Fairfax, VA.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *J. of Experimental Psychology*, 19(2):250–267.

Ballas, J. A. (1994). Delivery of information through sound. In Kramer, G., editor, *Auditory Display: Sonification, Audification and Auditory interfaces*, pages 79–94. Addison-Wesley Publishing Company, Reading, MA, USA.

Ballas, J. A. and Barnes, M. E. (1988). Everyday sound perception and aging. In *Proceedings of the 32nd Annual Meeting of the Human Factors Society*, pages 194–197, Santa Monica, CA. Human Factors Society.

Ballas, J. A., Dick, K. N., and Groshek, M. R. (1987). Failure to identify “identifiable” sounds. In *Proceedings of the 31st Annual Meeting of the Human Factors Society*, pages 144–146, Santa Monica, CA. Human Factors Society.

Ballas, J. A. and Howard, J. H. (1987). Interpreting the language of environmental sounds. *Environ. Behav.*, 1(9(1)):91–114.

Ballas, J. A. and Mullins, R. (1991). Effects of context on the identification of everyday sounds. *Human performance*, 4:199–219.

Ballas, J. A., Sliwinsky, M., and Harding, J. (1986). Uncertainty and response time in identifying non-speech sounds. *Acoustical Society of America*, 79.

Ballas, J. A. and Sliwinsky, M. J. (1986). Causal uncertainty in the identification of environmental sounds. Tech Report ONR-86-1, Office of Naval Research, Department of Psychology, Georgetown University, Washington, D. C.

Basso, K. H. (1996). *Senses of Place*, chapter Wisdom Sits in Places - notes on a Western Apache landscape, page 55. School of American Research Press.

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

BBC (2006). BBC sound effects library: sound effects library.

Bech, S. and Zacharov, N. (2006). *Perceptual Audio Evaluation- Theory, Method and Application*. John Wiley and Sons.

Bell, G. (1999). Making sense of museums. Tech report, Intel Labs, Portland, Oregon, USA.

Bell, G. (2001). Looking across the atlantic: Using ethnographic methods to make sense of europe. *Intel Technology Journal*, (3):1–10.

Benford, S. and Greenhalgh, C. (1997). Introducting third party objects into the spatial model of interaction. In Thimbleby, H., O'Conaill, B., and Thomas, P., editors, *Fifth European Conference on Computer Supported Cooperative Work*, pages 189–204, Lancaster, UK. Kluwer Academic Publishers.

Berg, J. (2005). Opaque - a tool for the elicitation and grading of audio quality attributes. In *Proceedings of the 118th Convention of the Audio Engineering Society*, Barcelona, Spain.

Berg, J. and Rumsey, F. (1999). Spatial attribute identification and scaling by repertory grid technique and other methods. In *Proceedings of the 16th AES International Conference on Spatial Sound Reproduction*, Rovaniemi, Finland.

Bertin, J. (1983). *Semiology of Graphics*. University of Wisconsin Press, Milwaukee, Wis.

Best, V. A. (2004). *Spatial Hearing with Simultaneous Sound Sources: A Psychophysical Investigation*. Ph.d., Faculty of Medicine, University of Sydney.

Beyer, H. and Holtzblatt, K. (1997). *Contextual Design: Defining Customer-Centered Systems*. Morgan Kaufmann.

Blattner, M., Sumikawa, D., and Greenberg, R. (1989). Earcons and icons: Their structure and common design principles. In *ACM CHI '89*, pages 548–574, Austin, Texas, USA. ACM Press.

Bliss, J. P. and Dunn, M. C. (2000). Behavioural implications of alarm mistrust as a function of task workload. *Ergonomics*, 43(18):1283–1300.

Bodden, J. L. and Klein, A. J. (1972). Cognitive differentiation and affective stimulus valus in vocational judgements. *J. of Vocational Behaviour*, 3:75–79.

Boltzmann, L. (1872). Weitere studien uber das warmegleichgewicht unter gasmolekulen [further studies on thermal equilibrium between gas molecules]. *Sitzungsberichte der Akademie der Wissenschaften*, Mathematische-Naturwissenschaftliche Klasse(Bd. 66):275–370.

Bonebright, T. (2001). Perceptual structure of everyday sounds: A multidimensional scaling approach. In Hiipakka, J., Zacharov, N., and Takala, T., editors, *ICAD 2001*, pages 73–78, Helsinki, Finland. Laboratory of Acoustics and Audio Signal Processing and the Telecommunications Software and Multimedia Laboratory, Univeristy of Helsinki, Espoo, Finland.

Bonebright, T. L. (1996). An investigation of data collection methods for auditory stimuli-paired comparisons versus a computer sorting task. *Behaviour Research Methods, Instruments, and Computers*, 28(2):275–278.

Bonebright, T. L., Miner, N. E., Goldsmith, T. E., and Caudell, T. P. (2005). Data collection and analysis techniques for evaluating the perceptual qualities of auditory stimuli. *ACM Transactions on Applied Perceptions*, 2(4):505–516.

Borg, I. and Groenen, P. J. F. (1996). *Modern Multidimensional Scaling: Theory and Applications*. Springer Series in Statistics. Springer-Verlag.

Borg, I. and Shye, S. (1995). *Facet Theory: Form and Content*, volume 5 of Advanced Quantitative Techniques in the Social Science Series. Sage, Thousand Oaks, California.

Bovermann, T., Hermann, T., and Ritter, H. (2006). A tangible environment for ambient data representation. In McGookin, D. and Brewster, S., editors, *First International Workshop on Haptic and Audio Interaction Design*, volume 2, pages 26–30.

Brazil, E. (2003). *Investigation of multiple visualisation techniques and dynamic queries in conjunction with direct sonification to support the browsing of audio resources*. Msc dissertation, University of Limerick.

Brazil, E., Fernström, M., and Ottaviani, L. (2003). A new experimental technique for gathering similarity ratings for sounds. In Brazil, E. and Shinn-Cunningham, B., editors, *International Conference on Auditory Display (ICAD-03)*, pages 238–242.

Bregman, A. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. M.I.T. Press, Cambridge, MA, USA.

Brewster, S. A. (1994). *Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces*. Ph.d., University of York.

Brewster, S. A. (2002). *The Human Computer Interaction Handbook*, chapter Non-speech auditory output, pages 220–239. Lawrence Erlbaum Associates.

Brungart, D., Ericson, M., and Simpson, B. (2002). Design considerations for improving the effectiveness of multitalker speech displays. In *International Conference on Auditory Display ICAD-02*, pages 424–430, Kyoto, Japan.

Brungart, D. and Simpson, B. (2002). The effects of spatial separation in distance on the information and energetic masking of a nearby speech signal. *J. Acoustic Society of America*, 112(2):664–676.

Brunswik, E. (1952). *International encyclopedia of unified science*, chapter The conceptual framework of psychology, pages 656–760. University of Chicago Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62:193–217.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. University of California Press, 2nd edition.

Buja, A. and Swayne, D. F. (2002). Visualization methodology for multidimensional scaling. *Journal of Classification*, 19(1):7–44.

Calder, B. J., Phillips, L. W., and Tybout, A. M. (1981). Designing research for application. *J. of Consumer Research*, 8(2):197–207.

Carello, C., Wagman, J. B., and Turvey, M. T. (2003). *Moving Image Theory: Ecological considerations*, chapter Acoustical specification of object properties, pages 79–104. Southern Illinois University Press, Carbondale, IL.

Carreria-Perpinan, M. A. (2001). *Continuous latent variable models for dimensionality reduction and sequential data reconstruction*. PhD thesis, Department of Computer Science, University of Sheffield.

Catsoluis, J. (2002). *Designing Embedded Hardware*. O'Reilly Press, 1st edition.

Cherry, E. C. (1953). Some experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America*, 25:975–979.

Cherry, E. C. and Taylor, W. K. (1954). Some further experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America*, 26:549–554.

Chion, M. (1994). *Audio-Vision: Sound on Screen*. Columbia University Press, 1st edition.

Choisel, C. and Wickelmaier, F. (2005). Extraction of auditory features and elicitation of attributes for theassessment of multichannel reproduced sound. In *Proceedings of the 118th Convention of the Audio Engineering Society*, Barcelona, Spain.

Cohen, J. (1992). A power prime. *Psychological Bulletin*, 112(1):155–159.

Cohen, J. (1994a). Monitoring background activities. In Kramer, G., editor, *Auditory Display: Sonification, Audification and Auditory interfaces*, pages 499–532. Addison-Wesley Publishing Company, Reading, MA, USA.

Cohen, J. (1994b). Out to lunch: Further adventures monitoring background activity. In *Second International Conference on Auditory Display (ICAD '94)*, pages 15–20, New Mexico. Santa Fe Institute.

Cook, P. R. (2002a). Modeling bill's gait: Analysis and parametric synthesis of walking sounds. In *22nd International Conference of the AES: Virtual, Synthetic, and Entertainment Audio*, pages 73–78. AES.

Cook, P. R. (2002b). *Real sound synthesis for interactive applications*. A.K. Peters, Natick, MA, USA, 1st edition.

Costall, A. and Still, A. (1989). Gibson's theory of direct perception and the problem of cultural relativism. *Journal for the Theory of Social Behaviour*, 19(4):433–441.

Courage, C. and Baxter, K. (2005). *Understanding your users: A practical guide to user requirements*. Morgan Kaufmann.

Crawley, M. J. (2005). *Statistics: An Introduction Using R*. John Wiley And Sons, Chichester, 1st edition.

Csikszentmihalyi, M. (2000). *Beyond Boredom and Anxiety: Experiencing Flow in Work and Play*. Jossey Bass Wiley.

Dahl, D. B. (2007). *xtable: Export tables to LaTeX or HTML*. R package version 1.5-2.

Dietz, P. and Leigh, D. (2001). Diamondtouch: A multi-user touch technology. In *Proceedings of UIST 2001*, pages 219–226.

Dourish, P. and Bellotti, V. (1992). Awareness and coordination in shared workspaces. In *ACM Conference on Computer Supported Cooperative Work (CSCW)*, pages 107–114.

Draffan, J. W. (1973). Randomness in grid test scores. *British Journal of Medical Psychology*, (46):391–392.

Ellis, D. (1996). *Prediction-driven computational auditory scene analysis*. Phd thesis, Massachusetts Institute of Technology.

Erickson, T., Halverson, C., Kellogg, W. A., Laff, M., and Wolf, T. (2002). Social translucence: Designing social infrastructures that make collective activity visible. *Communications of the ACM*, 45(4):40–44.

Erickson, T. and Kellogg, W. A. (2000). Social translucence: An approach to designing systems thas support social processes. *ACM Transactions on Computer-Human Interaction*, 7(1):59–83.

Everitt, B. S. and Hothorn, T. (2006). *A Handbook of Statistical Analyses using R*. Chapman and Hall.

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

Fishkin, K. P., Gujar, A., Harrison, B. L., Moran, T. P., and Want, R. (2000). Embodied user interfaces for really direct manipulation. *Communications of the ACM*, 43(9):74–80.

Fitch, W. and Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory over visual display in a complex, multivariate system. In Kramer, G., editor, *Auditory Display: Sonification, Audification and Auditory interfaces*, pages 307–326. Addison-Wesley Publishing Company, Reading, MA, USA.

Fodor, J. A. and Pylyshyn, Z. W. (1981). How direct is visual perception?: Some reflections on gibson's "ecological approach". *Cognition*, 9:139–196.

Fraden, J. (2004). *Handbook of Modern Sensors: Physics, Designs, and Applications*. Springer-Verlag, 3rd edition.

Fraley, C. and Raftery, A. (2007). *mclust: Model-Based Clustering / Normal Mixture Modeling*. R package version 3.1-1.

Fransella, F., Bell, R., and Bannister, D. (2004). *A manual for repertory grid technique*. John Wiley and Sons.

Frauenberger, C., Stockman, T., and Bourguet, M. L. (2007). A survey on common practice in designing audio in the user interface. In *Proceedings of BCS HCI2007*, pages 187–194. British HCI Group.

Garbarini, F. and Adenzato, M. (2004). At the root of embodied cognition: Cognitive science meets neurophysiology. *Brain and Cognition*, (56):100–106.

Gaver, W. W. (1988). *Everyday Listening and Auditory Icons*. PhD thesis, Cognitive Science and Psychology, San Diego, University of California.

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

Gaver, W. W. (1991). Sound support for collaboration. In *ECSCW'91*, Amsterdam. Kluwer, Dordrecht, Holland.

Gaver, W. W. (1993a). How do we hear in the world? explorations of ecological acoustics. *Ecological Psychology*, 5(4):285 – 313.

Gaver, W. W. (1993b). Synthesizing auditory icons. In *Interchi'93*, pages 228–235. ACM Press.

Gaver, W. W. (1993c). What in the world do we hear? an ecological approach to auditory source perception. *Ecological Psychology*, 5(1):1–29.

Gaver, W. W. (1994). Using and creating auditory icons. In Kramer, G., editor, *Auditory Display: Sonification, Audification and Auditory interfaces*, pages 417–446. Addison-Wesley Publishing Company, Reading, MA, USA.

Gaver, W. W. (1997). Auditory interfaces. In Helander, M., Landauer, T., and Prabhu, P., editors, *Handbook of Human-Computer Interaction*, pages 1003–1042. Elsevier Science, Amsterdam, The Netherlands, 2 edition.

Gaver, W. W., Beaver, J., and Benford, S. (2003). Ambiguity as a resource for design. In *Proceedings of CHI 2003 - New Horizons*, volume 5, pages 233–240. ACM.

Gaver, W. W., Dunne, T., and Pacenti, E. (1999). Cultural probes. *Interactions*, 6(1):21–29.

Gaver, W. W. and Smith, R. (1990). Auditory icons in large-scale collaborative environments. In Diaper, D., editor, *INTERACT'90*, pages 735–740. Elsevier Science Publishers B.V. (North-Holland).

Gaver, W. W., Smith, R., and O'Shea, T. (1991). Effective sounds in complex systems: the arkola simulation. In *CHI'91*, pages 85–90, New Orleans, Louisiana, USA. ACM Press.

Gawron, V. J. (2000). *Human Performance Measures Handbook*. Lawrence Erlbaum Associates Inc. Publishers, Mahwah, New Jersey.

Gérard, Y. (2004). *Mémoire sémantique et sons de l'environnement*. PhD thesis, Universiteé de Bourgogne.

Gibson, J. J. (1966). *The Senses Considered As Perceptual Systems*. Unwin Brothers, London.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates Inc. Publishers, Hillsdale, NJ, USA, 10th (1986) edition.

Gibson, J. J. (2001). *The Essential Brunswik: Beginnings, explications, applications*, chapter Survival in a world of probably objects, pages 244–246. Oxford University Press Inc.

Gilfix, M. and Crouch, A. (2000). Peep (the network auralizaer): Analyzing your network with sound. In *LISA XIV*, pages 109–118, New Orleans, USA.

Giordano, B. (2003). *The Sounding Object*, chapter Material categorization and hardness scaling in real and synthetic impact sounds, pages 73–93. Mondo Estremo, Firenze, Italy.

Giordano, B. and McAdams, S. (2006). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *J. Acoustic Society of America*, 119(2):1171–1181.

Glinsky, A. (2000). *Theremin: Ether Music and Espionage*. University of Illinois Press.

Goldstone, R. L. (1999). *MIT Encyclopedia Of The Cognitive Sciences*, chapter Similarity, pages 763–765. MIT Press, Cambridge, MA.

Greene, J. C., Caracelli, V. J., and Graham, W. F. (1989). Towards a conceptual framework for mixed-method evaluation designs. *Education Evaluation and Policy Analysis*, 11(3):255–274.

Guyot, F. (1996). *Etude de la perception sonore en terme de reconnaissance et d'appréciation qualitative: une approache par la catégorisation*. PhD thesis, Universityé du Maine.

Gygi, B. (2001). *Factors in the Identification of Environmental Sounds*. Ph.d., Indiana University.

Gygi, B. and Shafiro, V. (2007). General functions and specific applications of environmental sound research. *Frontiers in Bioscience*, 12:3152–3166.

Hart, S. G. and Staveland, L. E. (1988). *Human mental workload*, chapter Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, pages 139–183. Elsevier Science Publishers B.V. (North-Holland), Amsterdam.

Hassenzahl, M. and Wessler, R. (2000). Capturing design space from a user perspective: the repertory grid revisited. *International Journal of Human-Computer Interaction*, 12(3 & 4):441–459.

Heft, H. (1989). Affordances and the body: An intentional analysis of gibson's ecological approach to visual perception. *Journal of the Theory of Social Behaviour*, 19(1):1–30.

Hermann, T. and Hunt, A. (2005). An introduction to interactive sonification. *IEEE Multimedia*, 12(2):20–24.

Herrington, J. (2003). *Code Generation In Action*. Manning.

Hollywood-Edge (1990). The Hollywood Edge [sound recording] : sound effects library.

Holyoak, K. J. and Mah, W. A. (1982). Cognitive reference points in judgements of symbolic magnitude. *Cognitive Psychology*, 14:328–352.

Houix, O., Lemaitre, G., Misdariis, N., and Susini, P. (2007a). Closing the loop of sound evaluation and design (closed) deliverable 4.1 everyday sound classification part 2 experimental classification of everyday sounds. (4.1 Part 2).

Houix, O., Lemaitre, G., Misdariis, N., Susini, P., Franinovic, K., Hug, D., Otten, J., Scott, J., Visell, Y., Devallez, D., Fontana, F., Papetti, S., Polotti, P., and Rocchesso, D. (2007b). Closing the loop of sound evaluation and design (closed) deliverable 4.1 everyday sound classification: Sound perception, interaction, and synthesis part 1 state of the art. (4.1 Part 1).

Howard, J. H. and Ballas, J. A. (1980). Syntactic and semantic factors in the classification of nonspeech transient patterns. *Perception and Psychophysics*, 28:431–439.

Howard, J. H. and Ballas, J. A. (1982). Acquisition of acoustic pattern categories by exemplar observation. *Organisiational Behaviour and Human Performance*, 30:157–173.

Howard, J. H. and Ballas, J. A. (1983). Perception of simulated propeller cavitation. *Human Factors*, 25(6):643–655.

Huang, E. M. and Mynatt, E. D. (2003). Semi-public displays for small, co-located groups. In *Proceedings of CHI 2003 - New Horizons*, volume 5, pages 49–56, Ft. Lauerdale, FL, USA. ACM.

Hunt, A. and Hermann, T. (2005). An introduction to interactive sonification. *IEEE Multimedia*, 12(2):20–24.

Isaacs, E., Walendowski, A., and Ranganthan, D. (2002). Hubbub: A sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. In *Proceedings of CHI 2002*, pages 179–186.

Jenkins, J. J. (1985). Acoustic information for objects, places, and events. In Warren, W. H. J. and Shaw, R. E., editors, *Persistence and Change - Proceedings of the First International Conference on Event Perception*, pages 115–138, Hillsdale. Erlbaum.

Jolliffe, I. T. (2002). *Principal Component Analysis*. Springer-Verlag, 2nd edition.

Jordà, S., Kaltenbrunner, M., Geiger, G., and Bencina, R. (2005). The reactable*. In *Proceedings of the International Computer Music Conference (ICMC 2005)*, Barcelona, Spain.

Kanwal, J. S. and Rauschecker, J. P. (2007). Auditory cortex of bats and primates: managing species-specific calls for social communication. *Frontiers in Bioscience*, 12:4621–4640.

Kawai, K., Kojima, T., Hirate, K., and Yasuoka, M. (2004). Personal evaluation structure of environmental sounds: experiments of subjective evaluation using subjects' own terms. *Journal of Sound and Vibration*, 277:523–533.

Kelly, G. A. (1955). *The Psychology of Personal Constructs*. Norton.

Kendall, G. S. (1991). Visualisation by ear: Auditory imagery for scientific visualisation and virtual reality. *Computer Music Journal*, 15(4):70–73.

Kilander, F. and Lönnqvist, P. (2002). A whisper in the woods - an ambient soundscape for peripheral awareness of remote processes. In *International Conference on Auditory Display ICAD-02*, Kyoto, Japan.

Kraut, R. E., Fish, R. S., Root, R. W., and Chalfonte, B. L. (1990). *People's Reaction to Technology*, chapter Informal comunication in organizations: Form, function, and technology, pages 145–199. Sage, Newbury Park.

Kraut, R. E. and Streeter, L. A. (1995). Coordination in software development. *Communications of the ACM*, 38(3):69–81.

Kuniacsky, M. (2003). *Observing the user experience: A practitioner's guide to user research*. Morgan Kaufmann.

Lakatos, S., Scavone, G. P., and Cook, P. R. (2000). Obtaining perceptual spaces for large numbers of complex sounds: Sensory, cognitive, and decisional constraints. In Bonnet, C., editor, *Proceedings of the Sixteenth Annual Meeting of the International Psychophysics Society*, pages 245–250.

Lakhoff, G. and Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic Books, New York.

Lemon, J., Bolker, B., Oom, S., Klein, E., Rowlingson, B., Wickham, H., Tyagi, A., Eterradossi, O., Grothendieck, G., and Toews, M. (2007). *plotrix: Various plotting functions*. R package version 2.2-9.

Luce, R. D. (1990). Psychophysical laws: cross-modal matching. *Psychological Review*, 97:66–77.

Lumsden, J. and Brewster, S. A. (2001). A survey of audio-related knowledge amongst software engineers developing human computer interfaces. Tech Report TR-2001-97, Department of Computer Science, University of Glasgow.

Macaulay, C. and Crerar, A. (1998). 'Observing' the workplace soundscape: Ethnography and auditory interface design. In *International Conference on Auditory Display ICAD-98*, University of Glasgow, Glasgow.

Maechler, M., Rousseeuw, P., Struyf, A., and Hubert, M. (2005). Cluster analysis basics and extensions. Rousseeuw et al provided the S original which has been ported to R by Kurt Hornik and has since been enhanced by Martin Maechler: speed improvements, silhouette() functionality, bug fixes, etc. See the 'Changelog' file (in the package source).

Malandrino, D., Mea, D., Negro, A., Palmieri, G., and Scarano, V. (2003). Nemos: Network monitoring with sound. In Brazil, E. and Shinn-Cunningham, B., editors, *International Conference on Auditory Display (ICAD-03)*, pages 251–254, Boston, MA, USA.

Marcell, M. M., Borella, D., Greene, M., Kerr, E., and Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clincial and Experimental Neuropsychology*, 22(6):830–864.

Marcell, M. M., Malatanos, M., Leahy, C., and Comeaux, C. (2007). Identifying, rating, and remembering environmental sound events. *Behaviour Research Methods*, 39(3):561–569.

Martens, W. L. and Sungyoung, K. (2007). Verbal elicitation and scale construction for evaluating perceptual differences between four multichannel microphone techniques. In *Proceedings of the 122nd Convention of the Audio Engineering Society*, pages 1–22, Vienna, Austria.

Martin, G. and Soren, B. (2005). Attribute identification and quantification in automotive audio - part 1: Introduction to the descriptive analysis technique. In *Proceedings of the 118th Convention of the Audio Engineering Society*, pages 1–14, Barcelona, Spain. AES.

Matheny Jr., A. P., Wachs, T. D., Ludwif, J. L., and Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Applied Development Psychology*, 16:429–444.

Mauney, B. S. and Walker, B. N. (2004). Creating functional and livable soundscapes for peripheral monitoring of dynamic data. In *International Conference on Auditory Display ICAD-04*, Sydney, Australia.

McAdam, R., Mason, B., and McCrory, J. (2007). Exploring the dichotomies within the tacit knowledge literature: towards a process of tacit knowing in organizations. *Journal of Knowledge Management*, 11(2):43–59.

McAdams, S. (1993). *Thinking in sound: The cognitive psychology of human audition*, chapter Recognition of sound sources and events, pages 146–198. Oxford, Clarendon.

McAdams, S., Winsberg, S., Donnadieu, S., Soete, G. D., and Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: common dimensions, specificities and latent subject classes. *Psychological Research*, 58:177–192.

McGookin, D. (2004). *Understanding and Improving the Identification of Concurrently Presented Earcons*. PhD thesis, University of Glasgow.

McGookin, D. and Brewster, S. A. (2004). Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Transactions on Applied Perceptions*, 1(2):130–155.

McGregor, I., Leplâtre, G., Crerar, A., and Benyon, D. (2006). Sound and soundscape classification: Establishing key auditory dimensions and their relative importance. In *International Conference on Auditory Display ICAD-06*, pages 105–112.

Melles, G. (2008). An enlarged pragmatist inquiry paradigm for methodological pluralism in academic design research. *Artifact*, 2(1):3–11.

Millen, D. R. (2000). Rapid ethnography: Time deepening strategies for hci field research. In *Proceedings of DIS2000 Designing interactive systems: processes, practices, methods, and techniques*, pages 280–286.

Mistry, P., Maes, P., and Chang, L. (2009). Wuw- wear ur world: A wearable gestural interface. In *Proceedings of CHI 2009*, pages 4111–4116.

Moore, B. C. J. and Hedwig, G. (2002). Factors influencing sequential stream segregation. *Acta Acustica united with Acustica*, 88(3):329–333.

Morse, J. M. (2003). *Handbook of Mixed Methods in Social and Behavioral Research*, chapter Principles of mixed methods and multimethod research design, pages 189–208. Sage, Thousand Oaks, California.

Mynatt, E. (1994). Designing with auditory icons. In Kramer, G. and Smith, S., editors, *Second International Conference on Auditory Display (ICAD '94)*, pages 109–119, Santa Fe, New Mexico. Santa Fe Institute.

Mynatt, E., Back, M., Want, R., Baer, M., and Ellis, J. (1998). Designing audio aura. In Karat, C., Lund, A., Coutaz, J., and Karat, J., editors, *CHI'98: Human Factors in Computing Systems*, pages 566–573, Los Angeles, CA, USA. ACM Press.

Mynatt, E. D. (1995). *Transforming Graphical Interfaces into Auditory Interfaces*. Phd thesis, Georgia Institute of Technology.

Nakagawa, S. and Waas, J. R. (2004). 'O sibling, where art thou?' - a review of avian sibling recognition with respect to the mammalian literature. *Biological Review*, 49:101–119.

Nardi, B. A., Whittaker, S., and Bradner, E. (2000). Interaction and outeraction: Instant messaging in action. In *Proceedings of CSCW'00*, pages 79–88.

Neuhoff, J. (2004). *Ecologial Psychoacoustics*, chapter Ecologial Psychoacoustics: Introduction and History, pages 1–13. Elsevier, London, 1st edition.

Nichols, J., Wobbrock, J. O., Gergle, D., and Forlizzi, J. (2002). Mediator and medium: Doors as interruption gateways and aesthetic displays. In *Proceedings of DIS2002 Designing interactive systems: processes, practices, methods, and techniques*, pages 379–386, London, England.

Nielsen, J. (1989). Usability engineering at a discount. In *International Conference on Human-Computer Interaction*, pages 394–401, Boston, MA, USA. Elsevier Science.

Noble, W. G. (1991). *Against Cognitivism: Alternative Foundations for a Cognitive Psychology*, chapter Ecological realism and the fallacy of "objectification", pages 199–223. Harvester Wheatsheaf, London.

Osgood, C. E. (1964). Semantic differential technique in the comparative study of cultures. *American Anthropologist*, 66(3):171–200.

Papp, A. L. (1997). *Presentation of Dynamically Overlapping Auditory Messages in User Interfaces*. Phd thesis, University of California.

Paradis, E., Claude, J., and Strimmer, K. (2004). Ape: analyses of phylogenetics and evolution in r language. *Bioinformatics*, 20(289-290).

Parr, J. and Neimeyer, G. J. (1994). Effects of gender, construct types, occupational information and career relevance on vocational differentiation. *J. of Counseling Psychology*, 41:27–33.

Patterson, R. (1989). Guidelines for the design of auditory warnings. *Proceedings of the Institute of Acoustics*, 11(5):17–24.

Patterson, R. D. and Datta, A. J. (1999). Extending the domain of auditory warning sounds: creative use of high frequencies and temporal assymetry. In Stanton, N. A. and Edworthy, J., editors, *Human Factors in Auditory Warnings*, pages 73–90. Ashgate, Aldershot, Hants, UK.

Pawson, M. and Greenberg, S. (2009). Extremely rapid usability testing. *J. of Usability Studies*, 4(3):124–135.

Peltola, L., Erkut, C., Cook, P. R., and Valimaki, V. (2007). Synthesis of hand clapping sounds. *IEEE Transactions on Speech and Audio Processing*, 15(3):1021–1029.

Pentland, A. (2005). Socially aware computation and communication. *IEEE Computer*, 38(3):33–40.

Pentland, A. (2007). On the collective nature of human intelligence. *Adaptive Behavior*, 15(2):189–198.

Perry, N. C., Stevens, C., Wiggins, M. W., and Howell, C. E. (2007). Cough once for danger: Icons versus abstract warnings as informative alerts in civil aviation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(11):1061–1071(11).

Peterson, P. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, 28(3):450–461.

Polanyi, M. (1966). *The Tacit Dimension*. Routledge and Kegan Paul, London.

Port, R. F., Cummins, F., and McAuley, J. D. (1995). *Mind as Motion: Exploration in the dynamics of cognition*, chapter Naive time, temporal patterns, and human audition, pages 339–371. M.I.T. Press, Cambridge.

Pousman, Z. and Stasko, J. (2006). A taxonomy of ambient information systems: Four patterns of design. In *Proceedings of AVI'06*, pages 67–74, Venezia, Italy.

Project, L. I. (2006). Pipes: A brief introduction. Webpage.

Rath, M. (2004). *Interactive Sound-model for Human-Computer-Interaction: A sound design concept and applications*. Phd, Dipartimento di Informatics, Universita di Verona.

Rath, M. and Fontana, F. (2003). *The Sounding Object*, chapter High-level models: bouncing, breaking, rolling, crumpling, pouring, pages 173–204. Mondo Estremo, Firenze, Italy.

Rath, M. and Rocchesso, D. (2005). Continuous sonic feedback from a rolling ball. *IEEE MultiMedia*, 12(2):60–69.

Richardson, M. W. (1938). Multidimensional psychophysics. *Psychological Bulletin*, (35):659.

Ringel Morris, M., Morris, D., and Winograd, T. (2004). Individual audio channels with single display groupware: Effects on communication and task strategy. In *Proceedings of CSCW'04*, pages 242–251.

Roberts, E., Bornstein, M. H., Slater, A. M., and Barrett, J. (1999). Early cognitive development and parental education. *Infant and Child Development*, 8:49–62.

Rocchesso, D. (2004). Physically-based sounding objects, as we develop them today. *Journal of New Music Research*, 33(3):305–313.

Rocchesso, D., Bresin, R., and Fernström, M. (2003). Sounding objects. *IEEE Multimedia*, 10(2):42–52.

Rocchesso, D. and Fontana, F., editors (2003). *The Sounding Object*. Mondo Estremo, Firenze, Italy.

Rosch, E. (1975). Family resemblance: Studies in the internal structure of categories. *Cognitive Psychology*, (7):573–605.

Rosch, E. (1977). *Study in cross-cultural psychology*, chapter Human categorisation, pages 1–49. Academic Press, London.

Rosch, E. (1978). *Cognition and Categorisation*, chapter Principles of Categorisation, pages 27–48. Lawrence Erlbaum Associates, Hillsdale, NY.

Ross, B. and Murphy, G. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology*, 38(4):495–553.

Runeson, S. (1989). A note on the utility of ecologically incomplete invariants. *Newsletter of the International Society for Ecological Psychology*, 4(1):6–8.

Runeson, S. and Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender, recognition, and deceptive intention. *Journal of Experimental Psychology*, (112):585–615.

Saary, J. M. (2008). Radar plots: a useful way for presenting multivariate health care data. *Journal of Clinical Epidemiology*, 60:311–317.

Sanders, E. B.-N. (2005). Information, inspiration and co-creation. In *Proceedings of the 6th Conference of the European Academy of Design*, University of Arts, Bremen, Germany.

Sanders, J. T. (1997). An ontology of affordances. *Ecological Psychology*, 9(1):97–112.

Scavone, G. P., Lakatos, S., and Cook, P. R. (2000). Knowledge acquisition by listeners in a source learning task using physical models. *J. Acoustic Society of America*, 107(5):2817–2818.

Schafer, R. M. (1977). *The soundscape- Our sonic environment and the tuning of the world*. Destiny Books, Rochester, Vermont, USA.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3):515–530.

Shafiro, V. and Gygi, B. (2004). How to select stimuli for environmental sound research and where to find them? *Behaviour Research Methods, Instruments, and Computers*, 36(4):590–598.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423.

Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 39:373–422.

Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, 210:390–398.

Shepard, R. N. (1991). *The Perception Of Structure: Essays in Honour of Wendell R. Garner*, chapter Integrality versus separability of stimulus dimensions: From an early convergence of evidence to a proposed theoretical basis, pages 53–71. American Psychological Association, Washington, D.C.

Shinn-Cunningham, B. and Ihlefeld, A. (2004). Selective and divided attention: Extracting information from simultaneous sound sources. In *International Conference on Auditory Display (ICAD-04)*, Sydney, Australia.

Shinn-Cunningham, B. G., Ihlefeld, A., Satyavarta, and Larson, E. (2005). Bottom-up and top-down influences on spatial unmasking. *Acta Acustica united with Acustica*, 91(6):967–979.

Shinn-Cunningham, B. G., Lee, A. K. C., and Oxenham, A. J. (2007). A sound element gets lost in perceptual competition. *Proceedings of the National Academy of Sciences*, 104(29):12223–12227.

Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. *IEEE, Computer*, 16(8):57–69. direct manipulation.

Shneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley, Reading, MA, USA, 2 edition.

Simpson, B. D., Brungart, D. S., Gilkey, R. H., Iyer, N., and Hamil, J. T. (2007). Localization in multiple source environments: Localizing the missing source. In Scavone, G. P., editor, *International Conference on Auditory Display (ICAD-07)*, pages 280–284, Montreal, Canada.

Slater, M. (2004). How colorful was your day? why questionnaires cannot assess presence in virtual environments. *Presence*, 13(4):484–493.

Smith, R. (1988). A prototype futuristic technology for distance education. In *Proceedings of the NATO Advanced Workshop on New Directions in Educational Technology*, Cranfield, UK.

Stanton, N. and Edworthy, J., editors (1999). *Human Factors in Auditory Warnings*. Ashgate, Aldershot, Hants, UK.

Stevens, S. S. (1974). Perceptual magnitude and its measurement. In Carterette, E. C. and Friedman, M. P., editors, *Handbook of perception, II*, volume 2, pages 361–389. Academic Press, New York.

Stewart, J., Benderson, B. B., and Druin, A. (1999). Single display groupware: A model for co-present collaboration. In *CHI'99: Human Factors in Computing Systems*, pages 286–293, Pittsburgh, PA, USA. ACM.

Tardieu, J., Susini, P., Poisson, F., Lazareff, P., and McAdams, S. (2008). Perceptual study of soundscapes in train stations. *Applied Acoustics*, 69(12):1224–1239.

Team, R. D. C. (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Tomico, O. P. (2006). Experience landscapes. In *Proceedings of DESIGN 2006 Conference*, pages 393–400, Dubrovnik, Croatia.

Tomico, O. P. (2007). *Subjective experience gathering techniques for interaction design: subjective psychological exploration techniques based in the constructivism paradigm for informational and inspirational purposes*. PhD thesis, Universitat Politecnica De Catalunya.

Tridgell, A. (1999). *Efficient Algorithms for Sorting and Synchronization*. PhD thesis, The Australian National University.

Turner, P., McGregor, I., Turner, S., and Carroll, F. (2003). Evaluating sounscapes as a means of creating a sense of place. In Brazil, E. and Shinn-Cunningham, B., editors, *International Conference on Auditory Display ICAD-03*, pages 148–151, Boston, MA, USA.

van der Kouwe, A. J. W., Wang, D., and Brown, G. (2001). A comparison of auditory and blind separation techniques for speech segregation. *IEEE Transactions on Speech and Audio Processing*, 9(3):189–195.

van Noorden, L. (1975). *Temporal Coherence in the Perception of Tone Sequences*. Phd thesis, University of Eindhoven.

Vanderveer, N. J. (1979). *Ecological Acoustics- Human Perception of Environmental Sounds*. Phd, University Of Cornell.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press, Boston.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, fourth edition. ISBN 0-387-95457-0.

Walker, B. and Kramer, G. (2004). *Ecological Psychoacoustics*, chapter Ecological Psychoacoustics and Auditory Displays: Hearing, Grouping and Meaning Making, pages 149–174. Elsevier, London, 1st edition.

Walker, B. N., Nance, A., and Lindsay, J. (2006). Spearcons: Speech-based earcons improve navigation performance in auditory menus. In *Proceedings of ICAD 2006 - The 12th International Conference on Auditory Display*, pages 63–68.

Warren, W. and Verbrugge, R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. *Journal of Experimental Psychology: Human Perception and Performance*, 10:704–712.

Wickens, C. and Hollands, J. (2000). *Engineering Psychology and Human Performance*. Prentice Hall, New Jersey, USA, third edition edition.

Wickramasuriya, J., Vasudevan, V., and Narasimhan, N. (2007). Interstitial interfaces for mobile media. In Hazlewood, W. R., Coyle, L., and Consolvo, S., editors, *Proceedings of 1st Workshop on Ambient Information Systems. Colocated with Pervasive 2007*, volume 254 of *CEUR Workshop Proceedings ISSN 1613-0073*, Toronto, Canada.

Wightman, F. L. and Jenison, R. (1995). *Perception of Space and Motion*, chapter Auditory spatial layout, pages 365–397. Academic Press, San Diego.

Wikipedia (2007). Pipeline (unix). Webpage.

Williamson, J., Murray-Smith, R., and Hughes, S. (2007). Shoogle: excitatory multimodal interaction on mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 121–124, San Jose, California, USA. ACM Press.

Windsor, L. W. (1997). Frequency structure in electroacoustic music: ideology, function and perception. *Organized Sound*, 2(2):77–82.

Windsor, L. W. (2000). *Music, Electronic Media and Culture*, chapter Through and around the acousmatic, pages 7–35. Ashgate, Aldershot.

Windsor, L. W. (2004). An ecological approach to semiotics. *Journal for the Theory of Social Behaviour*, 34(2):179–198.

Wisneski, C., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B., and Yarin, P. (1998). Ambient displays: Turning architecture space into an interface between people and digitial information. In *Proceedings of the First International Workshop on Cooperative Buildings (CoBuild '98)*, pages 1–11. Springer-Verlag.

Wrigley, S. and Brown, G. (2004). A computational model of auditory selective attention. *IEEE Transactions on Neural Networks*, 15(5):1151–1163.

www.freesound.org (2007). Freesound creative commons: sound effects library.

Yorke, D. M. (1983). Straight or bent? an inquiry into rating scales in repertory grids. *British Educational Research Journal*, 9(2):141–151.

Yorke, D. M. (2001). Bipolarity or not ? some conceptual problems relating to bipolar rating scales. *British Educational Research Journal*, 27(2):171–186.

Zacharov, N. and Lorho, G. (2005). Sensory analysis of sound (in telecommunications). In *Proc. of European Sensory Network Conference 05*, pages 1–6, Madrid, Spain.

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

Zheng, C. and James, D. L. (2009). Harmonic fluids. *ACM Transactions on Graphics (SIGGRAPH 2009)*, 28(3):To Appear.

Zuberbühler, K. (2000). Interspecies semantic communication in two forest primates. *Royal Society of London*, 267:713–718.

Zwicker, E. and Fastl, H. (1990). *Psychoacoustics: Facts and Models*. Springer-Verlag, Heidelberg, Germany.

Part IV

Appendices

The part of the thesis provides the various appendices referred to in the main body of the thesis.

A.1 Appendix A - Chapter 4 - An Introduction to the Sonic Browser

This appendix presents an introduction to the Sonic Browser (Brazil, 2003) which was used as the experimental platform in the study presented in Chapter 4. In the Sonic Browser, a central idea is to map sound clips to aural and/or visual objects with properties that convey information about the sound clips and use the objects in order to create browsing spaces. The foundations of the design approach for the Sonic Browser are based on the principles of direct manipulation and interactive visualisation interfaces proposed by Shneiderman (1992). The three primary facets of this foundation are *overview first, zoom and filter, then detail on demand*. The interface was design to exploit the "*cocktail party effect*" (Cherry, 1953, Cherry and Taylor, 1954) and an "*aura*" mechanism (Benford and Greenhalgh, 1997). The cocktail party effect highlights a listener's ability to selectively attend to multiple different simultaneously occurring streams or sounds. The aura concept uses the ability of a listener to switch at will their auditory range of perception or the sound being focused upon (Wickens and Hollands, 2000). The aura mechanism implemented in the Sonic Browser allows for the browsing of sound files which were represented by visual icons under a circular cursor. The cursor is user controller and dynamically resizable. The resizable idea of the cursor is used to represent the listener's range of perception and allows for a zooming or broadening of interest by changing the size of the cursor which represents the aura. The sound feedback in the Sonic Browser used multiple stream stereo-spatialised audio playback controlled by the cursor/aura-over-icons. The sounds in the Sonic Browser were represented by icons whose playback was triggered when the aura/cursor moved over them. A limited set of the features of the Sonic Browser was used for the experiments as shown in Figure A-1 as compared to Figure A-2. A more detailed review and discussion of the Sonic Browser and its features can be found in Brazil (2003).

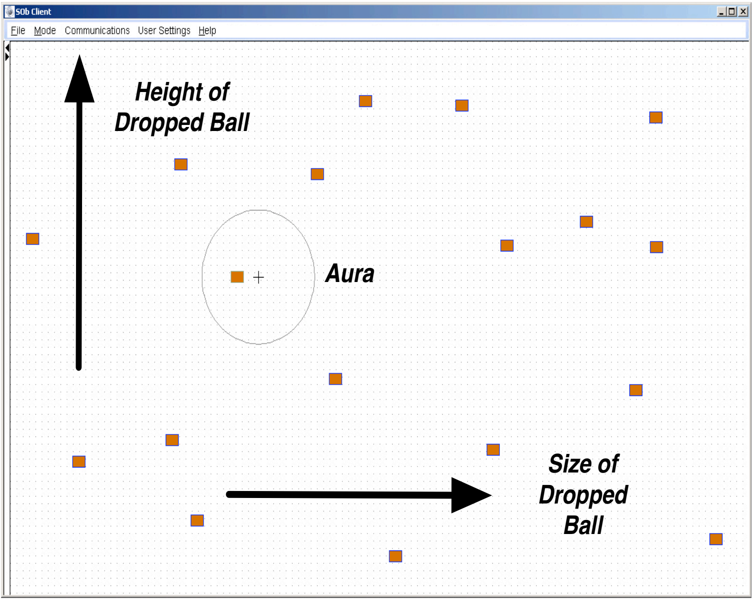


Fig. A-1: *The Sonic Browser interface (Brazil, 2003) as presented to participants in the experiment in Chapter 4.*

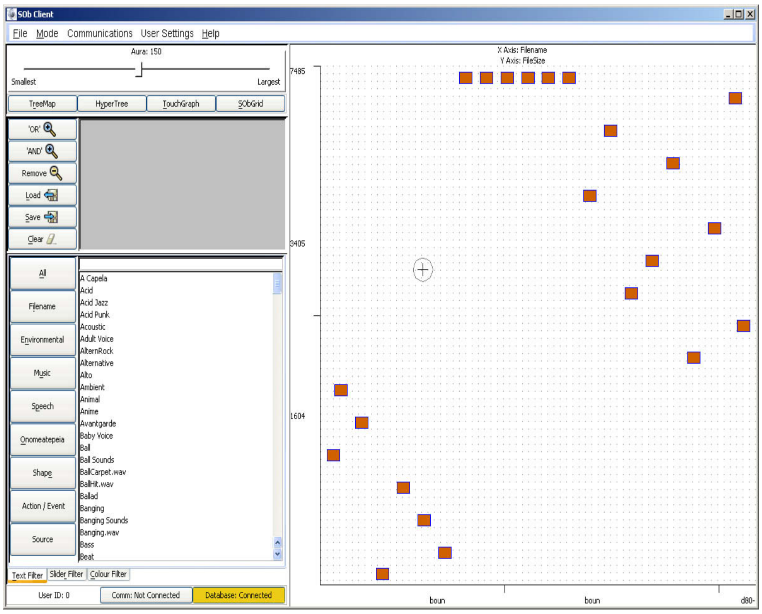


Fig. A-2: *A more complete view of the Sonic Browser interface (Brazil, 2003) and its functionality.*

B.1 Appendix B - Chapter 4 - Sounds and Participants Detailed Results

This appendix presents the detailed results from participants and the sounds in Chapter 4.

First Probe Results The individual participant's scaling and "tagging" information as shown in Figure B-1. Two participants (participant two and participant three) made a particular distinction between the real and synthetic sounds. This could indicate that other factors are affecting their scaling judgements. This could be anything from the realism to the recording conditions in the real sounds which are not present in the synthesised sounds. This could be related to the lack of acoustical richness inherent in many synthetic sounds as suggested by Carello et al. (2003).

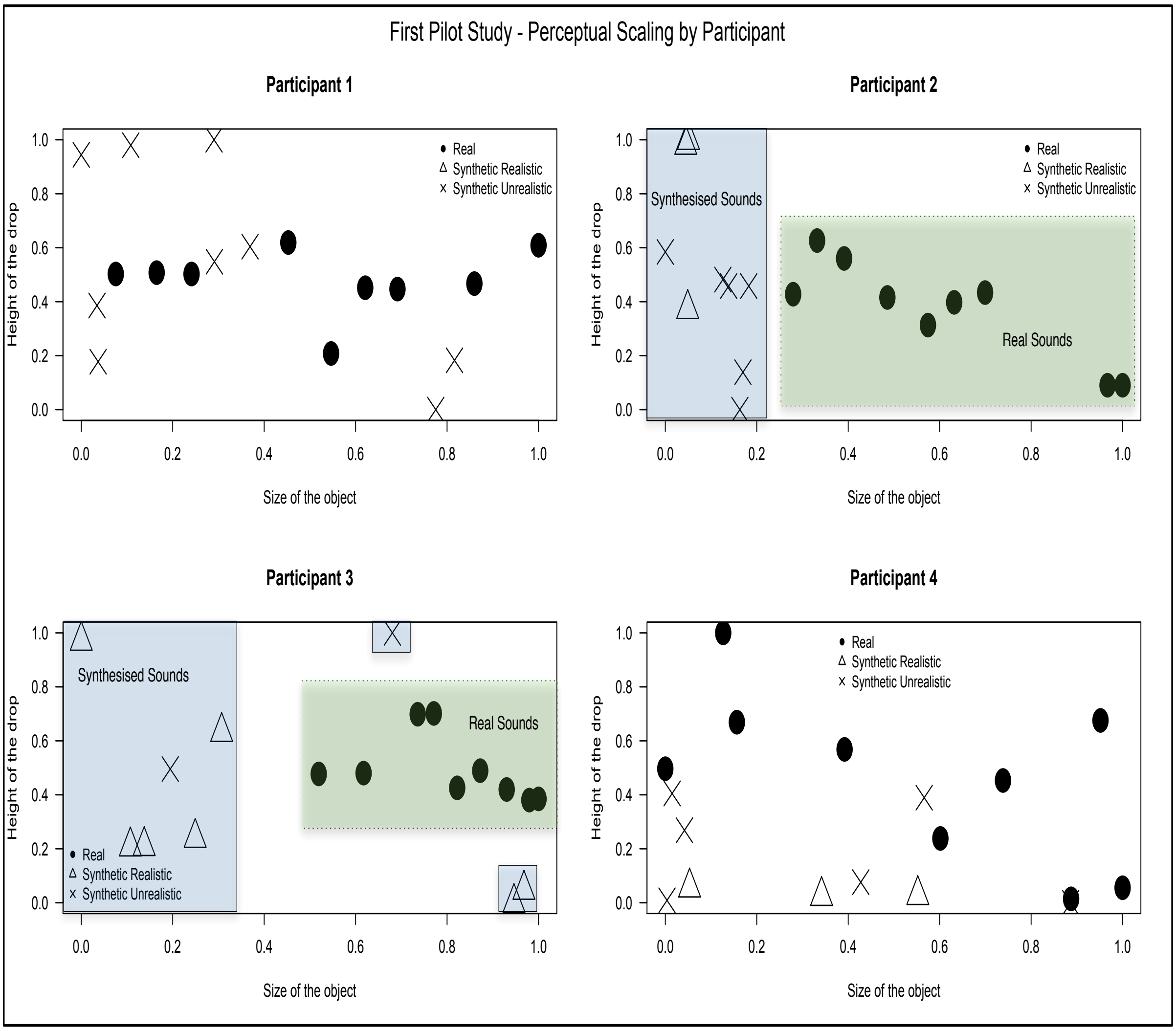


Fig. B-1: *The perceptual scaling and "tagging" information for each of the participants in the first probe.*

The individual scaling of the stimuli and the ranges of scaling are shown in Figure B-2 for the perceptual scaling of size and of height.

Second Pilot Results The perceptual scaling and "tagging" information sorted by participant for the second pilot study is shown in Figure B-3. The first three participants show clear distinctions in scaling between real and synthesised sounds. However, only participant three limited the scaling of synthesised sounds, where they alone felt that all the drops in the synthesised sounds were small. In a similar manner to the first pilot probe, the classification made by participants was grouped according to the type of sound. In particular, two participants (participant 1 and participant 2) only made minor judgements on the size of the real sounds and their judgements focused on the height aspect of the sounds

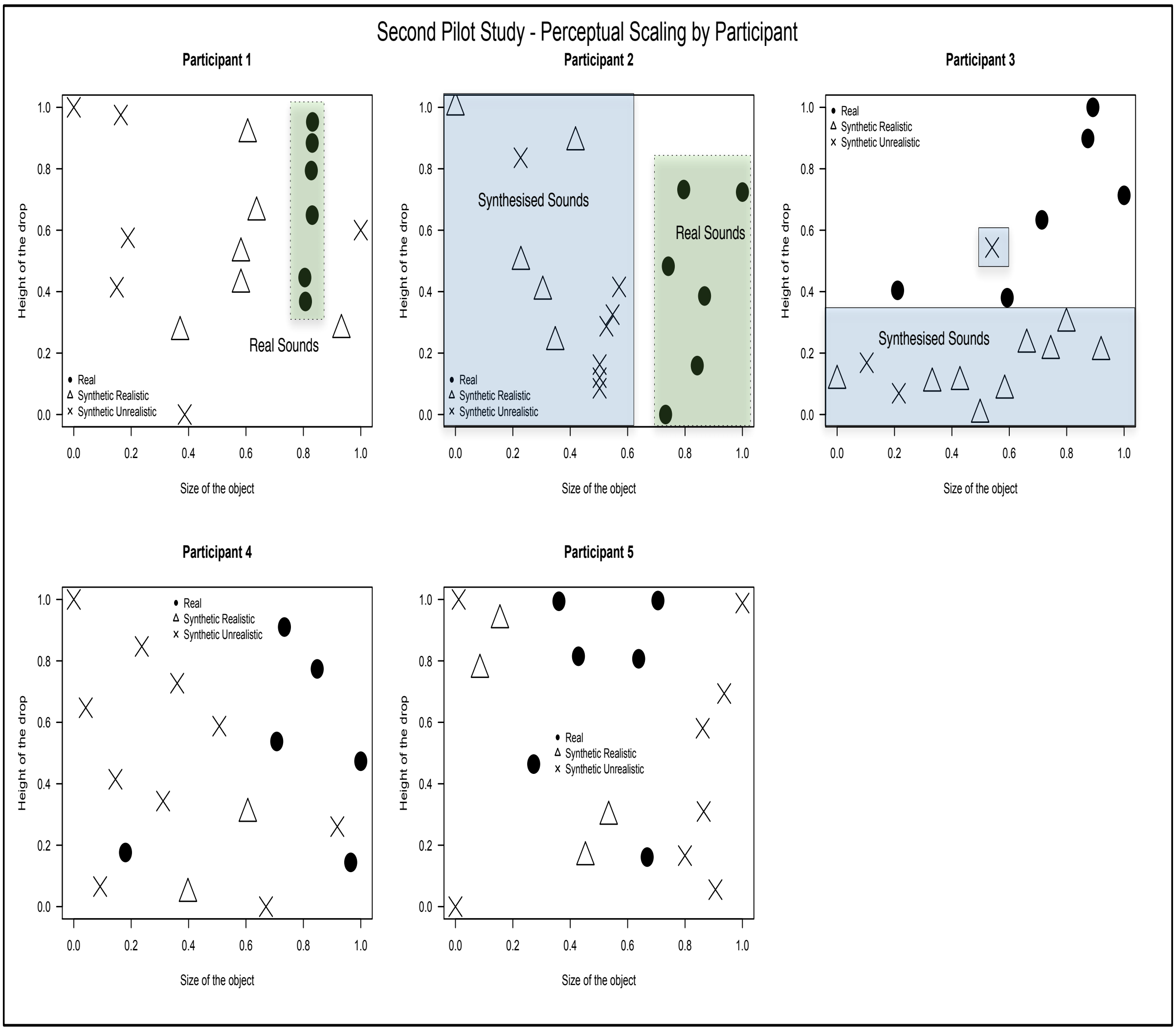


Fig. B-3: *The perceptual scaling and "tagging" information sorted by participant for the second pilot study.*

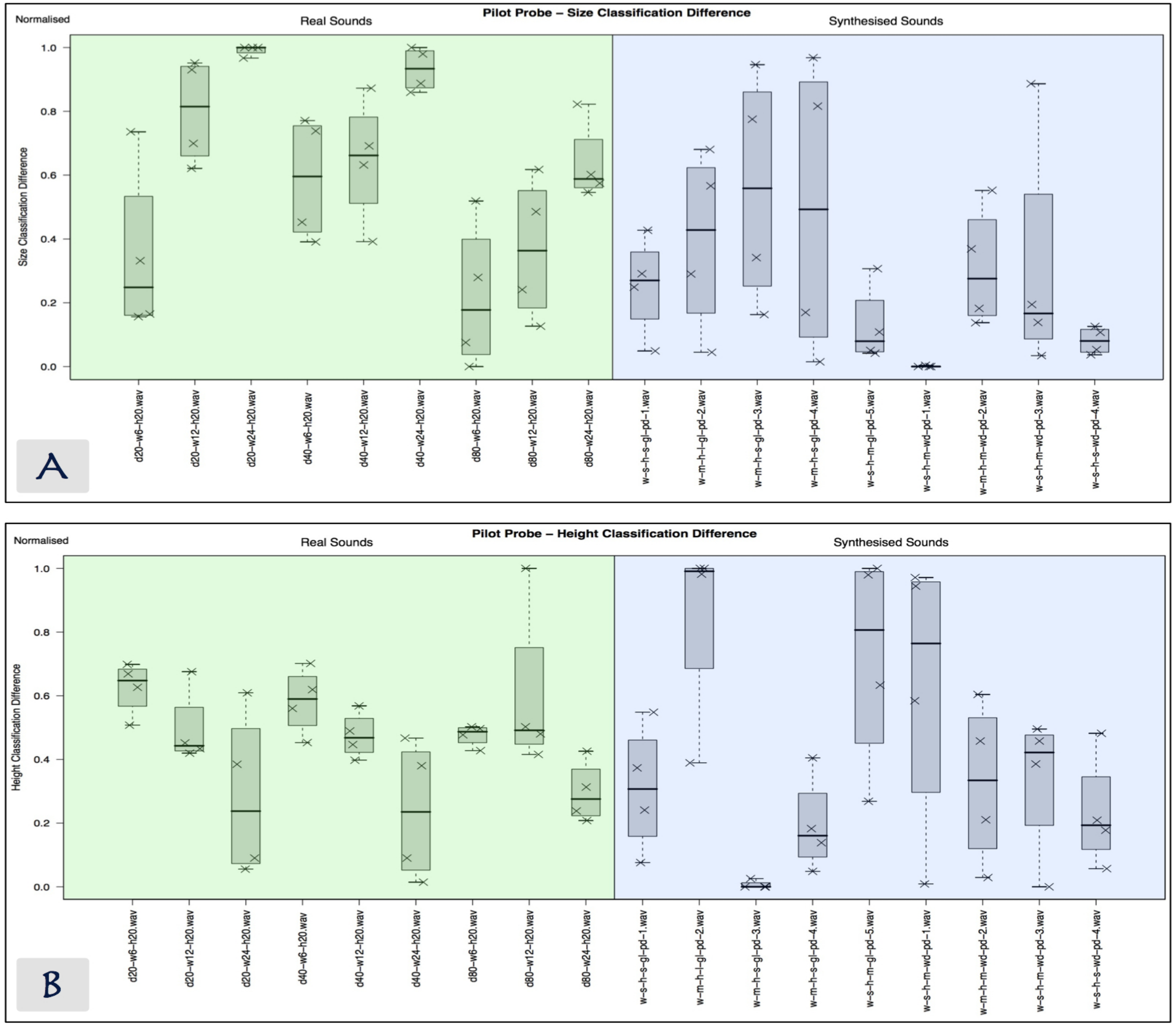


Fig. B-2: *The perceptual scaling by size (A) and by height (B) for each of the stimuli in the first probe.*

Participant ID	Filename	Normalised X	Normalised Y	Realistic	Synthetic
1	d20-w6-h20.wav	0.164889253	0.507614213	Y	N
2	d20-w6-h20.wav	0.331925676	0.626506024	Y	N
3	d20-w6-h20.wav	0.735738255	0.698484848	Y	N
4	d20-w6-h20.wav	0.156438676	0.668903803	Y	N
1	d20-w12-h20.wav	0.62100082	0.45177665	Y	N
2	d20-w12-h20.wav	0.699324324	0.43373494	Y	N
3	d20-w12-h20.wav	0.930369128	0.41969697	Y	N
4	d20-w12-h20.wav	0.951668613	0.675615213	Y	N
1	d20-w24-h20.wav	1	0.609137056	Y	N
2	d20-w24-h20.wav	0.967060811	0.090361446	Y	N
3	d20-w24-h20.wav	1	0.384848485	Y	N
4	d20-w24-h20.wav	1	0.055538064	Y	N
1	d40-w6-h20.wav	0.452830189	0.61928934	Y	N
2	d40-w6-h20.wav	0.391047297	0.560240964	Y	N
3	d40-w6-h20.wav	0.770973154	0.701515152	Y	N
4	d40-w6-h20.wav	0.738427798	0.453020134	Y	N
1	d40-w12-h20.wav	0.691550451	0.446700508	Y	N
2	d40-w12-h20.wav	0.631756757	0.397590361	Y	N
3	d40-w12-h20.wav	0.872483221	0.489393939	Y	N
4	d40-w12-h20.wav	0.392027873	0.568232662	Y	N
1	d40-w24-h20.wav	0.859721083	0.467005076	Y	N
2	d40-w24-h20.wav	1	0.090361446	Y	N
3	d40-w24-h20.wav	0.979865772	0.38030303	Y	N
4	d40-w24-h20.wav	0.887417014	0.014541387	Y	N
1	d80-w6-h20.wav	0.075471698	0.502538071	Y	N
2	d80-w6-h20.wav	0.279560811	0.427710843	Y	N
3	d80-w6-h20.wav	0.519295302	0.477272727	Y	N
4	d80-w6-h20.wav	0	0.496644295	Y	N
1	d80-w12-h20.wav	0.241181296	0.502538071	Y	N
2	d80-w12-h20.wav	0.485641892	0.415662651	Y	N
3	d80-w12-h20.wav	0.617449664	0.48030303	Y	N
4	d80-w12-h20.wav	0.126640833	1	Y	N
1	d80-w24-h20.wav	0.546349467	0.208121827	Y	N
2	d80-w24-h20.wav	0.574324324	0.313253012	Y	N
3	d80-w24-h20.wav	0.822147651	0.425757576	Y	N
4	d80-w24-h20.wav	0.601543957	0.238255034	Y	N

Table B-1*:* The raw data (1 of 2) from participants scalings and judgements in the first pilot study for the real sounds.

Participant ID	Filename	Normalised X	Normalised Y	Realistic	Synthetic
1	w-s-h-s-gl-pd-1.wav	0.291222313	0.54822335	N	Y
2	w-s-h-s-gl-pd-1.wav	0.048986486	0.373493976	Y	Y
3	w-s-h-s-gl-pd-1.wav	0.249161074	0.240909091	Y	Y
4	w-s-h-s-gl-pd-1.wav	0.427412811	0.07606264	N	Y
1	w-m-h-l-gl-pd-2.wav	0.290401969	1	N	Y
2	w-m-h-l-gl-pd-2.wav	0.044763514	0.981927711	Y	Y
3	w-m-h-l-gl-pd-2.wav	0.680369128	1	N	Y
4	w-m-h-l-gl-pd-2.wav	0.566159018	0.389261745	N	Y
1	w-m-h-s-gl-pd-3.wav	0.775225595	0	N	Y
2	w-m-h-s-gl-pd-3.wav	0.163006757	0	N	Y
3	w-m-h-s-gl-pd-3.wav	0.946308725	0	Y	Y
4	w-m-h-s-gl-pd-3.wav	0.341744013	0.025727069	Y	Y
1	w-m-h-s-gl-pd-4.wav	0.816242822	0.182741117	N	Y
2	w-m-h-s-gl-pd-4.wav	0.169763514	0.138554217	N	Y
3	w-m-h-s-gl-pd-4.wav	0.968120805	0.048484848	Y	Y
4	w-m-h-s-gl-pd-4.wav	0.014898922	0.4049217	N	Y
1	w-s-h-m-gl-pd-5.wav	0.10828548	0.979695431	N	Y
2	w-s-h-m-gl-pd-5.wav	0.050675676	1	Y	Y
3	w-s-h-m-gl-pd-5.wav	0.30704698	0.633333333	Y	Y
4	w-s-h-m-gl-pd-5.wav	0.041903217	0.268456376	N	Y
1	w-s-h-m-wd-pd-1.wav	0	0.944162437	N	Y
2	w-s-h-m-wd-pd-1.wav	0	0.584337349	N	Y
3	w-s-h-m-wd-pd-1.wav	0	0.971212121	Y	Y
4	w-s-h-m-wd-pd-1.wav	0.00372473	0.008948546	N	Y
1	w-m-h-m-wd-pd-2.wav	0.369155045	0.604060914	N	Y
2	w-m-h-m-wd-pd-2.wav	0.182432432	0.457831325	N	Y
3	w-m-h-m-wd-pd-2.wav	0.137583893	0.210606061	Y	Y
4	w-m-h-m-wd-pd-2.wav	0.552191279	0.029082774	Y	Y
1	w-s-h-m-wd-pd-3.wav	0.034454471	0.385786802	N	Y
2	w-s-h-m-wd-pd-3.wav	0.138513514	0.457831325	N	Y
3	w-s-h-m-wd-pd-3.wav	0.194630872	0.495454545	N	Y
4	w-s-h-m-wd-pd-3.wav	0.886485831	0	N	Y
1	w-s-h-s-wd-pd-4.wav	0.036915505	0.177664975	N	Y
2	w-s-h-s-wd-pd-4.wav	0.125844595	0.481927711	N	Y
3	w-s-h-s-wd-pd-4.wav	0.10738255	0.209090909	Y	Y
4	w-s-h-s-wd-pd-4.wav	0.053077408	0.05704698	Y	Y

Table B-2*:* The raw data (2 of 2) from participants scalings and judgements in the first pilot study for the synthesised sounds.

The synthesised sounds, the individual scaling of the stimuli, and the ranges of scaling for these stimuli are shown in Figure B-4. In a similar result to the first pilot study, the users did agree in at least one of the perceptual scaling dimensions. The maximum consensus amongst participants was by 3 participants. Similarly to the first pilot study, all of the real sounds were judged to be realistic.

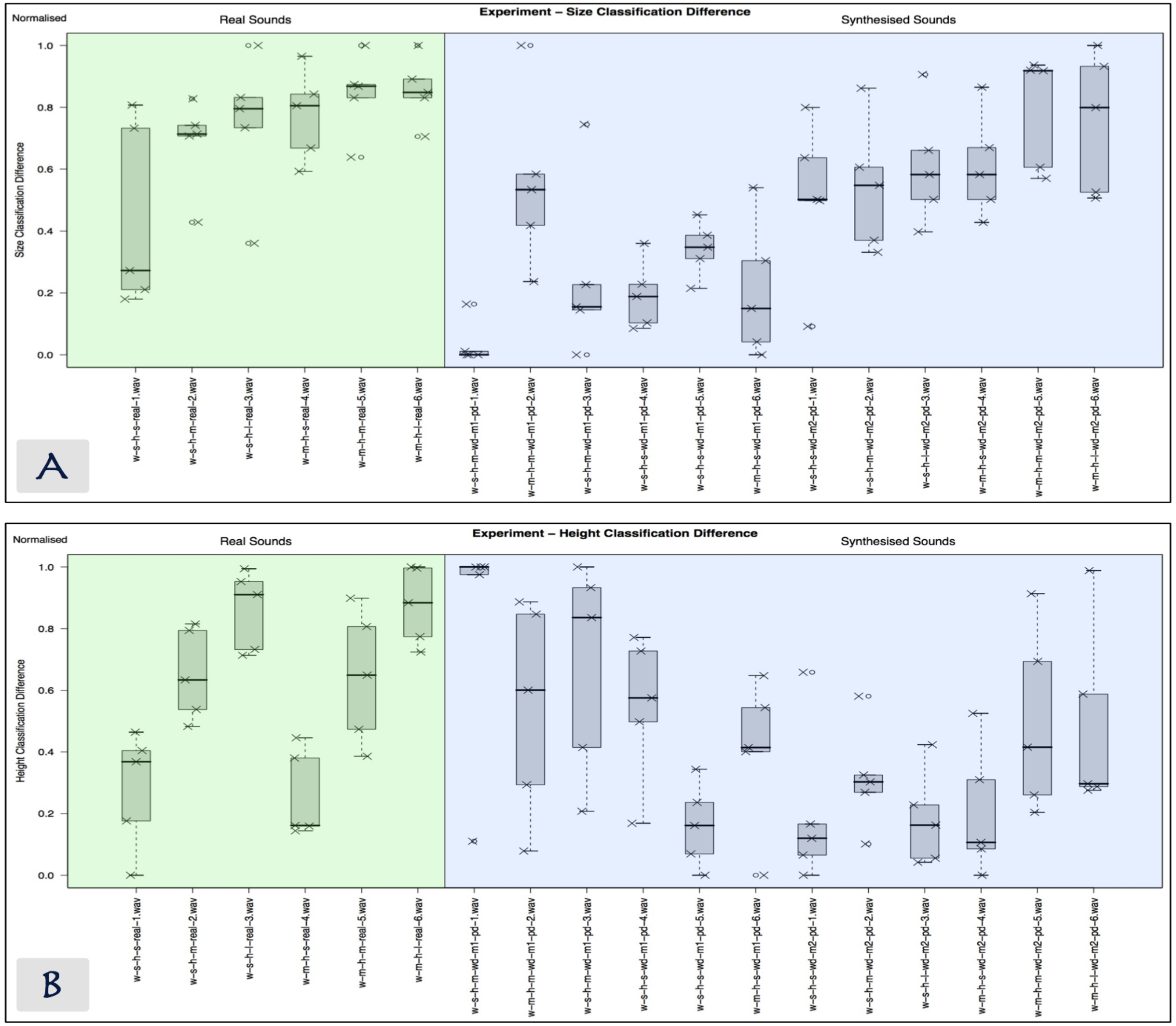


Fig. B-4: *The perceptual scaling by size (A) and by height (B) for each of the stimuli (see Table* **??** *for stimuli details) in the second pilot study.*

Participant ID	Filename	Normalised X	Normalised Y	Realistic	Synthetic
1	w-s-h-s-real-1.wav	0.8072805	0.3680739	Y	N
2	w-s-h-s-real-1.wav	0.732327992	0	Y	N
3	w-s-h-s-real-1.wav	0.210757409	0.40430925	Y	N
4	w-s-h-s-real-1.wav	0.180150125	0.1762749	Y	N
5	w-s-h-s-real-1.wav	0.272648084	0.46407538	Y	N
1	w-s-h-m-real-2.wav	0.8276231	0.7941953	Y	N
2	w-s-h-m-real-2.wav	0.741753063	0.48275862	Y	N
3	w-s-h-m-real-2.wav	0.713501647	0.63371356	Y	N
4	w-s-h-m-real-2.wav	0.707256047	0.537694	Y	N
5	w-s-h-m-real-2.wav	0.428571429	0.81507656	Y	N
1	w-s-h-l-real-3.wav	0.8319058	0.9525066	Y	N
2	w-s-h-l-real-3.wav	0.795475966	0.73234811	Y	N
3	w-s-h-l-real-3.wav	1	0.71356147	Y	N
4	w-s-h-l-real-3.wav	0.733944954	0.9101996	Y	N
5	w-s-h-l-real-3.wav	0.360627178	0.99411072	Y	N
1	w-m-h-s-real-4.wav	0.8051392	0.4459103	Y	N
2	w-m-h-s-real-4.wav	0.84260132	0.1592775	Y	N
3	w-m-h-s-real-4.wav	0.592755214	0.38022814	Y	N
4	w-m-h-s-real-4.wav	0.964970809	0.1441242	Y	N
5	w-m-h-s-real-4.wav	0.668118467	0.16136631	Y	N
1	w-m-h-m-real-5.wav	0.8308351	0.6490765	Y	N
2	w-m-h-m-real-5.wav	0.86804901	0.38587849	Y	N
3	w-m-h-m-real-5.wav	0.873765093	0.89860583	Y	N
4	w-m-h-m-real-5.wav	1	0.4733925	Y	N
5	w-m-h-m-real-5.wav	0.638501742	0.80683157	Y	N
1	w-m-h-l-real-6.wav	0.8308351	0.883905	Y	N
2	w-m-h-l-real-6.wav	1	0.72413793	Y	N
3	w-m-h-l-real-6.wav	0.891328211	1	Y	N
4	w-m-h-l-real-6.wav	0.848206839	0.7738359	Y	N
5	w-m-h-l-real-6.wav	0.705574913	0.99646643	Y	N

Table B-3*:* The raw data (1 of 2) from participants scalings and judgements in the second pilot study for the real sounds.

Participant ID	Filename	Normalised X	Normalised Y	Realistic	Synthetic
1	w-s-h-m-wd-m1-pd-1.wav	0.1638116	0.974934	N	Y
2	w-s-h-m-wd-m1-pd-1.wav	0	1	Y	Y
3	w-s-h-m-wd-m1-pd-1.wav	0	0.11026616	Y	Y
4	w-s-h-m-wd-m1-pd-1.wav	0	1	N	Y
5	w-s-h-m-wd-m1-pd-1.wav	0.011324042	1	N	Y
1	w-m-h-m-wd-m1-pd-2.wav	1	0.6002639	N	Y
2	w-m-h-m-wd-m1-pd-2.wav	0.418473139	0.88669951	Y	Y
3	w-m-h-m-wd-m1-pd-2.wav	0.583973655	0.07858048	Y	Y
4	w-m-h-m-wd-m1-pd-2.wav	0.236864053	0.8470067	N	Y
5	w-m-h-m-wd-m1-pd-2.wav	0.533972125	0.29328622	Y	Y
1	w-s-h-m-wd-m1-pd-3.wav	0	1	N	Y
2	w-s-h-m-wd-m1-pd-3.wav	0.227144204	0.83579639	N	Y
3	w-s-h-m-wd-m1-pd-3.wav	0.744422187	0.2075929	Y	Y
4	w-s-h-m-wd-m1-pd-3.wav	0.145120934	0.4146341	N	Y
5	w-s-h-m-wd-m1-pd-3.wav	0.155052265	0.93286219	Y	Y
1	w-s-h-s-wd-m1-pd-4.wav	0.1884368	0.5751979	N	Y
2	w-s-h-s-wd-m1-pd-4.wav	0.228086711	0.49753695	Y	Y
3	w-s-h-s-wd-m1-pd-4.wav	0.103183315	0.16856781	N	Y
4	w-s-h-s-wd-m1-pd-4.wav	0.36030025	0.7272727	N	Y
5	w-s-h-s-wd-m1-pd-4.wav	0.085365854	0.77149588	Y	Y
1	w-s-h-s-wd-m1-pd-5.wav	0.3865096	0	N	Y
2	w-s-h-s-wd-m1-pd-5.wav	0.347785108	0.2364532	Y	Y
3	w-s-h-s-wd-m1-pd-5.wav	0.215017315	0.06915632	N	Y
4	w-s-h-s-wd-m1-pd-5.wav	0.311092577	0.3436807	N	Y
5	w-s-h-s-wd-m1-pd-5.wav	0.452961672	0.16136631	Y	Y
1	w-m-h-s-wd-m1-pd-6.wav	0.1498929	0.414248	N	Y
2	w-m-h-s-wd-m1-pd-6.wav	0.304429783	0.40065681	Y	Y
3	w-m-h-s-wd-m1-pd-6.wav	0.540065862	0.54372624	N	Y
4	w-m-h-s-wd-m1-pd-6.wav	0.041701418	0.6474501	N	Y
5	w-m-h-s-wd-m1-pd-6.wav	0	0	N	Y
1	w-s-h-s-wd-m2-pd-1.wav	0.637045	0.6583113	Y	Y
2	w-s-h-s-wd-m2-pd-1.wav	0.502356268	0.11986864	N	Y
3	w-s-h-s-wd-m2-pd-1.wav	0.498353458	0	Y	Y
4	w-s-h-s-wd-m2-pd-1.wav	0.091743119	0.0654102	N	Y
5	w-s-h-s-wd-m2-pd-1.wav	0.799651568	0.16607774	N	Y
1	w-s-h-m-wd-m2-pd-2.wav	0.3704497	0.2691293	Y	Y
2	w-s-h-m-wd-m2-pd-2.wav	0.547596607	0.32512315	N	Y
3	w-s-h-m-wd-m2-pd-2.wav	0.331503842	0.10139417	Y	Y
4	w-s-h-m-wd-m2-pd-2.wav	0.606338616	0.3026608	Y	Y
5	w-s-h-m-wd-m2-pd-2.wav	0.861498258	0.58068316	N	Y
1	w-s-h-l-wd-m2-pd-3.wav	0.5824411	0.4234828	Y	Y
2	w-s-h-l-wd-m2-pd-3.wav	0.502356268	0.16256158	N	Y
3	w-s-h-l-wd-m2-pd-3.wav	0.660812294	0.22813688	Y	Y
4	w-s-h-l-wd-m2-pd-3.wav	0.397831526	0.0421286	Y	Y
5	w-s-h-l-wd-m2-pd-3.wav	0.905923345	0.05535925	N	Y
1	w-m-h-s-wd-m2-pd-4.wav	0.5824411	0.525066	Y	Y
2	w-m-h-s-wd-m2-pd-4.wav	0.502356268	0.08538588	N	Y
3	w-m-h-s-wd-m2-pd-4.wav	0.428100988	0.10646388	Y	Y
4	w-m-h-s-wd-m2-pd-4.wav	0.669724771	0	N	Y
5	w-m-h-s-wd-m2-pd-4.wav	0.864982578	0.30977621	N	Y
1	w-m-h-m-wd-m2-pd-5.wav	0.6059957	0.9129288	Y	Y
2	w-m-h-m-wd-m2-pd-5.wav	0.570216777	0.41543514	N	Y
3	w-m-h-m-wd-m2-pd-5.wav	0.919868277	0.20405577	Y	Y
4	w-m-h-m-wd-m2-pd-5.wav	0.918265221	0.2605322	N	Y
5	w-m-h-m-wd-m2-pd-5.wav	0.93641115	0.69375736	N	Y
1	w-m-h-l-wd-m2-pd-6.wav	0.9325482	0.2757256	Y	Y
2	w-m-h-l-wd-m2-pd-6.wav	0.525918944	0.28735632	N	Y
3	w-m-h-l-wd-m2-pd-6.wav	0.799121844	0.29657795	Y	Y
4	w-m-h-l-wd-m2-pd-6.wav	0.507089241	0.5875831	N	Y
5	w-m-h-l-wd-m2-pd-6.wav	1	0.98822144	N	Y

Table B-4: The raw data (2 of 2) from participants scalings and judgements in the second pilot study for the synthesised sounds.

C.1 Appendix C - Chapter 4 - Synthesised sounds parameters

This appendix presents the values used with the sound object models to create the synthesised sounds used in Chapter 4. The impact model used to create these sounds with its main features is discussed in further detail in Rath and Fontana (2003). The models used for synthesis in this experiment were created as part of the EU IST Disappearing Computer initiative and, in particular, the project "the Sounding Object" (SOb). The modal synthesis model of impact was written in *PureData* and created by the University of Verona, Italy as part of this project. A modal syntheses modal is a bank of damped harmonic oscillators. These oscillators are controlled and excited by a external stimuli, the settings of the oscillators such as their frequencies and dampings are based on the modelled geometry and material properties of the real object being modelled. The sound model parameters for a given real object are used as the basis for the synthesis model. These parameters are obtained experimentally using recordings of the object's impulse responses by fitting these synthesised model's parameters to these recordings of the object. The design, development and a detailed exploration of these modal synthesis models can be found in Rocchesso and Fontana (2003). The particular models used in this research were modal synthesis impact models with a higher level control model for bouncing with the additional parameters that included the materiality of the dropped object. A reference (Rath and Fontana, 2003), is available and covers the main features of these particular models and the meaning of their parameters.

C.1.1 Synthesised sounds used in the first pilot probe

The values used to create the synthesised sounds used in the first pilot probe are shown in Table C-1.

Short name	Elasticity	Damping	Gravity force	Strike velocity	Freq (Hz)	Decay time (s)
w-s-h-s-gl-pd-1.wav	15000	46.4159	990	630.957	1758.52	0.043070
w-m-h-l-gl-pd-2.wav	5540.1	8.57696	990	1318.26	1782.52	0.043070
w-m-h-s-gl-pd-3.wav	15000	21.5443	950	1584.89	1388.82	0.090315
w-m-h-s-gl-pd-4.wav	3161.6	21.5443	580	2290.87	1388.82	0.090315
w-s-h-m-gl-pd-5.wav	15000	46.4159	450	630.957	1782.52	0.043070
w-s-h-m-wd-pd-1.wav	15000	46.4159	450	630.957	1758.52	0.233307
w-m-h-m-wd-pd-2.wav	15000	63.0957	940	912.011	1113.23	0.603386
w-s-h-m-wd-pd-3.wav	11395	2.92864	860	301.995	1294.33	0.752992
w-s-h-s-wd-pd-4.wav	1309.5	4.64159	970	436.516	1294.33	0.784488

Table C-1: Parameter values for the impact model (Rath and Fontana, 2003) used to generate the synthesised sounds used in the first pilot probe.

C.1.2 Synthesised sounds used in the second pilot study

The values used to create the synthesised sounds used in the second pilot study are shown in Table C-2 for the single mode impact sounds and in Table C-3 for the two mode impact sounds. The sound model parameters for a given real object are used as the basis for the synthesis model. These parameters are obtained experimentally using recordings of the object's impulse responses by fitting these synthesised model's parameters to these recordings of the object. The design, development and a detailed exploration of these modal synthesis models can be found in Rocchesso and Fontana (2003).

Short name	Elasticity	Damping	Gravity force	Strike velocity	Freq (Hz)	Decay time (s)
w-s-h-m-wd-m1-pd-1.wav	15000	46.4159	450	630.957	1758.52	0.233307
w-m-h-m-wd-m1-pd-2.wav	15000	63.0957	940	912.011	1113.23	0.603386
w-s-h-m-wd-m1-pd-3.wav	11395	2.92864	860	301.995	1294.33	0.752992
w-s-h-s-wd-m1-pd-4.wav	1309.5	4.64159	970	436.516	1294.33	0.784488
w-s-h-s-wd-m1-pd-5.wav	1309.5	8.57696	990	1318.26	1254.95	0.784488
w-m-h-s-wd-m1-pd-6.wav	3162.28	25.1189	900	524.807	1322.83	0.233307

Table C-2*:* Parameter values for the single mode impact model (Rath and Fontana, 2003) used to generate the synthesised sounds used in the second pilot study.

Short name	Hammer mass	Initial interval (ms)	Acceleration - decceleration	Initial value
w-s-h-s-wd-m2-pd-1.wav	0.0215443	228.530	0.76	0.56
w-s-h-m-wd-m2-pd-2.wav	0.0215443	306.516	0.74	0.75
w-s-h-l-wd-m2-pd-3.wav	0.0398107	207.223	0.72	0.57
w-m-h-s-wd-m2-pd-4.wav	0.0398107	207.223	0.72	0.57
w-m-h-m-wd-m2-pd-5.wav	0.0398107	277.939	0.70	0.75
w-m-h-l-wd-m2-pd-6.wav	0.0398107	372.786	0.70	1.00

Table C-3*:* Parameter values for the two mode impact model (Rath and Fontana, 2003) used to generate the synthesised sounds used in the second pilot study.

D.1 Appendix D - Chapter 4 - Task Lists and Questionnaire

This appendix holds the tasks and the questions asked in the pilot studies presented in Chapter 4.

Task Sheet for Evaluators - Mapping Size versus Height of Dropped Objects
Evaluators Instructions:
Using the Sonic Browser application, please perform the task that is listed below. The observer will be present and will ask you to comment on your actions and reasoning for actions, this is known as the "Thinking Aloud" method. Please remember this is an evaluation of the perception of the sounds, not an evaluation of you! Do not feel afraid to comment on either the application or the sounds either positively or negatively, as the goal of the experiment is to evaluate the perception of the sound's not you.

Task List:

Drag the object's on screen to what you think is the correct X – Y axis where the size of the object being dropped is along the X axis and the height the object was dropped from is along the Y axis. Figure 1 shows an illustration of this idea.

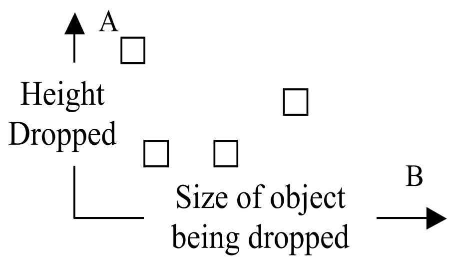


Figure 1: Illustrating the aim of the experiment

To drag the object, simply move the cursor over the object and press & hold the mouse button while you drag it to where you feel it belongs in the scale. In Figure 1 the object 'A' would represent a small object that has been dropped from a great height, whereas the object 'B' would represent a large object that has been dropped from only a small height. When you are satisfied with the layout of the objects, please indicate to the observer that you have completed the task.

Task Sheet for Evaluators - Determining Realism of Objects
Evaluators Instructions:
Using the Sonic Browser application, please perform the task that is listed below. The observer will be present and will ask you to comment on your actions and reasoning for actions, this is known as the “Thinking Aloud” method. Please remember this is an evaluation of the perception of the sounds, not an evaluation of you! Do not feel afraid to comment on either the application or the sounds either positively or negatively, as the goal of the experiment is to evaluate the perception of the sound’s not you.

Task List:

Drag the object’s on screen to what you think is the correct place for the object where the realism of the object is along the X axis. Figure 1 shows an illustration of this idea.

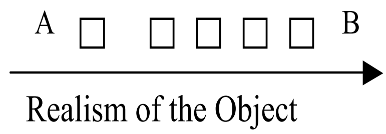


Figure 1: Illustrating the aim of the experiment

To drag the object, simply move the cursor over the object and press & hold the mouse button while you drag it to where you feel it belongs in the scale. In Figure 1 the object ‘A’ would represent an object that doesn’t sound real, whereas the object ‘B’ would represent an object that has sounds like a real sound. When you are satisfied with the layout of the objects, please indicate to the observer that you have completed the task.

Each of the users completed the scaling of the synthetic sounds task and then was asked to fill in a questionnaire for the set of synthetic sounds presented to them.

Q1: How difficult did you find the task?

easy - 0 ***neither - 3*** ***hard – 6***

Q2: Overall, how would rate the ease of use or usability of this design?

easy to use - 0 ***neither - 3*** ***hard to use – 6***

Q3: How realistic did you find the sounds?

realistic - 0 ***neither - 3*** ***unrealistic – 6***

Q4: How did you find the quality of the sounds?

good - 0 ***neither - 3*** ***poor – 6***

Q5: Was there a noticeable play lag before a sound was played?

no lag - 0 ***average play lag- 3*** ***distinct lag – 6***

The results of the questionnaire for the first pilot probe in Chapter 4 are shown in Figure D-1 using a boxplot with the cumulative participant responses represented.

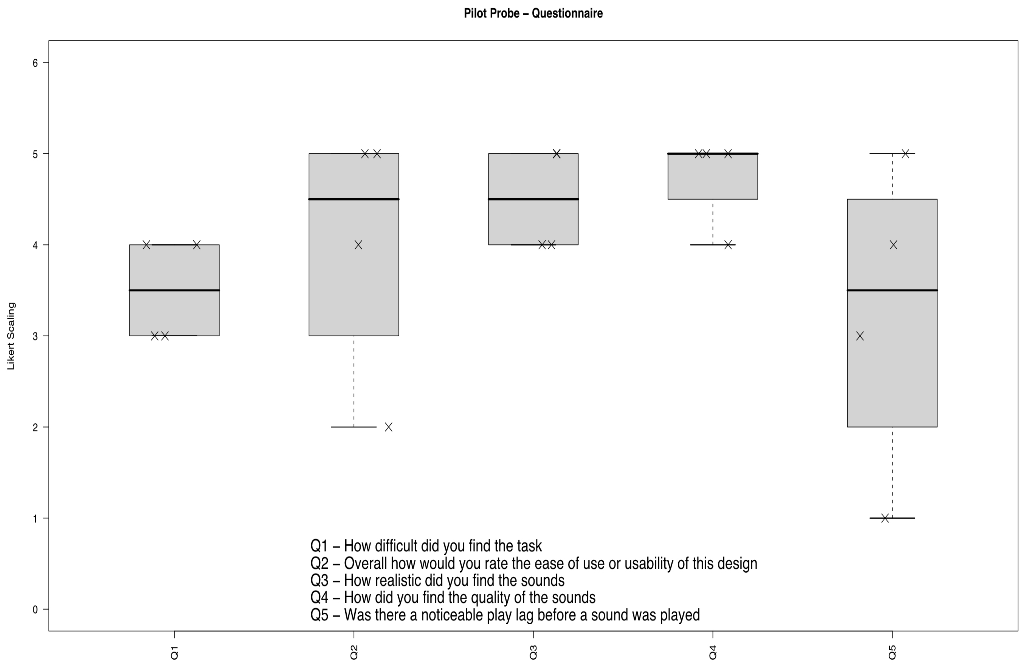


Fig. D-1: *The results of the questionnaire posed to participants in the first pilot study.*

In Figure D-2, the results of the questionnaire for the second pilot study in Chapter 4 are shown using a boxplot with the cumulative participant responses represented.

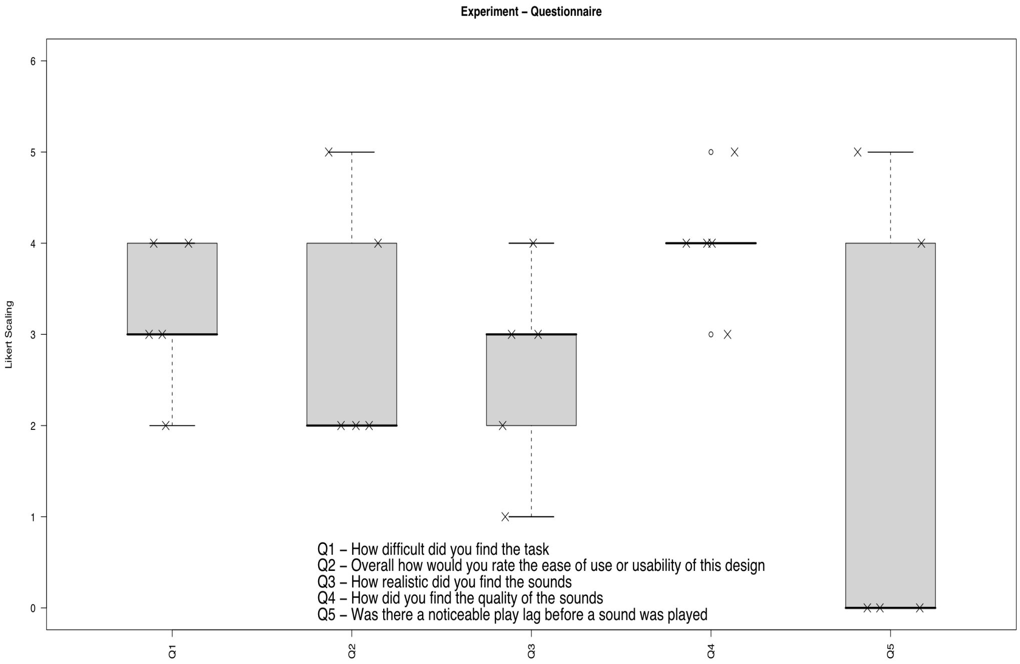


Fig. D-2*: The results of the questionnaire posed to participants in the second pilot study.*

E.1 Appendix E - Data related to the First Study in Chapter 5

This appendix presents the tasks and the data used for the analysis of the first study presented in Chapter 5. The file names and descriptions of the sounds used are shown in Table E-1 (prior categorisation) and in Table E-2 (no prior categorisation). The identification averages of the participant's descriptors are given in Table E-3 (prior categorisation) and in Table E-4 (no prior categorisation). The categories and the casual uncertainty for the object descriptors are shown in Tables E-7 and E-8 for those with prior categorisation and in Tables E-11 and E-8 for those with no prior categorisation. The categories and the casual uncertainty for the action descriptors are shown in Tables E-15 and E-16 for those with prior categorisation and in Tables E-19 and E-20 for those with no prior categorisation. In this study, 11 participants generated between 3 and 4 categories for each of the conditions when their descriptors where analysed.

Task List for Evaluators - DataSet 1 - Concurrent Auditory Icons Descriptions
Evaluators Instructions:
Using the experimental interface wizard, please perform the tasks that are listed below, you may do so in the order presented. The observer will be present and will ask you to comment on your actions and reasoning for actions, this is known as the "Thinking Aloud" method. Please remember this is an evaluation of the sounds, not an evaluation of you! Do not feel afraid to comment on the application or sounds either positively or negatively, as the goal of the experiment is to evaluate the application and sounds from a user's perspective.

Scenario:
You've been asked to listen to a set of sounds and write your own description of what you feel each sound is.

Number of tasks to complete: 4 (four)

Task List:

1. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

2. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

3. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

4. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

ID	Filename	Description
1	Waterfillingglassbottle	Water filling a glass bottle from a tap in a kitchen
2	BouncingHeavyBall	Ball bouncing three times
3	RunningUpstairs	A person running up carpeted stairs in a wooden hallway.
4	MotorbikePassing	A motorbike passing the house and driving up through the estate during the day.
5	BrushingTeeth	A person brushing their teeth
6	KnockingOnDoor2	Knocking on a door
7	AlarmClockElectric	An electic alarm block buzzing
8	BirdsSinging	Several birds singing in a rural setting
9	GlassWindowBreaking	A glass window breaking
10	Waterfrombathtap	Water flowing quickly from a bath tap in a tiled bathroom.
11	KnockingDoor	A person knocking on a wooden kitchen door.
12	ChainRattle	Rattling a metal chain several times in bare hands inside in a kitchen
13	RunningOnConcrete	Running on a concrete surface
14	SawingWood	Sawing a piece of wood in a kitchen on a wooden table.
15	BouncingWoodenBall	Ball bouncing three times
16	UsingVendingMachine	Using a vending machine and getting an item
17	MotorbikeStarting	A motorbike starting, revved and driven off out of the estate during the day.
18	ClinkingGlasses	Several glasses clinking off each other

Table E-1*:* The sound files and their descriptions as used in the 3, 4, 5, and 6 conditions with prior categorisation

ID	Filename	Description
1	DishWasher	Dishwasher in operation
2	Pouringoutofglass	Pouring water out of glass bottle into a sink in a kitchen
3	Watertiles	Water splashing slowly onto tiles in a kitchen
4	WashingHands	Washing hands in sink
5	KnockingOnDoor	Rapid Door Knocks
6	Stirringwater	Stirring water with a metal spoon in a cup in a kitchen
7	FingersDrumming	Bare fingers drumming on a wooden table in a kitchen.
8	BirdsSinging	Several birds singing in a rural setting
9	Raintrees	Rain falling through the trees
10	FootstepsSticks	Walking on sticks outdoors under trees and breaking them.
11	TractorReturningFromTheFields	Tractor returning from the fields
12	Windtrees	Wind rustling through the leaves on a summers day.
13	ManyGlassesClinkingInToast	Many glasses clinking in toast
14	PlatesInPress	Putting plate on top of another in a press in kitchen and closing the press after.
15	GlassesClashing	Two glasses clashing together 5 times held in hand in a kitchen.
16	Breaking3Glasses	Breaking 3 glasses in succession against cement brick outside on a summer's day.
17	CupsSliding	Sliding cups into a cupboard after drying them in a kitchen
18	CeramicCupsBreak	Breaking 2 ceramic cups in succession of a cement block outside during the day

Table E-2*:* The sound files and their descriptions as used in the 3, 4, 5, and 6 conditions with **no** prior categorisation

Identification	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	Average for Condition
Condition 3	3/3	2/3	3/3	2/3	2/3	2/3	2/3	3/3	3/3	3/3	3/3	**84.8%**
Condition 4	4/4	4/4	4/4	4/4	4/4	4/4	4/4	4/4	4/4	4/4	4/4	**100.0%**
Condition 5	5/5	5/5	5/5	4/5	4/5	4/5	4/5	4/5	4/5	4/5	4/5	**85.5%**
Condition 6	5/6	6/6	6/6	5/6	6/6	6/6	4/6	5/6	5/6	5/6	6/6	**89.4%**
Average for Participant	95.8%	91.7%	100.0%	82.5%	86.7%	86.7%	78.3%	90.8%	90.8%	90.8%	95.0%	

Table E-3*:* The results of identifications by participant and by condition for the 3, 4, 5, and 6 conditions with prior categorisation

Identification	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	Average for Condition
Condition 3	3/3	3/3	2/3	3/3	3/3	3/3	3/3	1/3	1/3	3/3	3/3	**84.8%**
Condition 4	3/4	4/4	4/4	3/4	4/4	4/4	3/4	4/4	4/4	3/4	3/4	**88.6%**
Condition 5	3/5	4/5	5/5	4/5	4/5	3/5	4/5	3/5	3/5	3/5	4/5	**72.7%**
Condition 6	3/6	4/6	3/6	3/6	3/6	4/6	4/6	1/6	3/6	2/6	3/6	**50.0%**
Average for Participant	71.3%	86.7%	79.2%	76.3%	82.5%	81.7%	80.4%	52.5%	60.8%	67.1%	76.3%	

Table E-4*:* The results of identifications by participant and by condition for the 3, 4, 5, and 6 conditions with **no** prior categorisation

Number	P1-Prior	P2-Prior	P3-Prior	P4-Prior	P5-Prior	P6-Prior
1	liquid, container	water, container	liquid, jug	water, glass	water, jug	liquid, container
2	?	ball	marble, wooden table	metal ball	marble	marble, table
3	wall	x	stairs	x	x	x
4	motorbike	engine/vehicle	motorbike	car	motorbike	motorbike
5	teeth, brush	teeth, brush	teeth, brush	teeth, brush	teeth, brush	teeth, brush
6	door	door	wooden door	wooden door	door	surface
7	alarm	alarm	alarm	alarm	alarm	alarm
8	birds	birds	birds	birds	birds	birds
9	glass	glass	glass	glass jar with metal bits	glass	glass, floor
10	stream	water	water	water	water	water
11	drum	surface	wooden door	wooden door	wood, hammer	drum
12	chain	metal chain	metal chain	x	x	x
13	footsteps	girl shoes, surface	girls shoes	shoes, wooden floor	footsteps	shoes, surface
14	x	wood	wood	wood	material	material
15	table	ball	marble, wooden table	metal ball	marble	marble, table
16	drawer	cash register	metal toolbox	x	box	metal object
17	motorbike	vehicle	motorbike	car	car	vehicle
18	glasses	glasses	metal, glass	glasses	glass	metal, glass

Table E-5*:* The first stage classification (1 of 2) of the object responses from the participants of the 3, 4, 5, and 6 conditions with prior categorisation. P(NUM) represents the participant's identification number.

Number	P7-Prior	P8-Prior	P9-Prior	P10-Prior	P11-Prior
1	glass, water	bottle, tap	water, bottle	water	water, glass container
2	glass ball, table	metal ball	surface	marble, table	marble, table
3	x	fingers	stairs	surface	stairs
4	car	vechicle	vehicle	vechicle	vehicle
5	teeth, brush	teeth, brush	teeth, brush	teeth, brush	teeth, brush
6	door	door / table	door	door / table	surface, wood
7	alarm	alarm	alarm	alarm	alarm
8	birds	birds	birds	birds	birds
9	glass	glass	glass	glass	glass
10	water	water, tap	water	water, tap	water
11	wood, hammer	door	door	wood door / table	window
12	x	x	x	x	x
13	footsteps	girls shoes, surface	girls shoes, surface	shoes, surface	girls shoes, surface
14	x	wood	wood	x	wood
15	x	metal ball	table	marble ball	marble, table
16	metal object	metal cabinet	metal toolbox	metal toolbox	vending machine
17	car	car	motorbike	vehicle	vehicle
18	x	x	x	glass	teaspoon, cup

Table E-6*:* The first stage classification (2 of 2) of the object responses from the participants of the 3, 4, 5, and 6 conditions with prior categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03	cat04
1	liquid	water		
2	ball	metal ball	glass ball	surface
3	wall	stairs	fingers	
4	motorbike	vehicle	car	
5	teeth			
6	door	surface		
7	alarm			
8	birds			
9	glass			
10	water			
11	drum	surface	door	wood
12	metal chain			
13	footsteps			
14	wood	material		
15	table	ball	metal ball	marble
16	drawer	cash register	metal toolbox	metal object
17	motorbike	vehicle	car	
18	glasses			

Table E-7*:* Categories of object responses from participants descriptions of the 3, 4, 5, and 6 conditions with prior categorisation

Number	cat01	cat02	cat03	cat04					*Hcu*	
1	11				0	0	0	0	0	
2	9	1			-0.2369	-0.3145	0	0	0.5514	
3	5	1			-0.5170	-0.3145	0	0	0.8315	**0.461**
4	11				0	0	0	0	0	
5	11				0	0	0	0	0	
6	9	2			-0.2369	-0.4472	0	0	0.6840	
7	11				0	0	0	0	0	**0.171**
8	11				0	0	0	0	0	
9	11				0	0	0	0	0	
10	11				0	0	0	0	0	
11	2	1	4	4	-0.4472	-0.3145	-0.5307	-0.5307	1.8231	
12	3				-0.5112	0	0	0	0.5112	**0.467**
13	11				0	0	0	0	0	
14	8				-0.3341	0	0	0	0.3341	
15	9	1			-0.2369	-0.3145	0	0	0.5514	
16	3	7			-0.5112	-0.4150	0	0	0.9262	
17	11				0	0	0	0	0	
18	8				-0.3341	0	0	0	0.3341	**0.358**

Table E-8*:* Analysis of object responses from participants descriptions of the 3, 4, 5, and 6 conditions with prior categorisation

Number	P1-Random	P2-Random	P3-Random	P4-Random	P5-Random	P6-Random
1	machine	machine	machine	machine	machine	machine
2	water, container	water, container	water, container	water, container	water, container	water, container
3	water, container	water, surface	x	water, surface	water, surface	water, surface
4	water, surface	water, surface	water, surface	water, surface	water, surface	water, surface
5	x	door	surface	x	door	door
6	cup, spoon	glass, spoon, surface	water, glass, spoon	spoon, glass, ceramic, surface	glass, spoon, surface	glass, spoon, surface
7	fingers, surface	fingers, surface	fingers, surface	fingers, surface	wood	fingers, surface
8	birds	birds	birds	birds	birds	birds
9	x	water	water, surface	water, surface	x	x
10	wood, surface	wood, surface	wood, surface	wood, surface	wood, surface	wood, surface
11	vehicle	vehicle	vehicle	vehicle	vehicle	vehicle
12	x	x	wind, mic	x	wind	x
13	glass	x	glass	x	x	x
14	door	wood	wood	ceramic	container	ceramic
15	x	glass	x	x	x	glass
16	glass	glass	glass	x	glass	glass
17	x	ceramic	x	wood, ceramic	ceramic	wood, ceramic
18	x	x	x	ceramic	x	x

Table E-9*:* The first stage classification (1 of 2) of the object responses from the participants of the 3, 4, 5, and 6 conditions with no prior (random) categorisation. P(NUM) represents the participant's identification number.

Number	P7-Random	P8-Random	P9-Random	P10-Random	P11-Random
1	machine	x	x	machine	machine
2	water, container	water, container	water, container	water, container	water, container
3	water, surface	x	x	water, surface	water, surface
4	water, surface	water, surface	water, surface	water, surface	x
5	x	door	door	x	door
6	glass, spoon, surface	glass, spoon, surface	glass, spoon, surface	glass, spoon, surface	glass, spoon, surface
7	fingers, surface	fingers, surface	fingers, surface	fingers, surface	fingers, surface
8	birds	birds	birds	birds	birds
9	x	x	x	x	water, surface
10	wood	wood	wood, surface	wood	surface
11	vehicle	vehicle	vehicle	vehicle	vehicle
12	wind	x	x	x	x
13	x	x	x	x	x
14	container	x	container	x	container
15	x	x	x	metal	x
16	glass	glass	glass	glass	glass
17	wood, ceramic	x	ceramic	x	ceramic
18	ceramic	x	x	x	x

Table E-10*:* The first stage classification (2 of 2) of the object responses from the participants of the 3, 4, 5, and 6 conditions with no prior (random) categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03	cat04
1	machine			
2	water, container			
3	water, container	water, surface		
4	water, surface			
5	door	surface		
6	spoon	glass	ceramic	surface
7	fingers	surface	wood	
8	birds			
9	water	surface		
10	wood	surface		
11	vehicle			
12	wind	mic		
13	glass	container		
14	door	wood	ceramic	container
15	glass	metal		
16	glass			
17	ceramic			
18	ceramic			

Table E-11*:* Categories of object responses from participants descriptions of the 3, 4, 5, and 6 conditions with **no** prior (random) categorisation

Number	cat01	cat02	cat03	cat04					*Hcu*	
1	9				-0.2369	0	0	0	0.2369	
2	11				0	0	0	0	0	
3	8				-0.3341	0	0	0	0.3341	**0.190**
4	10				-0.1250	0	0	0	0.1250	
5	6	1			-0.4770	-0.3145	0	0	0.7915	
6	11				0	0	0	0	0	
7	10	1			-0.1250	-0.3145	0	0	0.4395	**0.339**
8	11				0	0	0	0	0	
9	4				-0.5307	0	0	0	0.5307	
10	11				0	0	0	0	0	
11	11				0	0	0	0	0	
12	3				-0.5112	0	0	0	0.5112	**0.208**
13	2				-0.4472	0	0	0	0.4472	
14	1	2	3	3	-0.3145	-0.4472	-0.5112	-0.5112	1.7841	
15	2	1			-0.4472	-0.3145	0	0	0.7617	
16	10				-0.1250	0	0	0	0.1250	
17	7				-0.4150	0	0	0	0.4150	
18	2				-0.4472	0	0	0	0.4472	**0.663**

Table E-12*:* Analysis of object responses from participants descriptions of the 3, 4, 5, and 6 conditions with **no** prior categorisation

Number	P1-Prior	P2-Prior	P3-Prior	P4-Prior	P5-Prior	P6-Prior
1	filling	filling	filling	filling	filling	filling
2	dropping	bouncing	dropping	bouncing	bouncing	dropping
3	hitting	x	running	x	x	x
4	driving by	driving	passing by	passing by	driving	passing by
5	brushing	brushing	brushing	brushing	brushing	brushing
6	knocking	knocking	knocking	knocking	knocking	knocking
7	beeping	beeping	beeping	beeping	beeping	beeping
8	singing	whistling	chirping	chirping	singing	singing
9	breaking	breaking	breaking	breaking	breaking	breaking
10	flowing	flowing	flowing	flowing	pouring	flowing
11	banging	knocking	knocking	knocking	hitting	banging
12	rattling	dragging	falling	?	x	x
13	running	running	running	walking	walking	running
14	x	sawing	sawing	sawing	sawing	sawing
15	dropping	bouncing	dropping	bouncing	dropping	dropping
16	opening/closing	opening/closing	rattling	x	dropping	dropping
17	starting	starting	starting	starting	driving	starting
18	clanking	hitting	hitting	clinking	?	hitting

Table E-13*:* The first stage classification (1 of 2) of the action responses from the participants of the 3, 4, 5, and 6 conditions with prior categorisation. P(NUM) represents the participant's identification number.

Number	P7-Prior	P8-Prior	P9-Prior	P10-Prior	P11-Prior
1	filling	filling	filling	pouring	filling
2	dropping	dropping	dropping	dropping	dropping
3	x	hitting	running	hitting	running
4	passing by	passing by	driving	passing by	passing by
5	brushing	brushing	brushing	brushing	brushing
6	knocking	knocking	knocking	knocking	knocking
7	beeping	beeping	beeping	beeping	beeping
8	singing	singing	singing	singing	singing
9	breaking	breaking	breaking	breaking	breaking
10	flowing	flowing	flowing	flowing	flowing
11	knocking	knocking	knocking	knocking	knocking
12	x	x	x	x	x
13	running	running	running	running	running
14	x	sawing	sawing	x	sawing
15	dropping	dropping	dropping	dropping	dropping
16	moved	opening/closing	opening/closing	opening/closing	opening/closing
17	driving	driving away	driving away	driving away	driving away
18	x	x	x	clinking	clinking

Table E-14*:* The first stage classification (2 of 2) of the action responses from the participants of the 3, 4, 5, and 6 conditions with prior categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03	cat04
1	filling			
2	dropping	bouncing		
3	hitting	running		
4	driving	passing by		
5	brushing			
6	knocking			
7	beeping			
8	singing	whistling	chirping	
9	breaking			
10	flowing	pouring		
11	banging	knocking	hitting	
12	rattling	dragging	falling	
13	running	walking		
14	sawing			
15	dropping	bouncing		
16	opening/closing	rattling	dropping	moving
17	starting	driving		
18	clanking	hitting	clinking	

Table E-15*:* Categories of action responses from participants descriptions of the 3, 4, 5, and 6 conditions with prior categorisation

Number	cat01	cat02	cat03	cat04					*Hcu*	
1	10	1			-0.1250	-0.3145	0	0	0.4395	
2	8	3			-0.3341	-0.5112	0	0	0.8454	
3	3	3			-0.5112	-0.5112	0	0	1.0224	**0.769**
4	11				0	0	0	0	0	
5	11				0	0	0	0	0	
6	11				0	0	0	0	0	
7	11				0	0	0	0	0	**0**
8	11				0	0	0	0	0	
9	11				0	0	0	0	0	
10	11				0	0	0	0	0	
11	11				0	0	0	0	0	
12	1	1	1		-0.3145	-0.3145	-0.3145	0	0.9435	**0.189**
13	11				0	0	0	0	0	
14	8				-0.3341	0	0	0	0.3341	
15	11				0	0	0	0	0	
16	6	1	2	1	-0.4770	-0.3145	-0.4472	-0.3145	1.5531	
17	11				0	0	0	0	0	
18	7				-0.4150	0	0	0	0.4150	**0.384**

Table E-16: Analysis of action responses from participants descriptions of the 3, 4, 5, and 6 conditions with prior categorisation

Number	P1-Random	P2-Random	P3-Random	P4-Random	P5-Random	P6-Random
1	washing	spinning	humming	spinning / hitting	spinning / hitting	washing / hitting
2	filling	filling	x	pouring	pouring	filling
3	filling	pouring	pouring	pouring	pouring	pouring
4	pouring	pouring	pouring	pouring	pouring	pouring
5	hitting	knocking	hitting	x	knocking	hitting
6	stirring-hitting	hitting	hitting	hitting	hitting	hitting
7	hitting	hitting	hitting	hitting	hitting	hitting
8	singing	singing	singing	singing	singing	singing
9	x	pouring	pouring	pouring	x	x
10	breaking, walking	breaking, walking	breaking, walking	breaking, walking	breaking	breaking/walking
11	driving	driving	driving	driving	driving	driving
12	x	x	blowing	x	blowing	x
13	hitting	x	hitting	x	x	x
14	opening/closing	hitting	dropping	hitting	hitting	hitting
15	x	hitting	x	x	x	hitting
16	breaking	breaking	breaking	x	breaking	breaking
17	x	hitting	x	hitting	hitting	hitting
18	x	x	x	breaking	x	x

Table E-17: The first stage classification (1 of 2) of the action responses from the participants of the 3, 4, 5, and 6 conditions with no prior (random) categorisation. P(NUM) represents the participant's identification number.

Number	P7-Random	P8-Random	P9-Random	P10-Random	P11-Random
1	spinning	x	x	washing	spinning
2	filling	filling	filling	filling	pouring
3	pouring	x	x	pouring	pouring
4	pouring	pouring	pouring	pouring	x
5	x	knocking	knocking	x	knocking
6	hitting	hitting	hitting	hitting	hitting
7	hitting	hitting	hitting	hitting	hitting
8	singing	singing	whistling	singing	singing
9	x	x	x	x	pouring
10	burning	breaking/walking	breaking/walking	burning	breaking, walking
11	driving	driving	driving	driving	driving
12	blowing	x	x	x	x
13	x	x	x	x	x
14	hitting	x	hitting	x	hitting
15	x	x	x	hitting	x
16	breaking	breaking	breaking	breaking	breaking
17	hitting	x	hitting	x	hitting
18	breaking	x	x	x	x

Table E-18: The first stage classification (2 of 2) of the action responses from the participants of the 3, 4, 5, and 6 conditions with no prior (random) categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03
1	washing	spinning	humming
2	filling	pouring	
3	filling	pouring	
4	pouring		
5	hitting	knocking	
6	hitting		
7	hitting		
8	singing	whistling	
9	pouring		
10	breaking, walking	burning	
11	driving		
12	blowing		
13	hitting		
14	opening/closing	hitting	dropping
15	hitting		
16	breaking		
17	hitting		
18	breaking		

Table E-19: Categories of action responses from participants descriptions of the 3, 4, 5, and 6 conditions with **no** prior categorisation

Number	cat01	cat02	cat03				*Hcu*	
1	3	5	1	-0.5112	-0.5170	-0.3145	1.3428	
2	7	3		-0.4150	-0.5112	0	0.9262	
3	1	8		-0.3145	-0.3341	0	0.6486	**0.973**
4	10			-0.1250	0	0	0.1250	
5	3	5		-0.5112	-0.5170	0	1.0283	
6	11			0	0	0	0	
7	11			0	0	0	0	**0.288**
8	10	1		-0.1250	-0.3145	0	0.4395	
9	5			-0.5170	0	0	0.5170	
10	8	2		-0.3341	-0.4472	0	0.7813	
11	11			0	0	0	0	
12	3			-0.5112	0	0	0.5112	**0.450**
13	2			-0.4472	0	0	0.4472	
14	1	7	1	-0.3145	-0.4150	-0.3145	1.0439	
15	3			-0.5112	0	0	0.5112	
16	10			-0.1250	0	0	0.1250	
17	7			-0.4150	0	0	0.4150	
18	2			-0.4472	0	0	0.4472	**0.498**

Table E-20*:* Analysis of action responses from participants descriptions of the 3, 4, 5, and 6 conditions with **no** prior categorisation

F.1 Appendix F - Data related to the Second Study in Chapter 5

This appendix presents the tasks and the data used for the analysis of the second study presented in Chapter 5. The file names and descriptions of the sounds used are shown in Table F-1 (prior categorisation) and in Table F-2 (no prior categorisation). The identification averages of the participant's descriptors are given in Table F-3 (prior categorisation) and in Table F-4 (no prior categorisation). The categories and the casual uncertainty for the object descriptors are shown in Tables F-6 and F-7 for those with prior categorisation and in Tables F-9 and F-7 for those with no prior categorisation. The categories and the casual uncertainty for the action descriptors are shown in Tables F-12 and F-13 for those with prior categorisation and in Tables F-15 and F-16 for those with no prior categorisation. In this study, 5 participants generated between 3 and 4 categories for each of the conditions when their descriptors where analysed.

Task List for Evaluators - DataSet 1 - Concurrent Auditory Icons Descriptions
Evaluators Instructions:
Using the experimental interface wizard, please perform the tasks that are listed below, you may do so in the order presented. The observer will be present and will ask you to comment on your actions and reasoning for actions, this is known as the "Thinking Aloud" method. Please remember this is an evaluation of the sounds, not an evaluation of you! Do not feel afraid to comment on the application or sounds either positively or negatively, as the goal of the experiment is to evaluate the application and sounds from a user's perspective.

Scenario:
You've been asked to listen to a set of sounds and write your own description of what you feel each sound is.

Number of tasks to complete: 4 (four)

Task List:

1. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

2. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

3. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

4. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided on the application.

ID	Filename	Description
1	A2-10	Rain water in manhole
2	A3-4	Dentist - drilling - grinding teeth
3	A3-30	Front door - modern - outside - A person enters
4	B1-6	Vacuum cleaner - rewinding flex cord
5	B1-56	Pouring Beer - kitchen
6	B1-77	Venetian Blinds - down - adjusting - up
7	B2-32	Plates dropped on floor - kitchen
8	B2-64	Brushing pants
9	B1-47	Toaster - kitchen
10	B1-83	Light switch - on/off - 2 versions
11	B2-26	Cooking oatmeal - kitchen
12	B2-49	Clipping nails - bathroom
13	E3-16	Hammering on sheet metal - hard - tools
14	F1-7	Eraser - on paper - office
15	G4-39	Bicycle on asphalt - cycles off - passes - comes in/stops
16	A3-36	Beating on heavy door - copper clad - from outside
17	B1-8	Window cleaning
18	B1-38	Cereal - rice cripsies in bowl adding milk - kitchen
19	B1-69	Shaving - electric razor Philips - bathroom
20	B1-78	Drapes - curtains - closing then opening
21	B2-12	Whistling tea kettle - kitchen
22	B2-35	Scrubbing floor
23	C7-34	Small dog barking - mammal
24	F1-22	Binder being used - office
25	B1-58	Ice cube dropping into glass - kitchen
26	B1-89	Cloth being torn / being cut
27	I4-9	Child laughing - 8 months old
28	C3-45	Ice thaw - crushed ice in ocean bay - waves
29	C7-35	Dog - medium - barkings
30	F1-6	Writng with pencil - 3 versions - office
31	F1-28	Chair on castors - long sequence - office
32	G4-52	Bicycle on asphalt - passing with engaged dynamo
33	B1-2	Cigarette lighter - mehanical - electric
34	B1-22	Footsteps - wooden floor - male - stairs up/down

Table F-1*:* The sound files and their descriptions as used in the 7, 8, 9, and 10 conditions with prior categorisation

ID	Filename	Description
1	B1-7	Sweeping floor with brush - kitchen
2	B-35	Coffee brewin and filtering - kitchen
3	B1-71	Shower with person - shower curtain - bathroom
4	C6-21	Water splashing - kitchen
5	B1-50	Kitchen table being set - kitchen
6	B1-31	Water boiling - kitchen
7	B1-1	Match being struck and lit - 3 times
8	B1-81	Aquarium with bubbles pump - 2 version
9	B2-8	Window sliding glass - opening/closing
10	B2-28	Tab pulled off soft drink can and poured - kitchen
11	B2-43	Shower with person - showering - bathroom
12	B2-36	Wringing out water from rag - kitchen
13	B2-65	Brushing clothes
14	B2-60	Drapes - opening/closing
15	B2-66	Zipper - up/down
16	C7-22	Waterfall small
17	C3-54	Waves - long - pebble beach
18	C3-59	Raining heavily on roof and pavement - varying intensity
19	C4-30	Waves - small - pebble beach
20	C6-20	Water pouring - kitchen
21	C6-38	Plops - small - 4 versions - water
22	C7-52	Rain medium - splattering close-by
23	B1-40	Milk container opened and poured into tall glass - 2 versions - kitchen
24	B2-27	Ketchup squirt - kitchen
25	C6-59	Hail on window
26	C6-63	Wind across a lakeshore
27	C7-8	Brook flowing - atmosphere - water
28	C7-40	Fire - crackling fire
29	E1-19	Hammering nails - 1 1/4
30	G4-63	Bicycle stand up/down
31	I3-31	Foorsteps on dry snow with creaking - leaving - normal pace
32	C6-27	Water pouring into a container - kitchen
33	E3-12	Hammering nails - outdoors - tools
34	I3-33	Footsteps on dry snow with creaking - leaving - fast pace

Table F-2*:* The sound files and their descriptions as used in the 7, 8, 9, and 10 conditions with **no** prior categorisation

Identification	P1	P2	P3	P4	P5	Average for Condition
Condition 7	4/7	6/7	5/7	7/7	4/7	**74.3%**
Condition 8	6/8	4/8	3/8	3/8	4/8	**50.0%**
Condition 9	7/9	7/9	6/9	8/9	7/9	**77.8%**
Condition 10	7/10	9/10	7/10	8/10	6/10	**74.0%**
Average for Participant	70.0%	75.9%	61.4%	76.6%	61.2%	

Table F-3*:* The results of identifications by participant and by condition for the 7, 8, 9, and 10 conditions with prior categorisation

Identification	P1	P2	P3	P4	P5	Average for Condition
Condition 7	4/7	6/7	5/7	5/7	4/7	**68.6%**
Condition 8	5/8	5/8	5/8	5/8	6/8	**65.0%**
Condition 9	3/9	5/9	3/9	5/9	5/9	**46.7%**
Condition 10	3/10	4/10	3/10	5/10	3/10	**36.0%**
Average for Participant	45.7%	60.9%	49.3%	59.9%	54.4%	

Table F-4: The results of identifications by participant and by condition for the 7, 8, 9, and 10 conditions with **no** prior categorisation

Number	P1-Prior	P2-Prior	P3-Prior	P4-Prior	P5-Prior
1	water	water	?	liquid	liquid
2	machine	machine	machine	machine	machine
3	?	?	door	door	?
4	?	?	?	recording media	?
5	liquid	water	liquid	liquid	liquid
6	door	metal/ceramic	?	fabric/cloth	?
7	?	wood	metal	glass	?
8	people	?	?	wood	wood
9	?	?	?	?	?
10	?	metal	?	?	?
11	wind	liquid	liquid	?	wind
12	people	?	?	?	?
13	metal	plastic	metal	metal	metal
14	surface	?	paper	wood	wood
15	vehicle	?	?	?	?
16	metal	metal	wood	?	surface
17	glass	glass	glass	glass	?
18	container	?	container	container	glass
19	machine	machine	surface	machine	vehicle
20	?	?	door	animal	door
21	liquid	liquid	?	liquid	?
22	?	animal	recording media	machine	surface
23	animal	animal	?	liquid	animal
24	metal	metal	?	?	machine
25	glass	?	glass	glass	?
26	fabric/cloth	fabric/cloth	?	fabric/cloth	fabric/cloth
27	people	people	people	people	people
28	liquid	liquid	liquid	liquid	liquid
29	?	animal	animal	animal	animal
30	paper	paper	?	people	paper
31	?	wood	wood	?	machine
32	metal	?	?	metal	?
33	metal	metal	?	?	?
34	?	surface	surface	surface	?

Table F-5: The first stage classification of the object responses from the participants of the 7, 8, 9, and 10 conditions with prior categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03	cat04
1	liquid			
2	machine			
3	door			
4	recording media			
5	liquid			
6	door	metal	fabric/cloth	
7	wood	metal	glass	
8	people	wind		
9	?			
10	metal			
11	wind	liquid		
12	people			
13	metal	plastic		
14	surface	paper		
15	vehicle			
16	metal	wood	surface	
17	glass			
18	container	glass		
19	machine	surface	vehicle	
20	door	animal		
21	liquid			
22	animal	recording media	machine	surface
23	animal	liquid		
24	metal	machine		
25	glass			
26	fabric/cloth			
27	people			
28	liquid			
29	animal			
30	paper			
31	wood	machine		
32	metal			
33	metal			
34	surface			

Table F-6*:* Categories of object responses from participants descriptions of the 7, 8, 9, and 10 conditions with prior categorisation

Number	cat01	cat02	cat03	cat04					*Hcu*	
1	4				-0.2575	0	0	0	0.2575	
2	5				0	0	0	0	0	
3	2				-0.5288	0	0	0	0.5288	
4	1				-0.4644	0	0	0	0.4644	
5	5				0	0	0	0	0	
6	1	1	1		-0.4644	-0.4644	-0.4644	0	1.3932	
7	1	1	1		-0.4644	-0.4644	-0.4644	0	1.3932	**0.577**
8	1	2			-0.4644	-0.5288	0	0	0.9932	
9	0				0	0	0	0	0	
10	1				-0.4644	0	0	0	0.4644	
11	2	2			-0.5288	-0.5288	0	0	1.0575	
12	1				-0.4644	0	0	0	0.4644	
13	4	1			-0.2575	-0.4644	0	0	0.7219	
14	3	1			-0.4422	-0.4644	0	0	0.9066	
15	1				-0.4644	0	0	0	0.4644	**0.634**
16	2	1	1		-0.5288	-0.4644	-0.4644	0	1.4575	
17	4				-0.2575	0	0	0	0.2575	
18	3	1			-0.4422	-0.4644	0	0	0.9066	
19	3	1	1		-0.4422	-0.4644	-0.4644	0	1.3710	
20	2	1			-0.5288	-0.4644	0	0	0.9932	
21	3				-0.4422	0	0	0	0.4422	
22	1	1	1	1	-0.4644	-0.4644	-0.4644	-0.4644	1.8575	
23	3	1			-0.4422	-0.4644	0	0	0.9066	
24	2	1			-0.5288	-0.4644	0	0	0.9932	**1.021**
25	3				-0.4422	0	0	0	0.4422	
26	4				-0.2575	0	0	0	0.2575	
27	5				0	0	0	0	0	
28	5				0	0	0	0	0	
29	4				-0.2575	0	0	0	0.2575	
30	4				-0.2575	0	0	0	0.2575	
31	2	1			-0.5288	-0.4644	0	0	0.9932	
32	2				-0.5288	0	0	0	0.5288	
33	2				-0.5288	0	0	0	0.5288	
34	3				-0.4422	0	0	0	0.4422	**0.371**

Table F-7*:* Analysis of object responses from participants descriptions of the 7, 8, 9, and 10 conditions with prior categorisation

Number	P1-Random	P2-Random	P3-Random	P4-Random	P5-Random
1	?	?	?	?	?
2	?	fire	wind	machine	machine:liquid
3	liquid	liquid	liquid:surface	liquid	liquid
4	liquid	liquid	liquid	liquid:surface	liquid:container
5	metal	metal	metal	metal	metal
6	?	liquid	?	?	?
7	surface	surface	surface	metal	?
8	plastic:glass	plastic:glass	plastic:glass	plastic:glass	liquid
9	furniture	glass:surface	glass	?	surface
10	ceramic	plastic	fire	liquid	ceramic:glass
11	liquid	liquid:surface	liquid	liquid	?
12	liquid	?	surface	plastic	liquid
13	plastic:glass	plastic:glass	plastic:glass	plastic:glass	surface
14	?	fabric/cloth	?	?	?
15	?	?	?	metal:fabric/cloth	metal:fabric/cloth
16	?	liquid	?	liquid	liquid
17	?	wind	liquid:surface	liquid	liquid
18	liquid	?	liquid:surface	?	?
19	?	liquid	?	liquid:surface	?
20	?	?	?	liquid	liquid
21	liquid	?	?	liquid	liquid
22	liquid:people	liquid:people	liquid:people	?	liquid:people
23	?	wood:surface	?	?	?
24	?	?	?	?	?
25	liquid	?	liquid:surface	liquid:surface	liquid:surface
26	?	fire	?	liquid:surface	?
27	?	?	?	liquid	?
28	fire:wood	fire	fire:wood	fire:wood	fire:wood
29	?	fire:wood	?	wood	?
30	?	metal	?	?	?
31	plastic	?	fabric/cloth	?	animal
32	?	?	?	?	?
33	?	?	?	?	?
34	?	?	?	?	?

Table F-8*:* The first stage classification of the object responses from the participants of the 7, 8, 9, and 10 conditions with **no** prior (random) categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03	cat04
1	?			
2	fire	wind	machine	
3	liquid			
4	liquid			
5	metal			
6	liquid			
7	surface	metal		
8	glass	liquid		
9	furniture	glass		
10	ceramic	plastic	fire	liquid
11	liquid			
12	liquid	surface	plastic	
13	plastic	surface		
14	fabric/cloth			
15	metal			
16	liquid			
17	wind	liquid		
18	liquid			
19	liquid			
20	liquid			
21	liquid			
22	people			
23	wood			
24	?			
25	liquid			
26	fire	liquid		
27	liquid			
28	fire			
29	fire			
30	metal			
31	plastic	fabric/cloth	animal	
32	?			
33	?			
34	?			

Table F-9*:* Categories of object responses from participants descriptions of the 7, 8, 9, and 10 conditions with **no** prior categorisation

Number	cat01	cat02	cat03	cat04					*Hcu*	
1	0				0	0	0	0	0	
2	1	1	2		-0.4644	-0.4644	-0.5288	0	1.4575	
3	5				0	0	0	0	0	
4	5				0	0	0	0	0	
5	5				0	0	0	0	0	
6	1				-0.4644	0	0	0	0.4644	
7	3	1			-0.4422	-0.4644	0	0	0.9066	**0.404**
8	4	1			-0.2575	-0.4644	0	0	0.7219	
9	1	2	1		-0.4644	-0.5288	-0.4644	0	1.4575	
10	2	1	1	1	-0.5288	-0.4644	-0.4644	-0.4644	1.9219	
11	4				-0.2575	0	0	0	0.2575	
12	2	1	1		-0.5288	-0.4644	-0.4644	0	1.4575	
13	4	1			-0.2575	-0.4644	0	0	0.7219	
14	1				-0.4644	0	0	0	0.4644	
15	2				-0.5288	0	0	0	0.5288	**0.941**
16	3				-0.4422	0	0	0	0.4422	
17	1	3			-0.4644	-0.4422	0	0	0.9066	
18	2				-0.5288	0	0	0	0.5288	
19	2				-0.5288	0	0	0	0.5288	
20	2				-0.5288	0	0	0	0.5288	
21	3				-0.4422	0	0	0	0.4422	
22	4				-0.2575	0	0	0	0.2575	
23	1				-0.4644	0	0	0	0.4644	
24	0				0	0	0	0	0	**0.455**
25	4				-0.2575	0	0	0	0.2575	
26	1	1			-0.4644	-0.4644	0	0	0.9288	
27	1				-0.4644	0	0	0	0.4644	
28	5				0	0	0	0	0	
29	2				-0.5288	0	0	0	0.5288	
30	1				-0.4644	0	0	0	0.4644	
31	1	1	1		-0.4644	-0.4644	-0.4644	0	1.3932	
32	0				0	0	0	0	0	
33	0				0	0	0	0	0	
34	0				0	0	0	0	0	**0.404**

Table F-10*:* Analysis of object responses from participants descriptions of the 7, 8, 9, and 10 conditions with **no** prior categorisation

Number	P1-Prior	P2-Prior	P3-Prior	P4-Prior	P5-Prior
1	flowing	flowing	?	flowing	flowing
2	cutting	cutting	cutting	cutting	cutting
3	?	?	opening/closing	opening/closing	?
4	?	hitting	?	hitting	?
5	pouring	pouring	pouring	pouring	flowing
6	opening/closing	hitting	hitting	opening/closing	opening/closing
7	?	hitting	dropping	hitting	?
8	brushing	?	?	clicking	rubbing
9	?	?	?	?	?
10	?	clicking	?	?	?
11	boiling	boiling	boiling	?	cooking
12	cutting	?	?	?	?
13	hitting	hitting	hitting	hitting	hitting
14	brushing	rubbing	rubbing	rubbing	rubbing
15	cycling	?	?	?	?
16	hitting	hitting	hitting	hitting	hitting
17	?	rubbing	rubbing	rubbing	?
18	hitting	?	hitting	pouring	pouring
19	cutting	cutting	cutting	brushing	driving
20	?	?	opening/closing	rubbing	opening/closing
21	boiling	boiling	?	boiling	?
22	?	shouting/barking	rubbing	?	walking/running
23	shouting/barking	shouting/barking	?	shouting/barking	shouting/barking
24	hitting	hitting	?	?	cutting
25	hitting	?	hitting	hitting	?
26	cutting	cutting	cutting	opening/closing	cutting
27	crying/laughing	crying/laughing	crying/laughing	crying/laughing	crying/laughing
28	flowing	pouring	flowing	flowing	flowing
29	?	shouting/barking	shouting/barking	shouting/barking	shouting/barking
30	rubbing	rubbing	?	rubbing	rubbing
31	?	walking/running	rubbing	?	opening/closing
32	hitting	clicking	?	hitting	?
33	clicking	hitting	?	?	?
34	?	walking/running	walking/running	walking/running	?

Table F-11: The first stage classification of the action responses from the participants of the 7, 8, 9, and 10 conditions with prior categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03
1	flowing		
2	cutting		
3	opening/closing		
4	hitting		
5	flowing		
6	opening/closing	hitting	
7	hitting	dropping	
8	sweeping	clicking	
9	?		
10	clicking		
11	cooking		
12	cutting		
13	hitting		
14	sweeping		
15	cycling		
16	hitting		
17	rubbing		
18	hitting	pouring	
19	cutting	rubbing	driving
20	opening/closing	rubbing	
21	boiling		
22	shouting/barking	rubbing	walking/running
23	shouting/barking		
24	hitting	cutting	
25	hitting		
26	cutting	opening/closing	
27	crying/laughing		
28	flowing		
29	shouting/barking		
30	rubbing		
31	walking/running	rubbing	opening/closing
32	hitting	clicking	
33	clicking	hitting	
34	walking/running		

Table F-12*:* Categories of action responses from participants descriptions of the 7, 8, 9, and 10 conditions with prior categorisation

Number	cat01	cat02	cat03				*Hcu*	
1	4			-0.2575	0	0	0.2575	
2	5			0	0	0	0	
3	2			-0.5288	0	0	0.5288	
4	2			-0.5288	0	0	0.5288	
5	5			0	0	0	0	
6	3	2		-0.4422	-0.5288	0	0.9710	
7	2	1		-0.5288	-0.4644	0	0.9932	**0.468**
8	2	1		-0.5288	-0.4644	0	0.9932	
9	0			0	0	0	0	
10	1			-0.4644	0	0	0.4644	
11	4			-0.2575	0	0	0.2575	
12	1			-0.4644	0	0	0.4644	
13	5			0	0	0	0	
14	5			0	0	0	0	
15	1			-0.4644	0	0	0.4644	**0.330**
16	5			0	0	0	0	
17	3			-0.4422	0	0	0.4422	
18	2	2		-0.5288	-0.5288	0	1.0575	
19	3	1	1	-0.4422	-0.4644	-0.4644	1.3710	
20	2	1		-0.5288	-0.4644	0	0.9932	
21	3			-0.4422	0	0	0.4422	
22	1	1	1	-0.4644	-0.4644	-0.4644	1.3932	
23	4			-0.2575	0	0	0.2575	
24	2	1		-0.5288	-0.4644	0	0.9932	**0.772**
25	3			-0.4422	0	0	0.4422	
26	4	1		-0.2575	-0.4644	0	0.7219	
27	5			0	0	0	0	
28	5			0	0	0	0	
29	4			-0.2575	0	0	0.2575	
30	4			-0.2575	0	0	0.2575	
31	1	1	1	-0.4644	-0.4644	-0.4644	1.3932	
32	2	1		-0.5288	-0.4644	0	0.9932	
33	1	1		-0.4644	-0.4644	0	0.9288	
34	3			-0.4422	0	0	0.4422	**0.544**

Table F-13*:* Analysis of action responses from participants descriptions of the 7, 8, 9, and 10 conditions with prior categorisation

Number	P1-Random	P2-Random	P3-Random	P4-Random	P5-Random
1	?	?	?	?	?
2	?	burning	burning	flowing	flowing
3	flowing	flowing	flowing	flowing	flowing
4	flowing	flowing	flowing	flowing	hitting
5	rattling	rattling	rattling	rattling	rattling
6	?	boiling	?	?	?
7	rubbing	sweeping/brushing	sweeping/brushing	sweeping/brushing	?
8	rubbing	rubbing	rubbing	rubbing	flowing
9	opening/closing	rubbing	rubbing	?	rubbing
10	opening/closing	opening/closing	cooking	flowing	opening/closing
11	hitting:clicking	flowing	flowing	flowing	?
12	hitting	?	flowing	flowing	pouring
13	rubbing	rubbing	rubbing	rubbing	rubbing
14	?	rattling	?	?	?
15	?	?	?	opening/closing	opening/closing
16	?	flowing	?	flowing	flowing
17	?	blowing	hitting	flowing	flowing
18	hitting	?	hitting	?	?
19	?	flowing	?	flowing	?
20	?	?	?	flowing	hitting
21	hitting	?	?	flowing	blowing
22	walking/running	walking/running	walking/running	?	walking/running
23	?	sweeping/brushing	?	?	?
24	?	?	?	?	?
25	hitting	?	hitting	hitting	hitting
26	?	blowing	?	hitting	?
27	?	?	?	hitting	?
28	burning	burning	burning	burning	burning
29	?	hitting	?	hitting	?
30	?	rattling	?	?	?
31	hitting	?	opening/closing	?	shouting/barking
32	?	?	?	?	?
33	?	?	?	?	?
34	?	?	?	?	?

Table F-14*:* The first stage classification of the action responses from the participants of the 7, 8, 9, and 10 conditions with **no** prior (random) categorisation. P(NUM) represents the participant's identification number.

Number	cat01	cat02	cat03
1	?		
2	burning	flowing	
3	flowing		
4	flowing	hitting	
5	rattling		
6	boiling		
7	rubbing	sweeping/brushing	
8	rubbing	flowing	
9	opening/closing	rubbing	
10	opening/closing	cooking	flowing
11	hitting	flowing	
12	hitting	flowing	
13	rubbing		
14	rattling		
15	opening/closing		
16	flowing		
17	blowing	hitting	flowing
18	hitting		
19	flowing		
20	flowing	hitting	
21	hitting	flowing	blowing
22	walking/running		
23	sweeping/brushing		
24	?		
25	hitting		
26	blowing	hitting	
27	hitting		
28	burning		
29	hitting		
30	rattling		
31	hitting	opening/closing	shouting/barking
32	?		
33	?		
34	?		

Table F-15*:* Categories of action responses from participants descriptions of the 7, 8, 9, and 10 conditions with **no** prior categorisation

Number	cat01	cat02	cat03				*Hcu*	
1	0			0	0	0	0	
2	2	2		-0.5288	-0.5288	0	1.0575	
3	5			0	0	0	0	
4	4	1		-0.2575	-0.4644	0	0.7219	
5	5			0	0	0	0	
6	1			-0.4644	0	0	0.4644	
7	1	3		-0.4644	-0.4422	0	0.9066	**0.450**
8	4	1		-0.2575	-0.4644	0	0.7219	
9	1	3		-0.4644	-0.4422	0	0.9066	
10	3	1	1	-0.4422	-0.4644	-0.4644	1.3710	
11	1	3		-0.4644	-0.4422	0	0.9066	
12	1	3		-0.4644	-0.4422	0	0.9066	
13	5			0	0	0	0	
14	1			-0.4644	0	0	0.4644	
15	2			-0.5288	0	0	0.5288	**0.726**
16	3			-0.4422	0	0	0.4422	
17	1	1	2	-0.4644	-0.4644	-0.5288	1.4575	
18	2			-0.5288	0	0	0.5288	
19	2			-0.5288	0	0	0.5288	
20	1	1		-0.4644	-0.4644	0	0.9288	
21	1	1	1	-0.4644	-0.4644	-0.4644	1.3932	
22	4			-0.2575	0	0	0.2575	
23	1			-0.4644	0	0	0.4644	
24	0			0	0	0	0	**0.667**
25	4			-0.2575	0	0	0.2575	
26	1	1		-0.4644	-0.4644	0	0.9288	
27	1			-0.4644	0	0	0.4644	
28	5			0	0	0	0	
29	2			-0.5288	0	0	0.5288	
30	1			-0.4644	0	0	0.4644	
31	1	1	1	-0.4644	-0.4644	-0.4644	1.3932	
32	0			0	0	0	0	
33	0			0	0	0	0	
34	0			0	0	0	0	**0.404**

Table F-16*:* Analysis of action responses from participants descriptions of the 7, 8, 9, and 10 conditions with **no** prior categorisation

G.1 Appendix G - Data related to the Third Study in Chapter 5

This appendix presents the tasks and the data used for the analysis of the third study presented in Chapter 5. The file names and descriptions of the sounds used are shown in Tables G-1 and G-2 for Pool ‘A’, in Tables G-3 and G-4 for Pool ‘B’, and in Tables G-5 and G-6 for Pool ‘C’. The identification averages of the participant’s descriptors are given in Table G-7. The causal uncertainty results for each of the sounds used in this study is shown in one of ten tables (Tables G-8, G-9, G-10, G-11, G-12, G-13, G-14, G-15, G-16, and G-17).

Task List for Evaluators - Identifying multiple simultaneous everyday sounds

Instructions:

Please familiarise yourself with the idea of multiple simultaneous everyday sounds using the training interface.

Then, using the experimental interface wizard, please perform the tasks that are listed below, please do so in the order presented.

Please remember this is an evaluation of the sounds, not an evaluation of you! Do not feel afraid to comment on the application or on the sounds either positively or negatively.

Scenario:

You've been asked to listen to a set of sounds and write your own description for each of the sounds you heard.

Number of tasks to complete: 7 (seven)
Estimated Time: < 10 minutes

Task List:

1. - 6. Listening to the sounds presented, write your descriptions of the sounds into the textboxes provided in the computer application.

7. Fill in the eight short questions on the questionnaire.

Sound Filename	Object	Action	Length
C7-35.wav	animal	shouting/barking	1.89
Explsn.wav	surface	exploding	2.22
TELEPHON.wav	machine	singing	1.5
374_TwistedLemon_frontdoor_lock.wav	wood	opening/closing	5.19
Pingpong.wav	wood	hitting	3.49
50758_rutgermuller_Snapping_Fingers_www_rutgermuller_nl_short.wav	people	clicking	3.02
Hammerng.wav	metal	hitting	2.36
Gong.wav	metal	hitting:chiming	2.99
Crumpapr.wav	paper	crunching	2.05
B2-32.wav	ceramic	dropping	5.49
BouncingWoodenBall.wav	ball	bouncing	2.15
folding paper.wav	paper	rubbing:tearing	0.78
ELEPHANT.wav	animal	shouting/barking	2.98
Turnpage.wav	paper	sweeping	3.7
41918_duckboy80_DogYelping_LokiEdit_short.wav	animal	shouting/barking	2.8
Frog.wav	animal	shouting/barking	1.37
ROOSTER.wav	animal:countryside	shouting/barking	2.11
42706_K1m218_Toilet.wav	liquid:container	flowing	4.59
SONAR.wav	machine	hitting:chiming	2.21
cork pop.wav	glass	opening/closing	0.53
2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing	3.87
Chicough.wav	people	shouting/barking	2.83
33849_acclivity_NoisyDog_short.wav	animal	shouting/barking	2.06
B1-47.wav	machine	rattling	1.75
YAWN.wav	people	crying/laughing	2.7
MOSQUITO.wav	animal	flying	1.68
changing setting of cooker.wav	machine	rubbing	1.12
34855_jackstrebor_Clock_Ticking_short.wav	wood	hitting	3.38
14579_bjornredtail_typeing2_short.wav	surface	clicking	3.09
DONKEY.wav	animal:countryside	shouting/barking	2.37
Watrbubb.wav	liquid:container	boiling	2.95

Table G-1*:* Pool ‘A’ sounds (part 1 of 2) with descriptions and categories as used in the third study.

Sound Filename	Object	Action	Length
Cymbals.wav	metal	hitting	1.1
Laughing.wav	people	crying/laughing	1.9
32939_sagetyrtle_running_short.wav	people	walking/running	4.56
43382_AGFX_Rooster_chicken_calls_2.wav	animal:countryside	shouting/barking	2.98
Pourwatr.wav	liquid:container	flowing	3.24
B2-66.wav	fabric/cloth	opening/closing	5.07
Cashreg.wav	metal	hitting:chiming	3.4
Glassbrk.wav	glass	breaking	1.18
6174_NoiseCollector_dime.wav	metal:surface	hitting	3.34
spinning coin on table.wav	paper	rubbing:tearing	0.47
cork screw.wav	wood	rubbing	1.93
17502_Jace_Coin_dropping.wav	metal:surface	hitting:rubbing	2.25
ChainRattle.wav	metal	rattling	1.29
KnockingOnDoor.wav	surface:wood	knocking	5.19
GlassWindowBreaking.wav	glass	breaking	1.95
shaking large matchbox.wav	wood	cutting:tearing	1
43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping	2.62
Drill.wav	machine	cutting:tearing	2.77
420_TicTacShutUp_crickets_short.wav	animal	singing	3.03
WIND.wav	wind	sweeping	3.16
20438_AGFX_Water_slosh_spashing_8.wav	liquid:container	splashing	5.37
PlatesInPress.wav	ceramic:surface:wood	hitting:opening/closing	3.67
28303_HerbertBoland_Scissors_short.wav	metal	cutting:tearing	4.76
Monkey.wav	animal	shouting/barking	2.68
shaking coins in palm.wav	metal	rubbing	1.88
22877_Corsica_S_backup_truck_short.wav	vehicle	driving	5.11
Fryfood.wav	container:liquid	boiling	2.95
GlassesClashing.wav	glass	hitting	3.34
singlebellchime.wav	metal	hitting:chiming	5
tearing paper 02.wav	paper	cutting:tearing	0.31
shaking matchbox.wav	wood	rubbing	1.45

Table G-2*:* Pool ‘A’ sounds (part 2 of 2) with descriptions and categories as used in the third study.

Sound Filename	Object	Action	Length
Drums.wav	wood	hitting	1.07
turning pages.wav	paper	rubbing	2.37
B1-56.wav	liquid	flowing	5.04
cutting cardboard03.wav	paper	cutting:tearing	1.53
BABYCRY.wav	people	crying/laughing	1.8
Polsiren.wav	machine	shouting/barking	2.39
LAWNMOWR.wav	machine	sweeping	3.55
33657_Corsica_S_Meow.wav	animal	shouting/barking	4.08
JACKHAMM.wav	surface	hitting:breaking	2.52
dropping one coin on table1.wav	metal:surface	hitting	0.82
COW.wav	animal:countryside	shouting/barking	1.04
metal bowl on concrete.wav	metal:container	rubbing	2.21
4237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	3.85
18339_jppi_Stu_sw_paper_crumple_1.wav	paper	cutting:tearing	3.21
15416_pagancow_Zipper4.wav	metal	opening/closing	3.88
Birds.wav	animal:countryside	singing	1.86
Chrbell.wav	metal	hitting:chiming	3.18
River.wav	liquid:surface	flowing	4.7
9032_MisterTood_Dog_bark2.wav	plastic	singing	2.83
Sandpapr.wav	surface	rubbing	1.35
20732_megamart_mouse_click_short.wav	plastic	clicking	2.98
CARHORN.wav	vehicle	blowing	0.91
57876_dkustic_IkaBird.wav	animal	shouting/barking	3.49
Teakettl.wav	container:liquid	boiling	4.71
filling metal bowl.wav	liquid:container	flowing	2.64
Gunshots.wav	metal	hitting	1.21
Wolf.wav	animal	shouting/barking	3.81
Sawing.wav	wood	cutting:tearing	1.9
Watertiles.wav	liquid:surface	flowing	2.69
15478_elonen_stapler.wav	paper	rubbing:tearing	5.38

Table G-3*:* Pool ‘B’ sounds (part 1 of 2) with descriptions and categories as used in the third study.

Sound Filename	Object	Action	Length
cutting cardboard01.wav	paper	cutting:tearing	3.47
shaking shirt.wav	fabric/cloth	sweeping	2.01
Knocking.wav	wood	knocking	1.65
CUCKOOCL.wav	machine	shouting/barking	3.28
CARCRASH.wav	vehicle	hitting:breaking	2.83
cutting cardboard02.wav	paper	cutting:tearing	1.2
SNORING.wav	people	crying/laughing	3.65
21687_gbling_horn.wav	vehicle	shouting/barking	3.58
32247_ERH_robin_8_short.wav	animal:countryside	singing	4.09
BRUSHTEE.wav	people	rubbing	3.38
TRAIN.wav	vehicle	driving	4.12
Waterfillingglassbottle.wav	glass:liquid	flowing	2.18
26474_osivo_alarm_short.wav	machine	shouting/barking	5
flicking pages.wav	paper	rubbing	1.72
CAMERA.wav	machine	clicking	1.02
B1-78.wav	fabric/cloth	opening/closing	4.31
22694_Erdie_baby2_short.wav	people	crying/laughing	4.74
24640_dobroide_20061030_pigeon_wings_ms_01_short.wav	animal	flying	4.64
Coindrop.wav	metal	hitting	2.16
SawingWood.wav	wood	cutting	2.67
DOORBELL.wav	metal	hitting:chiming	1.79
Baskball.wav	plastic	hitting	2.73
crumbling paper.wav	paper	rubbing:tearing	2.93
17918_WIM_shovel10.wav	surface	cutting:tearing	4.05
Chewing.wav	people	chewing	3.25
BouncingHeavyBall.wav	ball	bouncing	1.95
14245_adcbicycle_50.wav	metal	hitting	3.99
Chickens.wav	animal:countryside	shouting/barking	1.08
16383_JonathanJansen_Metaal_19.wav	metal	hitting	1.5
7383_oyez_dogs_short.wav	animal	shouting/barking	4.59
7803_hanstimm_dieselB.wav	vehicle	driving	4.72
BLINDS.wav	fabric/cloth	opening/closing	2.22

Table G-4*:* Pool ‘B’ sounds (part 2 of 2) with descriptions and categories as used in the third study.

Sound Filename	Object	Action	Length
PINBALL.wav	machine	hitting:chiming	4.33
Crickets.wav	animal:countryside	singing	1.25
B2-27.wav	liquid	flowing	3.17
52226_mookie182_Gum_Chewing_short.wav	people	eating	3.32
Toilet.wav	container:liquid	flowing	4.59
Motorcyc.wav	machine	driving	3.75
closing lid of wheelie bin.wav	plastic	opening/closing	0.71
Boathorn.wav	machine	blowing	2.3
rolling wheelie bin.wav	machine	opening/closing	1.59
25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing	5
p4.wav	animal	shouting/barking	0.97
50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	4.46
35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating	2.7
cigarette lighter01.wav	metal	rubbing	0.67
CANCRUSH.wav	metal	crunching	1.28
B1-40.wav	container:liquid	flowing	5.48
31377_FreqMan_27_coins.wav	metal	hitting:rubbing	3.74
18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking	3.65
389_plagasul_glass_short.wav	glass	hitting:chiming	4.02
BLOWNOSE.wav	people	blowing	1.5
TearingPaper.wav	paper	cutting:tearing	4.01
WHISTLNG.wav	wind	shouting/barking	1.47
1928_RHumphries_rbh_Le_Mans_passby_03.wav	vehicle	driving	5.14
cigarette lighter02.wav	metal	opening/closing	1.76
Whip.wav	surface	hitting	1.28
KnockingDoor.wav	wood:surface	hitting	3.43
Clearthr.wav	people	shouting/barking	1.22
7893_schluppipuppie_bird001.wav	animal	singing	3.65
28113_HerbertBoland_Kukuklok_1_slag_short.wav	machine	shouting/barking	2.48
Seal.wav	animal	shouting/barking	1.39

Table G-5*:* Pool ‘C’ sounds (part 1 of 2) with descriptions and categories as used in the third study.

Sound Filename	Object	Action	Length
VELCRO.wav	fabric/cloth	cutting:tearing	1.29
BreakingGlass.wav	glass	hitting:breaking	2.04
BIKEBELL.wav	vehicle	hitting:chiming	1.34
FIRECRAK.wav	surface	exploding	2.06
C7-34.wav	animal	shouting/barking	3.19
19951_FreqMan_eating_chips_short.wav	people	eating	4.63
RunningUpstairs.wav	surface:room	walking/running	5.02
changing setting of micro.wav	machine	rubbing	2.33
Shuflcrd.wav	paper	sweeping	3.95
cutting paper02.wav	paper	cutting:tearing	1.91
9329_tigersound_pigeon_wings.wav	animal	flying	1.68
striking match.wav	wood	hitting:rubbing	1.12
SHEEP.wav	animal:countryside	shouting/barking	1.03
RAIN.wav	liquid:surface	hitting	3.84
GARGLING.wav	people	crying/laughing	2.48
Dropice.wav	liquid:container	hitting	1.88
LION.wav	animal	shouting/barking	2.28
tearing paper 01.wav	paper	cutting:tearing	1.94
crushing empty drink can.wav	metal	breaking	2.6
Horsegal.wav	animal	walking/running	2.5
ZIPPER.wav	metal	opening/closing	1.13
Cutpaper.wav	paper	cutting:tearing	1.81
42812_KidsCastTechy_Mixing_food_wooden_spoon.wav	glass:wood	hitting	4.76
cutting paper01.wav	paper	cutting:tearing	2.37
Woodpckr.wav	animal	singing	1.29
crumbling tinfoil.wav	metal	rubbing:tearing	2.79
Airplane.wav	vehicle	flying	3.81
375_TwistedLemon_light_switch.wav	surface	sweeping	2.03
51164_rutgermuller_Scissors_Cutting_Air_www.rutgermuller.nl_.wav	metal	cutting:tearing	3.2
42699_K1m218_Chicough.wav	people	crying/laughing	2.83
24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	3.75
writing on paper on table.wav	paper	rubbing	1.51

Table G-6*:* Pool ‘C’ sounds (part 2 of 2) with descriptions and categories as used in the third study.

Participant	3 AI Prior	6 AI Prior	9 AI Prior	3 AI No Prior	6 AI No Prior	9 AI No Prior	Participant Average
p1	2/3	1/2	1/3	0	1/3	4/9	37.96%
p2	1	2/3	5/9	1	5/6	4/9	75.00%
p3	2/3	2/3	5/9	1	1/2	4/9	63.89%
p4	2/3	1/2	5/9	2/3	5/6	4/9	61.11%
p5	1	2/3	5/9	1	2/3	5/9	74.07%
p6	2/3	5/6	2/3	2/3	1/2	4/9	62.96%
p7	2/3	1/2	5/9	2/3	1/2	4/9	55.56%
p8	1	1/2	4/9	2/3	1/2	4/9	59.26%
p9	2/3	1/3	4/9	1	1/2	1/3	54.63%
p10	2/3	1/3	1/9	1/3	1/6	1/3	32.41%
p11	2/3	2/3	4/9	1	1/6	4/9	56.48%
p12	2/3	2/3	4/9	1	2/3	1/3	62.96%
p13	2/3	1/2	2/9	1/3	1/2	1/9	38.89%
p14	1	1/2	5/9	1	1/2	4/9	66.67%
p15	2/3	2/3	5/9	1	5/6	1/3	67.59%
p16	1	2/3	2/3	1/3	1/2	4/9	60.19%
p17	1	2/3	4/9	1	1/2	4/9	67.59%
p18	1	5/6	5/9	1	2/3	4/9	75.00%
p19	2/3	2/3	5/9	2/3	2/3	1/3	59.26%
p20	1/3	1/2	4/9	2/3	2/3	1/3	49.07%
p21	1	2/3	4/9	1/3	1/3	2/9	50.00%
p22	1	2/3	5/9	2/3	1/2	2/9	60.19%
p23	1	2/3	5/9	2/3	1/6	5/9	60.19%
p24	2/3	1/2	1/3	2/3	1/2	1/3	50.00%
p25	2/3	1/2	1/3	1/3	1/2	1/3	44.44%
p26	1	2/3	4/9	2/3	2/3	5/9	66.67%
Condition Average	**79.49%**	**59.62%**	**47.44%**	**70.51%**	**52.56%**	**39.32%**	

Table G-7*:* The results of identifications for participants for the 3, 6, and 9 concurrent Auditory Icons with both prior and no categorisation

Sound Name	3 Prior Hcu Action	3 Prior Hcu Object	6 Prior Hcu Action	6 Prior Hcu Object	9 Prior Hcu Action	9 Prior Hcu Object	Not Heard By	Heard By	Total Hcu
14245_adcbicycle_50.wav			0	0	0	0	2	3	0
14579_bjornredtail_typeing2_short.wav							7	0	0
15416_pagancow_Zipper4.wav	0	0					2	1	0
15478_elonen_stapler.wav	0	0					1	1	0
16383_JonathanJansen_Metaal_19.wav							2	1	0
17502_Jace_Coin_dropping.wav							1	3	0
17918_WIM_shovel10.wav							2	0	0
18339_jppi_Stu_sw_paper_crumple_1.wav							5	1	0
18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav			0	0	0	0	8	2	0
1928_RHumphries_rbh_Le_Mans_passby_03.wav			0	0	0	0	1	5	0
19951_FreqMan_eating_chips_short.wav							1	0	0
20438_AGFX_Water_slosh_spashing_8.wav							0	1	0
20732_megamart_mouse_click_short.wav							3	0	0
21687_gbling_horn.wav							1	1	0
22694_Erdie_baby2_short.wav			0	0	0		0	3	0
22877_Corsica_S_backup_truck_short.wav			0	0	0	0	2	4	0
237_NoiseCollector_soopastarlaugh.wav							0	1	0
24640_dobroide_20061030_pigeon_wings_ms_01_short.wav					0	0	3	1	0
24965_mich3d_BigDogBarking_02.wav					0	0	0	8	0
2502_sdfalk_Car_door_slam_short.wav					0	0	6	3	0
25819_FreqMan_Splash_1_short.wav							3	4	0
26474_osivo_alarm_short.wav	0	0					0	3	0
28113_HerbertBoland_Kukuklok_1_slag_short.wav	0	0					1	2	0
28303_HerbertBoland_Scissors_short.wav							4	1	0
31377_FreqMan_27_coins.wav							2	5	0
32247_ERH_robin_8_short.wav			0	0			0	2	0
32939_sagetyrtle_running_short.wav			0	0			2	4	0
33657_Corsica_S_Meow.wav			0	0	0	0	0	3	0
33849_acclivity_NoisyDog_short.wav	0	0	0	0			1	5	0
34855_jackstrebor_Clock_Ticking_short.wav							3	0	0
35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	0	0	0	0			8	2	0
374_TwistedLemon_frontdoor_lock.wav							1	0	0
375_TwistedLemon_light_switch.wav							4	1	0
389_plagasul_glass_short.wav							8	2	0
41918_duckboy80_DogYelping_LokiEdit_short.wav							1	0	0
420_TicTacShutUp_crickets_short.wav	0	1			0	0	3	3	**1**
4237_NoiseCollector_soopastarlaugh.wav			0	0	0	0	0	6	0
42699_K1m218_Chicough.wav							1	3	0
42706_K1m218_Toilet.wav					0	0	0	3	0
42812_KidsCastTechy_Mixing_food_wooden_spoon.wav			0	0			3	3	0

Table G-8: The results of causal uncertainties of the sounds (part 1 of 10) used in the third study.

Sound Name	3 Random Hcu Action	3 Random Hcu Object	6 Random Hcu Action	6 Random Hcu Object	9 Random Hcu Action	9 Random Hcu Object	Not Heard By	Heard By	Total Hcu
14245_adcbicycle_50.wav			0	0			2	3	0
14579_bjornredtail_typeing2_short.wav							7	0	0
15416_pagancow_Zipper4.wav							2	1	0
15478_elonen_stapler.wav							1	1	0
16383_JonathanJansen_Metaal_19.wav			0	0			2	1	0
17502_Jace_Coin_dropping.wav	0	0					1	3	0
17918_WIM_shovel10.wav							2	0	0
18339_jppi_Stu_sw_paper_crumple_1.wav	0	0					5	1	0
18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav							8	2	0
1928_RHumphries_rbh_Le_Mans_passby_03.wav					0	0	1	5	0
19951_FreqMan_eating_chips_short.wav							1	0	0
20438_AGFX_Water_slosh_spashing_8.wav	0	0					0	1	0
20732_megamart_mouse_click_short.wav							3	0	0
21687_gbling_horn.wav					0	0	1	1	0
22694_Erdie_baby2_short.wav	0	0				0	0	3	0
22877_Corsica_S_backup_truck_short.wav					0	0	2	4	0
237_NoiseCollector_soopastarlaugh.wav	0	0					0	1	0
24640_dobroide_20061030_pigeon_wings_ms_01_short.wav							3	1	0
24965_mich3d_BigDogBarking_02.wav	0	0			0	0	0	8	0
2502_sdfalk_Car_door_slam_short.wav	0	0	0	0			6	3	0
25819_FreqMan_Splash_1_short.wav			0	0	0.918	0.918	3	4	**1.84**
26474_osivo_alarm_short.wav			0	0	0	0	0	3	0
28113_HerbertBoland_Kukuklok_1_slag_short.wav			0	0			1	2	0
28303_HerbertBoland_Scissors_short.wav			0	0			4	1	0
31377_FreqMan_27_coins.wav	0	0	0	0	0	0	2	5	0
32247_ERH_robin_8_short.wav	0	0					0	2	0
32939_sagetyrtle_running_short.wav	0	0					2	4	0
33657_Corsica_S_Meow.wav			0	0			0	3	0
33849_acclivity_NoisyDog_short.wav							1	5	0
34855_jackstrebor_Clock_Ticking_short.wav							3	0	0
35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav							8	2	0
374_TwistedLemon_frontdoor_lock.wav							1	0	0
375_TwistedLemon_light_switch.wav	0	0					4	1	0
389_plagasul_glass_short.wav	0	0	0	0			8	2	0
41918_duckboy80_DogYelping_LokiEdit_short.wav							1	0	0
420_TicTacShutUp_crickets_short.wav							3	3	**1**
4237_NoiseCollector_soopastarlaugh.wav	0	0	0	0			0	6	0
42699_K1m218_Chicough.wav					0	0	1	3	0
42706_K1m218_Toilet.wav	0	0			0	0	0	3	0
42812_KidsCastTechy_Mixing_food_wooden_spoon.wav			0	0			3	3	0

Table G-9: The results of causal uncertainties of the sounds (part 2 of 10) used in the third study.

Sound Name	3 Prior Hcu Action	3 Prior Hcu Object	6 Prior Hcu Action	6 Prior Hcu Object	9 Prior Hcu Action	9 Prior Hcu Object	Not Heard By	Heard By	Total Hcu
43745_gelo_papas_Lighter_Ignition.wav							7	0	0
50092_sunupi_stone_falling_water_short.wav	0	0			0	0	2	8	0
50758_rutgermuller_Snapping_Fingers_www_rutgermuller_nl_short.wav			0	0			2	2	0
51164_rutgermuller_Scissors_Cutting_Air_www.rutgermuller.nl_.wav							1	1	0
52226_mookie182_Gum_Chewing_short.wav	0	0					6	1	0
57876_dkustic_IkaBird.wav	0	0					2	2	0
6174_NoiseCollector_dime.wav							3	1	0
7383_oyez_dogs_short.wav					0	0	1	3	0
7803_hanstimm_dieselB.wav					0	0	0	8	0
7893_schluppipuppie_bird001.wav	0	0	0	0			1	5	0
9032_MisterTood_Dog_bark2.wav	0	0			0	0	4	3	0
9329_tigersound_pigeon_wings.wav					0	0	3	1	0
Airplane.wav					0	0	2	3	0
B1-40.wav							2	2	0
B1-47.wav							2	0	0
B1-56.wav	0	0			0	0	7	3	0
B1-78.wav							4	0	0
B2-27.wav			0	0			4	1	0
B2-32.wav	0	0			0	0	7	4	0
B2-66.wav			0	0			3	2	0
BABYCRY.wav	0	0					0	4	0
Baskball.wav					0	0	3	5	0
BIKEBELL.wav	0	0	0	0	0	0	0	4	0
Birds.wav			0	0			4	2	0
BLINDS.wav							5	0	0
BLOWNOSE.wav			0	0	0	0	1	3	0
Boathorn.wav					0	0	0	1	0
BouncingHeavyBall.wav					0	0	3	1	0
BouncingWoodenBall.wav			0	0			5	1	0
BreakingGlass.wav					0	0	2	2	0
BRUSHTEE.wav							5	0	0

Table G-10: The results of causal uncertainties of the sounds (part 3 of 10) used in the third study.

Sound Name	3 Random Hcu Action	3 Random Hcu Object	6 Random Hcu Action	6 Random Hcu Object	9 Random Hcu Action	9 Random Hcu Object	Not Heard By	Heard By	Total Hcu
43745_gelo_papas_Lighter_Ignition.wav							7	0	0
50092_sunupi_stone_falling_water_short.wav	0.72	0.72					2	8	**1.44**
50758_rutgermuller_Snapping_Fingers_www_rutgermuller_nl_short.wav			0	0			2	2	0
51164_rutgermuller_Scissors_Cutting_Air_www.rutgermuller.nl_.wav	0	0					1	1	0
52226_mookie182_Gum_Chewing_short.wav							6	1	0
57876_dkustic_IkaBird.wav					0	0	2	2	0
6174_NoiseCollector_dime.wav			0	0			3	1	0
7383_oyez_dogs_short.wav			0	0			1	3	0
7803_hanstimm_dieselB.wav			0	0			0	8	0
7893_schluppipuppie_bird001.wav					0	0	1	5	0
9032_MisterTood_Dog_bark2.wav							4	3	0
9329_tigersound_pigeon_wings.wav							3	1	0
Airplane.wav					0	0	2	3	0
B1-40.wav			0	0			2	2	0
B1-47.wav							2	0	0
B1-56.wav	0	0					7	3	0
B1-78.wav							4	0	0
B2-27.wav							4	1	0
B2-32.wav	0	0	0	0			7	4	0
B2-66.wav							3	2	0
BABYCRY.wav			0	0	0	0	0	4	0
Baskball.wav			0	0	0	0	3	5	0
BIKEBELL.wav	0	0					0	4	0
Birds.wav	0	0					4	2	0
BLINDS.wav							5	0	0
BLOWNOSE.wav							1	3	0
Boathorn.wav							0	1	0
BouncingHeavyBall.wav							3	1	0
BouncingWoodenBall.wav							5	1	0
BreakingGlass.wav	0	0					2	2	0
BRUSHTEE.wav							5	0	0

Table G-11: The results of causal uncertainties of the sounds (part 4 of 10) used in the third study.

Sound Name	3 Prior Hcu Action	3 Prior Hcu Object	6 Prior Hcu Action	6 Prior Hcu Object	9 Prior Hcu Action	9 Prior Hcu Object	Not Heard By	Heard By	Total Hcu
C7-34.wav							2	0	0
CAMERA.wav					0	1	2	4	**1**
CANCRUSH.wav					0	0	10	1	0
CARCRASH.wav			0	1			1	2	**1**
CARHORN.wav			0	0	0	0	0	5	0
Cashreg.wav	0	0					1	3	0
ChainRattle.wav	0	0					5	1	0
changing setting of cooker.wav			0	0			4	2	0
changing setting of micro.wav	0	0	0	0	0	0	0	5	0
Chewing.wav	0	0					5	4	0
Chickens.wav							1	2	0
Chicough.wav	0	0					0	1	0
Chrbell.wav			0	0	0	0	1	4	0
cigarette lighter01.wav	0	0					2	1	0
cigarette lighter02.wav					0	0	1	2	0
Clearthr.wav							0	1	0
closing lid of wheelie bin.wav							4	0	0
Coindrop.wav			0	0			4	1	0
cork pop.wav			0	0	0	0	1	2	0
cork screw.wav							3	2	0
COW.wav			0	0	0	0	1	4	0
Crickets.wav	0	0			0	0	2	3	0
crumbling paper.wav	0	0					4	2	0
crumbling tinfoil.wav							1	1	0
Crumpapr.wav							3	1	0
crushing empty drink can.wav							5	1	0
CUCKOOCL.wav							0	3	0
Cutpaper.wav							4	3	0
cutting cardboard01.wav							3	0	0
cutting cardboard02.wav			0	0			5	1	0
cutting cardboard03.wav							2	0	0
cutting paper01.wav							3	0	0
cutting paper02.wav							4	0	0
Cymbals.wav			0	0			1	1	0
DONKEY.wav			0	0			1	1	0
DOORBELL.wav							1	2	0
Drill.wav			0	0			3	2	0
Dropice.wav			0	0			2	4	0
dropping one coin on table1.wav							4	0	0
Drums.wav							1	1	0

Table G-12: The results of causal uncertainties of the sounds (part 5 of 10) used in the third study.

Sound Name	3 Random Hcu Action	3 Random Hcu Object	6 Random Hcu Action	6 Random Hcu Object	9 Random Hcu Action	9 Random Hcu Object	Not Heard By	Heard By	Total Hcu
C7-34.wav							2	0	0
CAMERA.wav					0	0	2	4	**1**
CANCRUSH.wav							10	1	0
CARCRASH.wav							1	2	**1**
CARHORN.wav			0	0	0	0	0	5	0
Cashreg.wav			0	0	0	0	1	3	0
ChainRattle.wav							5	1	0
changing setting of cooker.wav					0	0	4	2	0
changing setting of micro.wav					0	0	0	5	0
Chewing.wav	0	0	0	0			5	4	0
Chickens.wav			0	0			1	2	0
Chicough.wav							0	1	0
Chrbell.wav					0	0	1	4	0
cigarette lighter01.wav							2	1	0
cigarette lighter02.wav							1	2	0
Clearthr.wav					0	0	0	1	0
closing lid of wheelie bin.wav							4	0	0
Coindrop.wav							4	1	0
cork pop.wav							1	2	0
cork screw.wav					0	0	3	2	0
COW.wav			0	0			1	4	0
Crickets.wav					0	0	2	3	0
crumbling paper.wav							4	2	0
crumbling tinfoil.wav	0	0					1	1	0
Crumpapr.wav	0	0					3	1	0
crushing empty drink can.wav					0	0	5	1	0
CUCKOOCL.wav	0	0	0	0			0	3	0
Cutpaper.wav			0	0	1	1	4	3	**2**
cutting cardboard01.wav							3	0	0
cutting cardboard02.wav							5	1	0
cutting cardboard03.wav							2	0	0
cutting paper01.wav							3	0	0
cutting paper02.wav							4	0	0
Cymbals.wav							1	1	0
DONKEY.wav							1	1	0
DOORBELL.wav					0	0	1	2	0
Drill.wav	0	0					3	2	0
Dropice.wav	0	0			0	0	2	4	0
dropping one coin on table1.wav							4	0	0
Drums.wav					0	0	1	1	0

Table G-13: The results of causal uncertainties of the sounds (part 6 of 10) used in the third study.

Sound Name	3 Prior Hcu Action	3 Prior Hcu Object	6 Prior Hcu Action	6 Prior Hcu Object	9 Prior Hcu Action	9 Prior Hcu Object	Not Heard By	Heard By	Total Hcu
ELEPHANT.wav	0	0			0	0	1	7	0
Explsn.wav			0	0	0	0	0	6	0
filling metal bowl.wav	0	0					1	1	0
FIRECRAK.wav	0	0			0	0	3	2	0
flicking pages.wav	0	0					3	2	0
folding paper.wav					0	0	0	1	0
Frog.wav					0	0	1	4	0
Fryfood.wav							1	1	0
GARGLING.wav			0	0	0	0	0	5	0
Glassbrk.wav	0	0					3	2	0
GlassesClashing.wav							1	2	0
GlassWindowBreaking.wav	0	0			0	0	2	3	0
Gong.wav							0	4	0
Gunshots.wav							0	1	0
Hammerng.wav			0	0	0	0	1	5	0
Horsegal.wav					0	0	1	3	0
JACKHAMM.wav			0	0			0	3	0
Knocking.wav			0	0	0	0	1	8	0
KnockingDoor.wav							3	0	0
KnockingOnDoor.wav	0	0					0	3	0
Laughing.wav			0	0	0	0	1	8	0
LAWNMOWR.wav					0	0	4	4	0
LION.wav					0	0	1	4	0
metal bowl on concrete.wav							2	0	0
Monkey.wav			0	0	0	0	1	3	0
MOSQUITO.wav					0	0	1	2	0
Motorcyc.wav	0	0	0	0	0	0	0	5	0
p4.wav	0	0			0	0	0	6	0
PINBALL.wav			0	0			0	4	0
Pingpong.wav					0	0	1	2	0
PlatesInPress.wav							0	1	0
Polsiren.wav	0	0	0	0	0	0	1	10	0
Pourwatr.wav	0	0	0	0			0	8	0
RAIN.wav			0	0			1	6	0
River.wav			0	0	0	0	0	4	0
rolling wheelie bin.wav							4	0	0
ROOSTER.wav			0	0	0	0	0	5	0
RunningUpstairs.wav	0	0					1	1	0

Table G-14: The results of causal uncertainties of the sounds (part 7 of 10) used in the third study.

Sound Name	3 Random Hcu Action	3 Random Hcu Object	6 Random Hcu Action	6 Random Hcu Object	9 Random Hcu Action	9 Random Hcu Object	Not Heard By	Heard By	Total Hcu
ELEPHANT.wav			0	0			1	7	0
Explsn.wav	0	0	0	0	0	0	0	6	0
filling metal bowl.wav							1	1	0
FIRECRAK.wav							3	2	0
flicking pages.wav					0	0	3	2	0
folding paper.wav							0	1	0
Frog.wav			0	0	0	0	1	4	0
Fryfood.wav			0	0			1	1	0
GARGLING.wav					0	0	0	5	0
Glassbrk.wav							3	2	0
GlassesClashing.wav			0	0			1	2	0
GlassWindowBreaking.wav	0	0					2	3	0
Gong.wav	0	0	0	0	0	0	0	4	0
Gunshots.wav			0	0			0	1	0
Hammerng.wav	0	0	0	0			1	5	0
Horsegal.wav	0	0					1	3	0
JACKHAMM.wav			0	0			0	3	0
Knocking.wav			0	0			1	8	0
KnockingDoor.wav							3	0	0
KnockingOnDoor.wav					0	1	0	3	**1**
Laughing.wav					0	0	1	8	0
LAWNMOWR.wav					0	0	4	4	0
LION.wav					0	0	1	4	0
metal bowl on concrete.wav							2	0	0
Monkey.wav					0	0	1	3	0
MOSQUITO.wav					0	0	1	2	0
Motorcyc.wav	0	0					0	5	0
p4.wav			0	0	0	0	0	6	0
PINBALL.wav			0	0.918			0	4	**0.918**
Pingpong.wav	0	0					1	2	0
PlatesInPress.wav					0	0	0	1	0
Polsiren.wav	0	0			0	0	1	10	0
Pourwatr.wav			0	0			0	8	0
RAIN.wav			0	0			1	6	0
River.wav	0	0					0	4	0
rolling wheelie bin.wav							4	0	0
ROOSTER.wav			0	0	0	0	0	5	0
RunningUpstairs.wav							1	1	0

Table G-15: The results of causal uncertainties of the sounds (part 8 of 10) used in the third study.

Sound Name	3 Prior Hcu Action	3 Prior Hcu Object	6 Prior Hcu Action	6 Prior Hcu Object	9 Prior Hcu Action	9 Prior Hcu Object	Not Heard By	Heard By	Total Hcu
Sandpapr.wav					0	0	6	1	0
Sawing.wav					0	0	1	8	0
SawingWood.wav							7	0	0
Seal.wav	0	0					0	3	0
shaking coins in palm.wav	0	0	0	0			2	4	0
shaking large matchbox.wav							5	0	0
shaking matchbox.wav							3	1	0
shaking shirt.wav	0	0					4	1	0
SHEEP.wav	0	0	0	0			1	3	0
Shuflcrd.wav			0	0	0	0	4	2	0
singlebellchime.wav					0	0	2	2	0
SNORING.wav							3	0	0
SONAR.wav	0	0					3	2	0
spinning coin on table.wav							8	1	0
striking match.wav							1	0	0
Teakettl.wav					0	0	2	3	0
tearing paper 01.wav	0	0					2	1	0
tearing paper 02.wav							5	0	0
TearingPaper.wav							1	0	0
TELEPHON.wav	0	0			0	0	0	8	0
Toilet.wav	0	0			0	0	0	8	0
TRAIN.wav							0	1	0
turning pages.wav							3	0	0
Turnpage.wav			0	0	0	0	2	3	0
VELCRO.wav					0	0	6	1	0
Waterfillingglassbottle.wav	0	0	0	0			1	4	0
Watertiles.wav	0	0					2	3	0
Watrbubb.wav			0	0	0	0	1	5	0
Whip.wav							2	2	0
WHISTLNG.wav			0	0			1	1	0
Wolf.wav							2	2	0
Woodpckr.wav			0	0			4	3	0
writing on paper on table.wav							4	1	0
YAWN.wav					0	0	1	1	0
ZIPPER.wav			0	0	0	0	1	6	0

Table G-16: The results of causal uncertainties of the sounds (part 9 of 10) used in the third study.

Sound Name	3 Random Hcu Action	3 Random Hcu Object	6 Random Hcu Action	6 Random Hcu Object	9 Random Hcu Action	9 Random Hcu Object	Not Heard By	Heard By	Total Hcu
Sandpapr.wav							6	1	0
Sawing.wav	0	0	1	1	0	0	1	8	**2**
SawingWood.wav							7	0	0
Seal.wav					0	0	0	3	0
shaking coins in palm.wav			0	0	0	0	2	4	0
shaking large matchbox.wav							5	0	0
shaking matchbox.wav			0	0			3	1	0
shaking shirt.wav							4	1	0
SHEEP.wav							1	3	0
Shuflcrd.wav							4	2	0
singlebellchime.wav					0	0	2	2	0
SNORING.wav							3	0	0
SONAR.wav							3	2	0
spinning coin on table.wav	0	0					8	1	0
striking match.wav							1	0	0
Teakettl.wav	0	0					2	3	0
tearing paper 01.wav							2	1	0
tearing paper 02.wav							5	0	0
TearingPaper.wav							1	0	0
TELEPHON.wav			0	0			0	8	0
Toilet.wav			0	0	0	0	0	8	0
TRAIN.wav			0	0			0	1	0
turning pages.wav							3	0	0
Turnpage.wav					0	0	2	3	0
VELCRO.wav							6	1	0
Waterfillingglassbottle.wav			0	0			1	4	0
Watertiles.wav	0	0					2	3	0
Watrbubb.wav							1	5	0
Whip.wav	0	0	0	0			2	2	0
WHISTLNG.wav							1	1	0
Wolf.wav					0	0	2	2	0
Woodpckr.wav	0	0	0	0			4	3	0
writing on paper on table.wav			0	0			4	1	0
YAWN.wav							1	1	0
ZIPPER.wav			0	0	0	0	1	6	0

Table G-17*:* The results of causal uncertainties of the sounds (part 10 of 10) used in the third study.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
1	6	constrained	b	1	CARCRASH.wav	vehicle	hitting	car crash	Y	1	4
1	6	constrained	b	2	metal bowl on concrete.wav	metal:container	rubbing	uses found sounds	C	0	4
1	6	constrained	b	3	B1-78.wav	fabric/cloth	opening/closing		N	0	4
1	6	constrained	b	4	cutting cardboard02.wav	paper	cutting:tearing	man made	C	1	4
1	6	constrained	b	5	32247_ERH_robin_8_short.wav	animal:countryside	singing	bird tweet	Y	1	4
1	6	constrained	b	6	Watertiles.wav	liquid:surface	flowing		N	0	4
1	6	random	c	1	25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing	sound of liquid pouring or being displaced	Y	1	4
1	6	random	c	2	Cutpaper.wav	paper	cutting:tearing		N	0	4
1	6	random	c	3	closing lid of wheelie bin.wav	plastic	opening/closing	something being poured onto or dropped	C	0	4
1	6	random	c	4	VELCRO.wav	fabric/cloth	cutting:tearing		N	0	4
1	6	random	c	5	CANCRUSH.wav	metal	crunching		N	0	4
1	6	random	c	6	writing on paper on table	paper	rubbing	writing utencil (pencil)	Y	1	4
1	3	constrained	b	1	crumbling paper.wav	paper	rubbing:tearing	easily identifiable	C	1	3
1	3	constrained	b	2	dropping one coin on table1.wav	metal:surface	hitting		N	0	3
1	3	constrained	b	3	Chewing.wav	people	chewing	sounds we generally associate with breakfast and reminds me of some cheesy cereal adverts	Y	1	3
1	9	random	c	1	Airplane.wav	vehicle	flying	air craft	Y	1	6
1	9	random	c	2	ZIPPER.wav	metal	opening/closing	a zipper	Y	1	6
1	9	random	c	3	cutting paper01.wav	paper	cutting:tearing	disturbing	N	0	6
1	9	random	c	4	BLOWNOSE.wav	people	blowing		N	0	6
1	9	random	c	5	writing on paper on table.wav	paper	rubbing		N	0	6
1	9	random	c	6	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	6
1	9	random	c	7	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	a dog	Y	1	6
1	9	random	c	8	Seal.wav	animal	shouting/barking	a seal	Y	1	6
1	9	random	c	9	375_TwistedLemon_light_switch.wav	surface	sweeping		N	0	6
1	3	random	c	1	9329_tigersound_pigeon_wings.wav	animal	flying	schocking	N	0	3
1	3	random	c	2	striking match.wav	wood	hitting	disturbing	N	0	3
1	3	random	c	3	B2-27.wav	liquid	flowing	quite a prolonged latter sound	N	0	3
1	9	constrained	b	1	COW.wav	animal:countryside	shouting/barking	hit and run on a cow	M	1	3
1	9	constrained	b	2	Chrbell.wav	metal	hitting:chiming		N	0	3
1	9	constrained	b	3	4237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	the child seems to think its funny anyway	Y	1	3
1	9	constrained	b	4	Sandpapr.wav	surface	rubbing		N	0	3
1	9	constrained	b	5	Teakettl.wav	container:liquid	boiling		N	0	3
1	9	constrained	b	6	7803_hanstimm_dieselB.wav	vehicle	driving	hit and run on a cow	M	1	3
1	9	constrained	b	7	BLINDS.wav	fabric/cloth	opening/closing		N	0	3
1	9	constrained	b	8	9032_MisterTood_Dog_bark2.wav	plastic	singing	chaotic sounds	N	0	3
1	9	constrained	b	9	Sawing.wav	wood	cutting:tearing		N	0	3

Table G-18: The results from Participant 1 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
2	9	constrained	c	1	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	dog barking	Y	1	5
2	9	constrained	c	2	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating		N	0	5
2	9	constrained	c	3	389_plagasul_glass_short.wav	glass	hitting:chiming		N	0	5
2	9	constrained	c	4	Airplane.wav	vehicle	flying	Airplane landing or taking off	Y	1	5
2	9	constrained	c	5	cutting paper02.wav	paper	cutting:tearing		N	0	5
2	9	constrained	c	6	ZIPPER.wav	metal	opening/closing	radio controlled car	C	1	5
2	9	constrained	c	7	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking	Mouse being clicked	Y	1	5
2	9	constrained	c	8	FIRECRAK.wav	surface	exploding		N	0	5
2	9	constrained	c	9	Boathorn.wav	machine	blowing	ringing a copper bell manually	Y	1	5
2	6	random	b	1	Knocking.wav	wood	knocking	knocking on a door	Y	1	5
2	6	random	b	2	Gunshots.wav	metal	hitting	big kind of gun but more like a hammer hitting wood	Y	1	5
2	6	random	b	3	Baskball.wav	plastic	hitting		N	0	5
2	6	random	b	4	7803_hanstimm_dieselB.wav	vehicle	driving	a car engine as the car leaving	Y	1	5
2	6	random	b	5	CARHORN.wav	vehicle	blowing	old kind of horn you'd have on an old car	Y	1	5
2	6	random	b	6	4237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	baby laughing briefly	Y	1	5
2	3	random	b	1	Chewing.wav	people	chewing	taking a bite of cracker bread	Y	1	3
2	3	random	b	2	Watertiles.wav	liquid:surface	flowing	water hitting hard floor	Y	1	3
2	3	random	b	3	Polsiren.wav	machine	shouting/barking	police siren, like you'd hear in american movies	Y	1	3
2	6	constrained	c	1	GARGLING.wav	people	crying/laughing	Gurgling noise, someone using mouth water	Y	1	4
2	6	constrained	c	2	PINBALL.wav	machine	hitting:chiming	knocking sound / another bell	M	1	4
2	6	constrained	c	3	tearing paper 01.wav	paper	cutting:tearing		N	0	4
2	6	constrained	c	4	crushing empty drink can.wav	metal	breaking		N	0	4
2	6	constrained	c	5	7893_schluppipuppie_bird001.wav	animal	singing	A bird in the forest	Y	1	4
2	6	constrained	c	6	1928_RHumphries_rbh_Le_Mans_passby_03.wav	vehicle	driving	a light motorcycle passing by	Y	1	4
2	9	random	b	1	16383_JonathanJansen_Metaal_19.wav	metal	hitting		N	0	4
2	9	random	b	2	7383_oyez_dogs_short.wav	animal	shouting/barking		N	0	4
2	9	random	b	3	Wolf.wav	animal	shouting/barking	a wolf howling, but it sounds very artificial	Y	1	4
2	9	random	b	4	shaking shirt.wav	fabric/cloth	sweeping		N	0	4
2	9	random	b	5	LAWNMOWR.wav	machine	sweeping	a car leaving. A heavy car with a big engine	C	1	4
2	9	random	b	6	BABYCRY.wav	people	crying/laughing	a babies cry	Y	1	4
2	9	random	b	7	Coindrop.wav	metal	hitting		N	0	4
2	9	random	b	8	26474_osivo_alarm_short.wav	machine	shouting/barking	an alarm clock	Y	1	4
2	9	random	b	9	COW.wav	animal:countryside	shouting/barking		N	0	4
2	3	constrained	c	1	SHEEP.wav	animal:countryside	shouting/barking	a sheep	Y	1	3
2	3	constrained	c	2	52226_mookie182_Gum_Chewing_short.wav	people	eating	somebody chewing	Y	1	3
2	3	constrained	c	3	tearing paper 01.wav	paper	cutting:tearing	sound of paper ripping	Y	1	3

Table G-19*:* The results from Participant 2 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
3	3	constrained	a	1	shaking coins in palm.wav	metal	rubbing	money/keys been shook	Y	1	2
3	3	constrained	a	2	420_TicTacShutUp_crickets_short.wav	animal	singing	Birds sounds	Y	1	2
3	3	constrained	a	3	shaking large matchbox.wav	wood	cutting:tearing		N	0	2
3	9	random	c	1	Cutpaper.wav	paper	cutting:tearing	sawing/scraping noise	Y	1	4
3	9	random	c	2	31377_FreqMan_27_coins.wav	metal	hitting:rubbing	keys/coins jangling	Y	1	4
3	9	random	c	3	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	4
3	9	random	c	4	Toilet.wav	container:liquid	flowing	Toilet being flushed	Y	1	4
3	9	random	c	5	CANCRUSH.wav	metal	crunching		N	0	4
3	9	random	c	6	52226_mookie182_Gum_Chewing_short.wav	people	eating		N	0	4
3	9	random	c	7	7893_schluppipuppie_bird001.wav	animal	singing		N	0	4
3	9	random	c	8	LION.wav	animal	shouting/barking	animal	Y	1	4
3	9	random	c	9	42699_K1m218_Chicough.wav	people	crying/laughing		N	0	4
3	6	constrained	a	1	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	4
3	6	constrained	a	2	spinning coin on table.wav	paper	rubbing:tearing		N	0	4
3	6	constrained	a	3	33849_acclivity_NoisyDog_short.wav	animal	shouting/barking	Dog barking	Y	1	4
3	6	constrained	a	4	Pourwatr.wav	liquid:container	flowing	glass being fillled	Y	1	4
3	6	constrained	a	5	B2-66.wav	fabric/cloth	opening/closing	velco	Y	1	4
3	6	constrained	a	6	32939_sagetyrtle_running_short.wav	people	walking/running	person stomping on wooden floor	Y	1	4
3	3	random	c	1	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	something being dropped in plastic basin of water	Y	1	3
3	3	random	c	2	Horsegal.wav	animal	walking/running	horse galloping	Y	1	3
3	3	random	c	3	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	dog barking	y	1	3
3	6	random	c	1	389_plagasul_glass_short.wav	metal	hitting:chiming	glass being struck	Y	1	3
3	6	random	c	2	B1-40.wav	container:liquid	flowing	liquid being poured	Y	1	3
3	6	random	c	3	crushing empty drink can.wav	metal	breaking		N	0	3
3	6	random	c	4	Woodpckr.wav	animal	singing		N	0	3
3	6	random	c	5	42812_KidsCastTechy_Mixing_food_wooden_spoon.wav	glass:wood	hitting		N	0	3
3	6	random	c	6	RAIN.wav	liquid:surface	hitting	water flowing	C	1	3
3	9	constrained	a	1	B2-32.wav	ceramic	dropping		N	0	5
3	9	constrained	a	2	spinning coin on table.wav	paper	rubbing		N	0	5
3	9	constrained	a	3	Laughing.wav	people	crying/laughing	person laughing	Y	1	5
3	9	constrained	a	4	Hammerng.wav	metal	hitting	hammer being struck	Y	1	5
3	9	constrained	a	5	33849_acclivity_NoisyDog_short.wav	animal	shouting/barking		N	0	5
3	9	constrained	a	6	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing		N	0	5
3	9	constrained	a	7	TELEPHON.wav	machine	singing	phone ringing	Y	1	5
3	9	constrained	a	8	Watrbubb.wav	liquid:container	boiling	water bubbling	Y	1	5
3	9	constrained	a	9	Explsn.wav	surface	exploding	explosion	Y	1	5

Table G-20*:* The results from Participant 3 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
4	3	constrained	c	1	CANCRUSH.wav	metal	crunching	pepples being thrown	S	0	3
4	3	constrained	c	2	BIKEBELL.wav	vehicle	hitting:chiming	bicycle bell (Country context)	Y	1	3
4	3	constrained	c	3	7893_schluppipuppie_bird001.wav	animal	singing	bird singing	Y	1	3
4	6	constrained	c	1	Motorcyc.wav	machine	driving	Motor bike taking off	Y	1	3
4	6	constrained	c	2	52226_mookie182_Gum_Chewing_short.wav	people	eating		N	0	3
4	6	constrained	c	3	Dropice.wav	liquid:container	hitting	something being thrown into a glass jar (pennies or stones)	Y	1	3
4	6	constrained	c	4	SHEEP.wav	animal:countryside	shouting/barking	Sheep bleating	Y	1	3
4	6	constrained	c	5	closing lid of wheelie bin.wav	plastic	opening/closing		N	0	3
4	6	constrained	c	6	375_TwistedLemon_light_switch.wav	surface	sweeping		N	0	3
4	3	random	a	1	17502_Jace_Coin_dropping.wav	metal:surface	hitting:rubbing	bell of some description being rung	C	0	3
4	3	random	a	2	20438_AGFX_Water_slosh_spashing_8.wav	liquid:container	splashing	water being poured	Y	1	3
4	3	random	a	3	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing	car door opening and closing	Y	1	3
4	9	constrained	c	1	p4.wav	animal	shouting/barking	a cat	Y	1	5
4	9	constrained	c	2	cigarette lighter02.wav	metal	opening/closing	a lighter being lit	Y	1	5
4	9	constrained	c	3	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating		N	0	5
4	9	constrained	c	4	RunningUpstairs.wav	surface:room	walking/running		N	0	5
4	9	constrained	c	5	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	something being dropped into water	Y	1	5
4	9	constrained	c	6	VELCRO.wav	fabric/cloth	cutting:tearing	cloth being torn	Y	1	5
4	9	constrained	c	7	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	5
4	9	constrained	c	8	Shuflcrd.wav	paper	sweeping		N	0	5
4	9	constrained	c	9	changing setting of micro.wav	machine	rubbing	a cash register	C	1	5
4	6	random	a	1	Frog.wav	animal	shouting/barking	a frog	Y	1	5
4	6	random	a	2	6174_NoiseCollector_dime.wav	metal:surface	hitting		N	0	5
4	6	random	a	3	ELEPHANT.wav	animal	shouting/barking	an elephant	Y	1	5
4	6	random	a	4	TELEPHON.wav	machine	singing	a phone ringing	Y	1	5
4	6	random	a	5	Pourwatr.wav	liquid:container	flowing	water being poured into a container	Y	1	5
4	6	random	a	6	ROOSTER.wav	animal:countryside	shouting/barking	a cock crowing	Y	1	5
4	9	random	a	1	singlebellchime.wav	metal	hitting:chiming	a small bell	Y	1	4
4	9	random	a	2	shaking coins in palm.wav	metal	rubbing		N	0	4
4	9	random	a	3	374_TwistedLemon_frontdoor_lock.wav	wood	opening/closing		N	0	4
4	9	random	a	4	22877_Corsica_S_backup_truck_short.wav	vehicle	driving	a truck reversing	Y	1	4
4	9	random	a	5	changing setting of cooker.wav	machine	rubbing	a microwave being set	Y	1	4
4	9	random	a	6	YAWN.wav	people	crying/laughing		N	0	4
4	9	random	a	7	50758_rutgermuller_Snapping_Fingers_www_rutgermuller_nl_short.wav	people	clicking		N	0	4
4	9	random	a	8	B2-66.wav	fabric/cloth	opening/closing		N	0	4
4	9	random	a	9	Explsn.wav	surface	exploding	an explosion	Y	1	4

Table G-21: The results from Participant 4 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
5	3	random	c	1	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	water splashing	Y	1	5
5	3	random	c	2	Horsegal.wav	animal	walking/running	horse running	Y	1	5
5	3	random	c	3	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	dog darking	Y	1	5
5	9	constrained	b	1	Sawing.wav	wood	cutting:tearing	sawing	Y	1	6
5	9	constrained	b	2	BLINDS.wav	fabric/cloth	opening/closing		N	0	6
5	9	constrained	b	3	7383_oyez_dogs_short.wav	animal	shouting/barking	dogs barking	Y	1	6
5	9	constrained	b	4	16383_JonathanJansen_Metaal_19.wav	metal	hitting		N	0	6
5	9	constrained	b	5	B1-56.wav	liquid	flowing	steam gush	C	1	6
5	9	constrained	b	6	9032_MisterTood_Dog_bark2.wav	plastic	singing	squeaky toy	C	1	6
5	9	constrained	b	7	flicking pages.wav	paper	rubbing		N	0	6
5	9	constrained	b	8	BouncingHeavyBall.wav	ball	bouncing		N	0	6
5	9	constrained	b	9	LAWNMOWR.wav	machine	sweeping	lawnmower running	Y	1	6
5	6	random	c	1	31377_FreqMan_27_coins.wav	metal	hitting:rubbing	rattling keys/wind chimes	Y	1	4
5	6	random	c	2	B1-40.wav	container:liquid	flowing	pouring water	Y	1	4
5	6	random	c	3	crushing empty drink can.wav	metal	breaking		M	0	4
5	6	random	c	4	Woodpckr.wav	animal	singing		N	0	4
5	6	random	c	5	42812_KidsCastTechy_Mixing_food_wooden_spoon.wav	glass:wood	hitting	thud	M	1	4
5	6	random	c	6	RAIN.wav	liquid:surface	hitting	waterfall/river rapids	C	1	4
5	6	constrained	b	1	CARCRASH.wav	vehicle	hitting:breaking	explosion	Y	1	6
5	6	constrained	b	2	River.wav	liquid:surface	flowing	water flowing	Y	1	6
5	6	constrained	b	3	4237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	laughter	Y	1	6
5	6	constrained	b	4	SawingWood.wav	wood	cutting		N	0	6
5	6	constrained	b	5	33657_Corsica_S_Meow.wav	animal	shouting/barking	bird cheeping	Y	1	6
5	6	constrained	b	6	CAMERA.wav	machine	clicking	opera	C	0	6
5	9	random	c	1	Cutpaper.wav	paper	cutting:tearing	tearing paper	Y	1	6
5	9	random	c	2	31377_FreqMan_27_coins.wav	metal	hitting:rubbing	keys jingling	Y	1	6
5	9	random	c	3	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	6
5	9	random	c	4	Toilet.wav	container:liquid	flowing	flushing water	Y	1	6
5	9	random	c	5	CANCRUSH.wav	metal	crunching	metal chimes	C	0	6
5	9	random	c	6	52226_mookie182_Gum_Chewing_short.wav	people	eating		N	0	6
5	9	random	c	7	7893_schluppipuppie_bird001.wav	animal	singing	birdsong	Y	1	6
5	9	random	c	8	LION.wav	animal	shouting/barking		N	0	6
5	9	random	c	9	42699_K1m218_Chicough.wav	people	crying/laughing	coughing	Y	1	6
5	3	constrained	b	1	B2-32.wav	ceramic	dropping	glass dropping	Y	1	5
5	3	constrained	b	2	Chewing.wav	people	chewing	crunching crisps	Y	1	5
5	3	constrained	b	3	Waterfillingglassbottle.wav	glass:liquid	flowing	water plinking into bottle	Y	1	5

Table G-22*:* The results from Participant 5 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
6	3	constrained	c	1	375_TwistedLemon_light_switch.wav	surface	sweeping		M	0	2
6	3	constrained	c	2	Seal.wav	animal	shouting/barking	seal noises	Y	1	2
6	3	constrained	c	3	cigarette lighter01.wav	metal	rubbing	typewriter keys being pressed	M	1	2
6	9	random	b	1	24640_dobroide_20061030_pigeon_wings_ms_01_short.wav	animal	flying		N	0	6
6	9	random	b	2	21687_gbling_horn.wav	vehicle	shouting/barking	car horn, dukes of hazard like (general lee)	Y	1	6
6	9	random	b	3	BABYCRY.wav	people	crying/laughing	baby crying	Y	1	6
6	9	random	b	4	Watertiles.wav	liquid:surface	flowing		CM	0	6
6	9	random	b	5	Wolf.wav	animal	shouting/barking	wolf howl	N	0	6
6	9	random	b	6	SawingWood.wav	wood	cutting	background hum	CM	0	6
6	9	random	b	7	flicking pages.wav	paper	rubbing	fire / burning like noise	C	1	6
6	9	random	b	8	BLINDS.wav	fabric/cloth	opening/closing		CM	0	6
6	9	random	b	9	Drums.wav	wood	hitting	single drum	Y	1	6
6	6	random	b	1	Chickens.wav	animal:countryside	shouting/barking	cockrel crowing in the morning	N	0	4
6	6	random	b	2	turning pages.wav	paper	rubbing		N	0	4
6	6	random	b	3	Waterfillingglassbottle.wav	glass:liquid	flowing	water filling up an item perhaps glass with funnel top	Y	1	4
6	6	random	b	4	7383_oyez_dogs_short.wav	animal	shouting/barking	dog barking	Y	1	4
6	6	random	b	5	crumbling paper.wav	paper	rubbing:tearing		N	0	4
6	6	random	b	6	CARHORN.wav	vehicle	blowing	car horn	Y	1	4
6	9	constrained	c	1	LION.wav	animal	shouting/barking	lion roar	Y	1	6
6	9	constrained	c	2	BLOWNOSE.wav	people	blowing	someone blowing a nose in hanky	Y	1	6
6	9	constrained	c	3	rolling wheelie bin.wav	machine	opening/closing		N	0	6
6	9	constrained	c	4	BreakingGlass.wav	glass	hitting:breaking	sheet of glass smashing	Y	1	6
6	9	constrained	c	5	FIRECRAK.wav	surface	exploding	popcorn popping or bangers exploding	Y	1	6
6	9	constrained	c	6	B2-27.wav	liquid	flowing		N	0	6
6	9	constrained	c	7	writing on paper on table.wav	paper	rubbing		N	0	6
6	9	constrained	c	8	Airplane.wav	vehicle	flying	airplane taking off overhead	Y	1	6
6	9	constrained	c	9	CANCRUSH.wav	metal	crunching	crushing of a can or tin	Y	1	6
6	6	constrained	c	1	Shuflcrd.wav	paper	sweeping	deck of cards flicked through	Y	1	5
6	6	constrained	c	2	35032_cognito_perceptu_smacking_and ... ... popping_bubblegum_3_short.wav	people	eating	dripping sound	C	1	5
6	6	constrained	c	3	RAIN.wav	liquid:surface	hitting	rain falling in the background	Y	1	5
6	6	constrained	c	4	9329_tigersound_pigeon_wings.wav	animal	flying		N	0	5
6	6	constrained	c	5	1928_RHumphries_rbh_Le_Mans_passby_03.wav	vehicle	driving	fast car passing by (doppler effect)	Y	1	5
6	6	constrained	c	6	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking	mouse click	Y	1	5
6	3	random	b	1	18339_jppi_Stu_sw_paper_crumple_1.wav	paper	cutting:tearing	salt shaker/rice in tin/box	C	1	3
6	3	random	b	2	Polsiren.wav	machine	shouting/barking	police car siren	Y	1	3
6	3	random	b	3	B1-78.wav	fabric/cloth	opening/closing		N	0	3

Table G-23: The results from Participant 6 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
7	6	constrained	a	1	Hammerng.wav	metal	hitting		N	0	4
7	6	constrained	a	2	spinning coin on table.wav	paper	rubbing:tearing		M	0	4
7	6	constrained	a	3	41918_duckboy80_DogYelping_LokiEdit_short.wav	animal	shouting/barking		N	0	4
7	6	constrained	a	4	BouncingWoodenBall.wav	ball	bouncing	doors opening	CM	1	4
7	6	constrained	a	5	Pourwatr.wav	liquid:container	flowing	running water	Y	1	4
7	6	constrained	a	6	Explsn.wav	surface	exploding	explosion	Y	1	4
7	3	constrained	a	1	TELEPHON.wav	machine	singing	phone ringing once continously	Y	1	2
7	3	constrained	a	2	33849_acclivity_NoisyDog_short.wav	animal	shouting/barking	dog barking	Y	1	2
7	3	constrained	a	3	tearing paper 02.wav	paper	cutting:tearing		N	0	2
7	3	random	b	1	4237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	young child laughing	Y	1	2
7	3	random	b	2	cutting cardboard03.wav	paper	cutting:tearing		N	0	2
7	3	random	b	3	Sawing.wav	wood	cutting:tearing	sawing of wood	Y	1	2
7	6	random	b	1	cutting cardboard02.wav	paper	cutting:tearing		N	0	3
7	6	random	b	2	Knocking.wav	wood	knocking	knocking on door	Y	1	3
7	6	random	b	3	CUCKOOCL.wav	machine	shouting/barking	cuckoo clock ringing	Y	1	3
7	6	random	b	4	SNORING.wav	people	crying/laughing		N	0	3
7	6	random	b	5	7803_hanstimm_dieselB.wav	vehicle	driving	sound of a engine, probably diesel	Y	1	3
7	6	random	b	6	BRUSHTEE.wav	people	rubbing		N	0	3
7	9	constrained	a	1	ELEPHANT.wav	animal	shouting/barking	elephant call	Y	1	5
7	9	constrained	a	2	34855_jackstrebor_Clock_Ticking_short.wav	wood	hitting		N	0	5
7	9	constrained	a	3	Turnpage.wav	paper	sweeping	paper sheets being disturbed	Y	1	5
7	9	constrained	a	4	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing		N	0	5
7	9	constrained	a	5	Laughing.wav	people	crying/laughing	man laughing	Y	1	5
7	9	constrained	a	6	Watrbubb.wav	liquid:container	boiling	bubbling liquid	Y	1	5
7	9	constrained	a	7	28303_HerbertBoland_Scissors_short.wav	metal	cutting:tearing		N	0	5
7	9	constrained	a	8	changing setting of cooker.wav	machine	rubbing		N	0	5
7	9	constrained	a	9	B2-32.wav	ceramic	dropping	breaking glass	C	1	5
7	9	random	b	1	57876_dkustic_IkaBird.wav	animal	shouting/barking		N	0	4
7	9	random	b	2	B1-56.wav	liquid	flowing		N	0	4
7	9	random	b	3	LAWNMOWR.wav	machine	sweeping	lawnmower or other engine sound moving off	Y	1	4
7	9	random	b	4	crumbling paper.wav	paper	rubbing:tearing		N	0	4
7	9	random	b	5	CAMERA.wav	machine	clicking	camera shutter sound	Y	1	4
7	9	random	b	6	Birds.wav	animal:countryside	singing		N	0	4
7	9	random	b	7	Coindrop.wav	metal	hitting		N	0	4
7	9	random	b	8	Baskball.wav	plastic	hitting	bouncing of a basketball	Y	1	4
7	9	random	b	9	Polsiren.wav	machine	shouting/barking	siren	Y	1	4

Table G-24*:* The results from Participant 7 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
8	3	constrained	b	1	shaking shirt.wav	fabric/cloth	sweeping	A fire being lit	C	1	3
8	3	constrained	b	2	B1-56.wav	liquid	flowing	Pouring carbonated drink	Y	1	3
8	3	constrained	b	3	26474_osivo_alarm_short.wav	machine	shouting/barking	Alarm signal from clock or small device	Y	1	3
8	9	constrained	b	1	Baskball.wav	plastic	hitting	Plastic ball bouncing	Y	1	4
8	9	constrained	b	2	57876_dkustic_IkaBird.wav	animal	shouting/barking		N	0	4
8	9	constrained	b	3	SawingWood.wav	wood	cutting		N	0	4
8	9	constrained	b	4	Sandpapr.wav	surface	rubbing	Scraping sound	Y	1	4
8	9	constrained	b	5	Chewing.wav	people	chewing		N	0	4
8	9	constrained	b	6	shaking shirt.wav	fabric/cloth	sweeping		N	0	4
8	9	constrained	b	7	Waterfillingglassbottle.wav	glass:liquid	flowing		N	0	4
8	9	constrained	b	8	7803_hanstimm_dieselB.wav	vehicle	driving	Car engine idealing	Y	1	4
8	9	constrained	b	9	CAMERA.wav	machine	clicking	Analog SLR camera taking pics	Y	1	4
8	6	constrained	b	1	JACKHAMM.wav	surface	hitting	Machine gun	C	1	4
8	6	constrained	b	2	Polsiren.wav	machine	shouting/barking	Police siren	Y	1	4
8	6	constrained	b	3	Knocking.wav	wood	knocking		N	0	4
8	6	constrained	b	4	Waterfillingglassbottle.wav	glass:liquid	flowing	pouring water	Y	1	4
8	6	constrained	b	5	flicking pages.wav	paper	rubbing		M	0	4
8	6	constrained	b	6	15416_pagancow_Zipper4.wav	metal	opening/closing	Electronic buzzing	CM	0	4
8	3	random	a	1	42706_K1m218_Toilet.wav	liquid:container	flowing	Toilet flushing	Y	1	2
8	3	random	a	2	spinning coin on table.wav	paper	rubbing	Heavy chain dropped on the ground	CM	1	2
8	3	random	a	3	shaking matchbox.wav	wood	rubbing		M	0	2
8	6	random	a	1	Cashreg.wav	metal	hitting:chiming		N	0	3
8	6	random	a	2	TELEPHON.wav	machine	singing	Phone ringing	Y	1	3
8	6	random	a	3	28303_HerbertBoland_Scissors_short.wav	metal	cutting:tearing	Scissors cutting in heavy fabric and nothing	Y	1	3
8	6	random	a	4	cork screw.wav	wood	rubbing		N	0	3
8	6	random	a	5	14579_bjornredtail_typeing2_short.wav	surface	clicking		N	0	3
8	6	random	a	6	ELEPHANT.wav	animal	shouting/barking	Elephant roar	Y	1	3
8	9	random	a	1	MOSQUITO.wav	animal	flying	Mosquito buzzing	Y	1	4
8	9	random	a	2	22877_Corsica_S_backup_truck_short.wav	vehicle	driving	Sound from heavy machinery	Y	1	4
8	9	random	a	3	SONAR.wav	machine	hitting:chiming		N	0	4
8	9	random	a	4	32939_sagetyrtle_running_short.wav	people	walking/running		N	0	4
8	9	random	a	5	420_TicTacShutUp_crickets_short.wav	animal	singing		N	0	4
8	9	random	a	6	ROOSTER.wav	animal:countryside	shouting/barking	Rooster	Y	1	4
8	9	random	a	7	Frog.wav	animal	shouting/barking	Frog	Y	1	4
8	9	random	a	8	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	4
8	9	random	a	9	tearing paper 02.wav	paper	cutting:tearing		N	0	4

Table G-25: The results from Participant 8 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
9	9	random	a	1	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	4
9	9	random	a	2	6174_NoiseCollector_dime.wav	metal:surface	hitting		N	0	4
9	9	random	a	3	PlatesInPress.wav	ceramic:surface	hittting:opening/closing	Door banging	Y	1	4
9	9	random	a	4	cork pop.wav	glass	opening/closing		N	0	4
9	9	random	a	5	Crumpapr.wav	paper	crunching		N	0	4
9	9	random	a	6	spinning coin on table.wav	paper	rubbing:tearing		N	0	4
9	9	random	a	7	Monkey.wav	animal	shouting/barking	Chicken being sodomised	C	1	4
9	9	random	a	8	Glassbrk.wav	glass	breaking		N	0	4
9	9	random	a	9	Gong.wav	metal	hitting:chiming	Drum or gong	Y	1	4
9	9	constrained	c	1	9329_tigersound_pigeon_wings.wav	animal	flying	A flapping sound or someone running	Y	1	4
9	9	constrained	c	2	25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing		N	0	4
9	9	constrained	c	3	Shuflcrd.wav	paper	sweeping		N	0	4
9	9	constrained	c	4	CANCRUSH.wav	metal	crunching		N	0	4
9	9	constrained	c	5	Motorcyc.wav	machine	driving	Motorbike	Y	1	4
9	9	constrained	c	6	GARGLING.wav	people	crying/laughing	A gasp	Y	1	4
9	9	constrained	c	7	closing lid of wheelie bin.wav	plastic	opening/closing		N	0	4
9	9	constrained	c	8	BIKEBELL.wav	vehicle	hitting:chiming	A chirp	C	1	4
9	9	constrained	c	9	WHISTLNG.wav	wind	shouting/barking		N	0	4
9	6	random	a	1	B2-32.wav	ceramic	dropping		N	0	4
9	6	random	a	2	17502_Jace_Coin_dropping.wav	metal:surface	hitting:rubbing	Footsteps	C	0	4
9	6	random	a	3	cork screw.wav	wood	rubbing		N	0	4
9	6	random	a	4	GlassesClashing.wav	glass	hitting	Glasses tinckling	Y	1	4
9	6	random	a	5	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing	Door opening and closing	Y	1	4
9	6	random	a	6	ELEPHANT.wav	animal	shouting/barking	Elephant	Y	1	4
9	3	constrained	c	1	changing setting of micro.wav	machine	rubbing	Beeping	Y	1	2
9	3	constrained	c	2	FIRECRAK.wav	surface	exploding	Gunshots	Y	1	2
9	3	constrained	c	3	cutting paper01.wav	paper	cutting:tearing		N	0	2
9	3	random	a	1	Hammerng.wav	metal	hitting	Hammering	Y	1	3
9	3	random	a	2	GlassWindowBreaking.wav	glass	breaking	Breaking glass	Y	1	3
9	3	random	a	3	Pingpong.wav	wood	hitting	Something bouncing along the ground	Y	1	3
9	6	constrained	c	1	BLOWNOSE.wav	people	blowing	Gasp/farting sound	M	1	2
9	6	constrained	c	2	51164_rutgermuller_Scissors_Cutting_Air_www.rutgermuller.nl_.wav	metal	cutting:tearing		N	0	2
9	6	constrained	c	3	rolling wheelie bin.wav	machine	opening/closing		M	0	2
9	6	constrained	c	4	Shuflcrd.wav	paper	sweeping		M	0	2
9	6	constrained	c	5	RAIN.wav	liquid:surface	hitting	running water	Y	1	2
9	6	constrained	c	6	Horsegal.wav	animal	walking/running		N	0	2

Table G-26: The results from Participant 9 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
10	3	random	c	1	Motorcyc.wav	machine	driving	motorcycle leaving the engine is a bit sluggish	Y	1	1
10	3	random	c	2	1928_RHumphries_rbh_Le_Mans_passby_03.wav	vehicle	driving		M	0	1
10	3	random	c	3	Shuflcrd.wav	paper	sweeping		N	0	1
10	3	constrained	a	1	ELEPHANT.wav	animal	shouting/barking	elephant	Y	1	2
10	3	constrained	a	2	ChainRattle.wav	metal	rattling		N	0	2
10	3	constrained	a	3	Pourwatr.wav	liquid:container	flowing	water - pissing? or drinking and spilling	Y	1	2
10	6	random	c	1	PINBALL.wav	machine	hitting:chiming	drumming and cash register	Y	1	2
10	6	random	c	2	C7-34.wav	animal	shouting/barking		N	0	2
10	6	random	c	3	389_plagasul_glass_short.wav	glass	hitting:chiming		N	0	2
10	6	random	c	4	Cutpaper.wav	paper	cutting:tearing		N	0	2
10	6	random	c	5	cutting paper02.wav	paper	cutting:tearing		N	0	2
10	6	random	c	6	VELCRO.wav	fabric/cloth	cutting:tearing		N	0	2
10	9	random	c	1	Dropice.wav	liquid:container	hitting		N	0	2
10	9	random	c	2	25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing		N	0	2
10	9	random	c	3	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	cat and dog disturbed in the night	Y	1	2
10	9	random	c	4	389_plagasul_glass_short.wav	glass	hitting:chiming		N	0	2
10	9	random	c	5	Crickets.wav	animal:countryside	singing	cat and dog disturbed in the night	CM	1	2
10	9	random	c	6	VELCRO.wav	fabric/cloth	cutting:tearing		N	0	2
10	9	random	c	7	p4.wav	animal	shouting/barking	cat and dog disturbed in the night	Y	1	2
10	9	random	c	8	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating		N	0	2
10	9	random	c	9	CANCRUSH.wav	metal	crunching		N	0	2
10	6	constrained	a	1	DONKEY.wav	animal:countryside	shouting/barking		N	0	3
10	6	constrained	a	2	Pourwatr.wav	liquid:container	flowing	water	Y	1	3
10	6	constrained	a	3	changing setting of cooker.wav	machine	rubbing		N	0	3
10	6	constrained	a	4	Cymbals.wav	metal	hitting	drum sound	C	1	3
10	6	constrained	a	5	tearing paper 02.wav	paper	cutting:tearing		N	0	3
10	6	constrained	a	6	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing	omph sound	C	0	3
10	9	constrained	a	1	ROOSTER.wav	animal:countryside	shouting/barking	gooooood morning, ducky quack quack	Y	1	1
10	9	constrained	a	2	Glassbrk.wav	glass	breaking		N	0	1
10	9	constrained	a	3	Fryfood.wav	container:liquid	boiling		N	0	1
10	9	constrained	a	4	22877_Corsica_S_backup_truck_short.wav	vehicle	driving		N	0	1
10	9	constrained	a	5	6174_NoiseCollector_dime.wav	metal:surface	hitting		N	0	1
10	9	constrained	a	6	BouncingWoodenBall.wav	ball	bouncing		N	0	1
10	9	constrained	a	7	shaking matchbox.wav	wood	rubbing		N	0	1
10	9	constrained	a	8	Laughing.wav	people	crying/laughing		N	0	1
10	9	constrained	a	9	Drill.wav	machine	cutting:tearing		N	0	1

Table G-27: The results from Participant 10 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
11	6	constrained	b	1	14245_adcbicycle_50.wav	metal	hitting	coin rattling in a glass bottle	Y	1	4
11	6	constrained	b	2	Knocking.wav	wood	knocking	knocking on door	Y	1	4
11	6	constrained	b	3	Birds.wav	animal:countryside	singing		N	0	4
11	6	constrained	b	4	4237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	children laughing	Y	1	4
11	6	constrained	b	5	18339_jppi_Stu_sw_paper_crumple_1.wav	paper	cutting:tearing		N	0	4
11	6	constrained	b	6	Polsiren.wav	machine	shouting/barking	rising falling sound	C	1	4
11	9	constrained	b	1	Sawing.wav	wood	cutting	sawing wood	Y	1	5
11	9	constrained	b	2	BouncingHeavyBall.wav	ball	bouncing		N	0	5
11	9	constrained	b	3	LAWNMOWR.wav	machine	sweeping		N	0	5
11	9	constrained	b	4	7383_oyez_dogs_short.wav	animal	shouting/barking	dog barking	Y	1	5
11	9	constrained	b	5	Teakettl.wav	container:liquid	boiling	kettle boiling	Y	1	5
11	9	constrained	b	6	Sandpapr.wav	surface	rubbing		N	0	5
11	9	constrained	b	7	DOORBELL.wav	metal	hitting:chiming		N	0	5
11	9	constrained	b	8	7803_hanstimm_dieselB.wav	vehicle	driving	car engine driving and in neutral	Y	1	5
11	9	constrained	b	9	Chewing.wav	people	chewing		N	0	5
11	3	random	c	1	Whip.wav	surface	hitting	whip crack	Y	1	4
11	3	random	c	2	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	water dripping	Y	1	4
11	3	random	c	3	31377_FreqMan_27_coins.wav	metal	hitting:rubbing	keys rattling	Y	1	4
11	6	random	c	1	PINBALL.wav	machine	hitting	old door bell or bell sound	C	1	3
11	6	random	c	2	C7-34.wav	animal	shouting/barking		N	0	3
11	6	random	c	3	389_plagasul_glass_short.wav	glass	hitting:chiming		N	0	3
11	6	random	c	4	Cutpaper.wav	paper	cutting:tearing		N	0	3
11	6	random	c	5	cutting paper02.wav	paper	cutting:tearing		N	0	3
11	6	random	c	6	ZIPPER.wav	metal	opening/closing		N	0	3
11	9	random	c	1	Dropice.wav	liquid:container	hitting	tablets dropped into glass	Y	1	4
11	9	random	c	2	25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing	sand falling	CM	1	4
11	9	random	c	3	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	dog	Y	1	4
11	9	random	c	4	B2-27.wav	liquid	flowing		M	0	4
11	9	random	c	5	Crickets.wav	animal:countryside	singing		N	0	4
11	9	random	c	6	VELCRO.wav	fabric/cloth	cutting:tearing		N	0	4
11	9	random	c	7	p4.wav	animal	shouting/barking	cat	Y	1	4
11	9	random	c	8	35032_cognito_perceptu_smacking ... ... _and_popping_bubblegum_3_short.wav	people	eating		N	0	4
11	9	random	c	9	CANCRUSH.wav	metal	crunching		N	0	4
11	3	constrained	b	1	crumbling paper.wav	paper	rubbing		N	0	2
11	3	constrained	b	2	Polsiren.wav	machine	shouting/barking	Police siren	Y	1	2
11	3	constrained	b	3	9032_MisterTood_Dog_bark2.wav	plastic	singing	some kind of small animal making chirping sounds	Y	1	2

Table G-28: The results from Participant 11 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
12	3	constrained	a	1	B2-66.wav	fabric/cloth	opening/closing		N	0	5
12	3	constrained	a	2	Cashreg.wav	metal	hitting:chiming	Bell-like and ringin	Y	1	5
12	3	constrained	a	3	Glassbrk.wav	glass	breaking	crackly	Y	1	5
12	9	random	b	1	15478_elonen_stapler.wav	paper	rubbing:tearing		N	0	5
12	9	random	b	2	cutting cardboard01.wav	paper	cutting:tearing		N	0	5
12	9	random	b	3	Wolf.wav	animal	shouting/barking	sweeping pitch	C	0	5
12	9	random	b	4	Sawing.wav	wood	cutting:tearing	saw like	Y	1	5
12	9	random	b	5	Coindrop.wav	metal	hitting		N	0	5
12	9	random	b	6	SawingWood.wav	wood	cutting		N	0	5
12	9	random	b	7	DOORBELL.wav	metal	hitting:chiming	doorbell like	Y	1	5
12	9	random	b	8	Baskball.wav	plastic	hitting	thumping and bouncing	C	1	5
12	9	random	b	9	17918_WIM_shovel10.wav	surface	cutting:tearing		N	0	5
12	3	random	b	1	River.wav	liquid:surface	flowing	rushing bubbly	M	1	6
12	3	random	b	2	B1-56.wav	liquid	flowing	swooshing	M	1	6
12	3	random	b	3	Birds.wav	animal:countryside	singing	tweeting bird like	Y	1	6
12	6	constrained	a	1	14579_bjornredtail_typeing2_short.wav	surface	clicking	busy	C	0	5
12	6	constrained	a	2	DONKEY.wav	animal:countryside	shouting/barking	neiging	Y	1	5
12	6	constrained	a	3	Watrbubb.wav	liquid:container	boiling	rumbling	Y	1	5
12	6	constrained	a	4	Cymbals.wav	metal	hitting		N	0	5
12	6	constrained	a	5	Laughing.wav	people	crying/laughing	throaty	Y	1	5
12	6	constrained	a	6	22877_Corsica_S_backup_truck_short.wav	vehicle	driving	beeping	Y	1	5
12	6	random	b	1	33657_Corsica_S_Meow.wav	animal	shouting/barking	cat-like call	Y	1	6
12	6	random	b	2	JACKHAMM.wav	surface	hitting:breaking	fast and slow thumping	Y	1	6
12	6	random	b	3	dropping one coin on table1.wav	metal:surface	hitting	fast and slow thumping	M	0	6
12	6	random	b	4	COW.wav	animal:countryside	shouting/barking	mooing	Y	1	6
12	6	random	b	5	metal bowl on concrete.wav	metal:container	rubbing	fast and slow thumping	M	0	6
12	6	random	b	6	Knocking.wav	wood	knocking	knocking	Y	1	6
12	9	constrained	a	1	YAWN.wav	people	crying/laughing	yawning	Y	1	5
12	9	constrained	a	2	MOSQUITO.wav	animal	flying	buzzing	Y	1	5
12	9	constrained	a	3	changing setting of cooker.wav	machine	rubbing		N	0	5
12	9	constrained	a	4	GlassWindowBreaking.wav	glass	breaking	shattering and glass smashing	Y	1	5
12	9	constrained	a	5	Pingpong.wav	wood	hitting	high-knocking	Y	1	5
12	9	constrained	a	6	BouncingWoodenBall.wav	ball	bouncing		N	0	5
12	9	constrained	a	7	Turnpage.wav	paper	sweeping		N	0	5
12	9	constrained	a	8	B2-32.wav	ceramic	dropping		N	0	5
12	9	constrained	a	9	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing		N	0	5

Table G-29: The results from Participant 12 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
13	6	random	b	1	Chickens.wav	animal:countryside	shouting/barking	in rural environment	Y	1	1
13	6	random	b	2	26474_osivo_alarm_short.wav	machine	shouting/barking	morning alarm	Y	1	1
13	6	random	b	3	dropping one coin on table1.wav	metal:surface	hitting		N	0	1
13	6	random	b	4	BRUSHTEE.wav	people	rubbing		N	0	1
13	6	random	b	5	15416_pagancow_Zipper4.wav	metal	opening/closing		N	0	1
13	6	random	b	6	JACKHAMM.wav	surface	hitting	beside to a construction site	Y	1	1
13	3	constrained	c	1	375_TwistedLemon_light_switch.wav	surface	sweeping		N	0	1
13	3	constrained	c	2	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating	a dog drinking from a bowl	C	1	1
13	3	constrained	c	3	28113_HerbertBoland_Kukuklok_1_slag_short.wav	machine	shouting/barking	an old cookoo clock	Y	1	1
13	9	random	b	1	cutting cardboard02.wav	paper	cutting:tearing	a saw felling large trees	M	0	1
13	9	random	b	2	20732_megamart_mouse_click_short.wav	plastic	clicking		N	0	1
13	9	random	b	3	24640_dobroide_20061030_pigeon_wings_ms_01_short.wav	animal	flying		N	0	1
13	9	random	b	4	Sawing.wav	wood	cutting:tearing	a saw felling large trees	M	1	1
13	9	random	b	5	14245_adcbicycle_50.wav	metal	hitting		N	0	1
13	9	random	b	6	B1-78.wav	fabric/cloth	opening/closing		N	0	1
13	9	random	b	7	cutting cardboard01.wav	paper	cutting:tearing	a saw felling large trees	M	0	1
13	9	random	b	8	CARCRASH.wav	vehicle	hitting:breaking	a saw felling large trees	M	0	1
13	9	random	b	9	BLINDS.wav	fabric/cloth	opening/closing		N	0	1
13	3	random	b	1	Teakettl.wav	container:liquid	boiling	Old kettle boiling on a gas hob	Y	1	1
13	3	random	b	2	18339_jppi_Stu_sw_paper_crumple_1.wav	paper	cutting:tearing		M	0	1
13	3	random	b	3	filling metal bowl.wav	liquid:container	flowing		M	0	1
13	6	constrained	c	1	RAIN.wav	liquid:surface	hitting	in the rain	Y	1	1
13	6	constrained	c	2	9329_tigersound_pigeon_wings.wav	animal	flying		N	0	1
13	6	constrained	c	3	1928_RHumphries_rbh_Le_Mans_passby_03.wav	vehicle	driving	with motor bike passing by	Y	1	1
13	6	constrained	c	4	WHISTLNG.wav	wind	shouting/barking	man whistling	Y	1	1
13	6	constrained	c	5	writing on paper on table.wav	paper	rubbing		N	0	1
13	6	constrained	c	6	crushing empty drink can.wav	metal	breaking		N	0	1
13	9	constrained	c	1	VELCRO.wav	fabric/cloth	cutting:tearing		N	0	1
13	9	constrained	c	2	GARGLING.wav	people	crying/laughing	morning bathroom scene with man gurguling	Y	1	1
13	9	constrained	c	3	Toilet.wav	container:liquid	flowing	morning bathroom scene and flushing loo	Y	1	1
13	9	constrained	c	4	FIRECRAK.wav	surface	exploding		N	0	1
13	9	constrained	c	5	SHEEP.wav	animal:countryside	shouting/barking		N	0	1
13	9	constrained	c	6	BreakingGlass.wav	glass	hitting:breaking		N	0	1
13	9	constrained	c	7	Airplane.wav	vehicle	flying		N	0	1
13	9	constrained	c	8	rolling wheelie bin.wav	machine	opening/closing		N	0	1
13	9	constrained	c	9	CANCRUSH.wav	metal	crunching		N	0	1

Table G-30*:* The results from Participant 13 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
14	3	constrained	a	1	Explsn.wav	surface	exploding	explosion bass sounding explosion	Y	1	3
14	3	constrained	a	2	TELEPHON.wav	machine	singing	ringing phone type ringing	Y	1	3
14	3	constrained	a	3	374_TwistedLemon_frontdoor_lock.wav	wood	opening/closing	sound of a window with light clasps closing lightly i.e. by wind	Y	1	3
14	3	random	b	1	Gunshots.wav	metal	hitting	shooting type or banging sound	Y	1	3
14	3	random	b	2	Wolf.wav	animal	shouting/barking	long howl or whine of a dog	Y	1	3
14	3	random	b	3	Sawing.wav	wood	cutting:tearing	hand saw sawing	Y	1	3
14	6	random	b	1	cutting cardboard02.wav	paper	cutting:tearing		N	0	3
14	6	random	b	2	Knocking.wav	wood	knocking	knocking on wood sound by hand	Y	1	3
14	6	random	b	3	CUCKOOCL.wav	machine	shouting/barking	sound from cookoo clock	Y	1	3
14	6	random	b	4	SNORING.wav	people	crying/laughing		N	0	3
14	6	random	b	5	7803_hanstimm_dieselB.wav	vehicle	driving	sound of engine ticking over	Y	1	3
14	6	random	b	6	BRUSHTEE.wav	people	rubbing		N	0	3
14	9	constrained	a	1	ELEPHANT.wav	animal	shouting/barking	elephant roar / sound	Y	1	5
14	9	constrained	a	2	34855_jackstrebor_Clock_Ticking_short.wav	wood	hitting		N	0	5
14	9	constrained	a	3	Turnpage.wav	paper	sweeping		N	0	5
14	9	constrained	a	4	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing	door/boot of car closing/clamping shut	Y	1	5
14	9	constrained	a	5	Laughing.wav	people	crying/laughing	laughter of a mad man	Y	1	5
14	9	constrained	a	6	Watrbubb.wav	liquid:container	boiling	sound of bubbling water	Y	1	5
14	9	constrained	a	7	28303_HerbertBoland_Scissors_short.wav	metal	cutting:tearing		N	0	5
14	9	constrained	a	8	changing setting of cooker.wav	machine	rubbing		N	0	5
14	9	constrained	a	9	B2-32.wav	ceramic	dropping	sound of small broken parts of window glass landing on hard surface	Y	1	5
14	6	constrained	a	1	Hammerng.wav	metal	hitting		N	0	3
14	6	constrained	a	2	spinning coin on table.wav	paper	rubbing:tearing		N	0	3
14	6	constrained	a	3	42706_K1m218_Toilet.wav	liquid:container	flowing	toilet water draining	Y	1	3
14	6	constrained	a	4	BouncingWoodenBall.wav	ball	bouncing	bounce of table tennis ball	Y	1	3
14	6	constrained	a	5	GlassWindowBreaking.wav	glass	breaking		N	0	3
14	6	constrained	a	6	Chicough.wav	people	shouting/barking	a childs cough	Y	1	3
14	9	random	b	1	57876_dkustic_IkaBird.wav	animal	shouting/barking	cricket sound in background	C	1	5
14	9	random	b	2	B1-56.wav	liquid	flowing		N	0	5
14	9	random	b	3	LAWNMOWR.wav	machine	sweeping	maybe a engine starting to purr	Y	1	5
14	9	random	b	4	crumbling paper.wav	paper	rubbing:tearing		N	0	5
14	9	random	b	5	CAMERA.wav	machine	clicking		N	0	5
14	9	random	b	6	Birds.wav	animal:countryside	singing		N	0	5
14	9	random	b	7	Coindrop.wav	metal	hitting	rattle	C	0	5
14	9	random	b	8	Baskball.wav	plastic	hitting	foot ball bouncing	Y	1	5
14	9	random	b	9	Polsiren.wav	machine	shouting/barking	siren of car	Y	1	5

Table G-31: The results from Participant 14 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
15	3	random	a	1	Gong.wav	metal	hitting	Gong clang	Y	1	3
15	3	random	a	2	Crumpapr.wav	paper	crunching	paper crumpling	Y	1	3
15	3	random	a	3	B2-32.wav	ceramic	dropping	Glass breaking	C	1	3
15	3	constrained	c	1	52226_mookie182_Gum_Chewing_short.wav	people	eating	Dropped Trigangle/Spoon	C	0	3
15	3	constrained	c	2	Toilet.wav	container:liquid	flowing	Toilet Flush	Y	1	3
15	3	constrained	c	3	Motorcyc.wav	machine	driving	Motorcycle	Y	1	3
15	9	random	a	1	cork screw.wav	wood	rubbing	Wood rasp/plain	Y	1	4
15	9	random	a	2	17502_Jace_Coin_dropping.wav	metal:surface	hitting		N	0	4
15	9	random	a	3	shaking coins in palm.wav	metal	rubbing	Coins being dropped	Y	1	4
15	9	random	a	4	ChainRattle.wav	metal	rattling		M	0	4
15	9	random	a	5	KnockingOnDoor.wav	surface:wood	knocking	Hammer	Y	1	4
15	9	random	a	6	GlassWindowBreaking.wav	glass	breaking		N	0	4
15	9	random	a	7	shaking large matchbox.wav	wood	cutting:tearing		N	0	4
15	9	random	a	8	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	4
15	9	random	a	9	Drill.wav	machine	cutting:tearing	Hoover/Vacuum	C	0	4
15	6	constrained	c	1	Woodpckr.wav	animal	singing	Woodpecker	Y	1	4
15	6	constrained	c	2	ZIPPER.wav	metal	opening/closing	Drawstring (lawnmower?)	Y	1	4
15	6	constrained	c	3	Dropice.wav	liquid:container	hitting		N	0	4
15	6	constrained	c	4	SHEEP.wav	animal:countryside	shouting/barking	Sheep bleating	Y	1	4
15	6	constrained	c	5	closing lid of wheelie bin.wav	plastic	opening/closing		N	0	4
15	6	constrained	c	6	BLOWNOSE.wav	people	blowing	Noseblowing	Y	1	4
15	6	random	a	1	Frog.wav	animal	shouting/barking	Frog	Y	1	5
15	6	random	a	2	6174_NoiseCollector_dime.wav	metal:surface	hitting	Coin	Y	1	5
15	6	random	a	3	ELEPHANT.wav	animal	shouting/barking	Elephant	Y	1	5
15	6	random	a	4	TELEPHON.wav	machine	singing	Phone	Y	1	5
15	6	random	a	5	50758_rutgermuller_Snapping_Fingers_www_rutgermuller_nl_short.wav	people	clicking	Finger Click	Y	1	5
15	6	random	a	6	singlebellchime.wav	metal	hitting:chiming		N	0	5
15	9	constrained	c	1	B1-47.wav	machine	rattling		N	0	5
15	9	constrained	c	2	cigarette lighter02.wav	metal	opening/closing	Cigarette Lighter	Y	1	5
15	9	constrained	c	3	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating		N	0	5
15	9	constrained	c	4	Horsegal.wav	animal	walking/running	Horse hoofs	Y	1	5
15	9	constrained	c	5	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	Stones being dropped in bucket/water	Y	1	5
15	9	constrained	c	6	VELCRO.wav	fabric/cloth	cutting:tearing		N	0	5
15	9	constrained	c	7	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	5
15	9	constrained	c	8	Shuflcrd.wav	paper	sweeping	Cards being shuffled	Y	1	5
15	9	constrained	c	9	changing setting of micro.wav	machine	rubbing	Beeps	Y	1	5

Table G-32: The results from Participant 15 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
16	3	constrained	b	1	CARHORN.wav	vehicle	blowing	horn from an old car done twice	Y	1	3
16	3	constrained	b	2	filling metal bowl.wav	liquid:container	flowing	water flowing down a drain followed by droplets droplets	Y	1	3
16	3	constrained	b	3	26474_osivo_alarm_short.wav	machine	shouting/barking	beep alarm probably from smoke alarm	Y	1	3
16	9	random	a	1	MOSQUITO.wav	animal	flying	bee sound	Y	1	4
16	9	random	a	2	22877_Corsica_S_backup_truck_short.wav	vehicle	driving	van reversing beep i.e. a warning beep	Y	1	4
16	9	random	a	3	SONAR.wav	machine	hitting:chiming		N	0	4
16	9	random	a	4	32939_sagetyrtle_running_short.wav	people	walking/running		N	0	4
16	9	random	a	5	420_TicTacShutUp_crickets_short.wav	animal	singing		N	0	4
16	9	random	a	6	ROOSTER.wav	animal:countryside	shouting/barking	cock a doodle do a rooster	Y	1	4
16	9	random	a	7	Frog.wav	animal	shouting/barking		N	0	4
16	9	random	a	8	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	4
16	9	random	a	9	42706_K1m218_Toilet.wav	liquid:container	flowing	water flowing in a toilet	Y	1	4
16	6	random	a	1	Cashreg.wav	metal	hitting:chiming	old cash register	Y	1	4
16	6	random	a	2	TELEPHON.wav	machine	singing	old telephone ringing	Y	1	4
16	6	random	a	3	28303_HerbertBoland_Scissors_short.wav	metal	cutting:tearing	someone cutting with a scissors	Y	1	4
16	6	random	a	4	cork screw.wav	wood	rubbing	grinding noise	C	0	4
16	6	random	a	5	tearing paper 02.wav	paper	cutting:tearing		N	0	4
16	6	random	a	6	spinning coin on table.wav	paper	rubbing:tearing		N	0	4
16	9	constrained	b	1	Baskball.wav	plastic	hitting	basketball bouncing	Y	1	6
16	9	constrained	b	2	Polsiren.wav	machine	shouting/barking	police alarm	Y	1	6
16	9	constrained	b	3	SawingWood.wav	wood	cutting		N	0	6
16	9	constrained	b	4	Sandpapr.wav	surface	rubbing	using sanding paper	Y	1	6
16	9	constrained	b	5	Chewing.wav	people	chewing	person eating crisps	Y	1	6
16	9	constrained	b	6	shaking shirt.wav	fabric/cloth	sweeping		N	0	6
16	9	constrained	b	7	B1-56.wav	liquid	flowing		N	0	6
16	9	constrained	b	8	7803_hanstimm_dieselB.wav	vehicle	driving	the engine of a van or large vehicle probably diesel	Y	1	6
16	9	constrained	b	9	CAMERA.wav	machine	clicking	camera taking photos	Y	1	6
16	6	constrained	b	1	JACKHAMM.wav	surface	hitting	not real gunfire from a computer game also contains surround sounds	C	1	4
16	6	constrained	b	2	57876_dkustic_IkaBird.wav	animal	shouting/barking		N	0	4
16	6	constrained	b	3	Knocking.wav	wood	knocking	someone knocking on a door	Y	1	4
16	6	constrained	b	4	Waterfillingglassbottle.wav	glass:liquid	flowing	water flowing	Y	1	4
16	6	constrained	b	5	flicking pages.wav	paper	rubbing		N	0	4
16	6	constrained	b	6	15416_pagancow_Zipper4.wav	metal	opening/closing	someone ziping up a zip	Y	1	4
16	3	random	a	1	B2-66.wav	fabric/cloth	opening/closing	zip from a tent or sleeping bag	Y	1	2
16	3	random	a	2	shaking matchbox.wav	wood	rubbing	shuffling noise	C	0	2
16	3	random	a	3	14579_bjornredtail_typeing2_short.wav	surface	clicking		N	0	2

Table G-33: The results from Participant 16 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
17	6	constrained	b	1	Chrbell.wav	metal	hitting:chiming	bell ringing	Y	1	4
17	6	constrained	b	2	Waterfillingglassbottle.wav	glass:liquid	flowing	water dripping	Y	1	4
17	6	constrained	b	3	COW.wav	animal:countryside	shouting/barking	cow scratching	M	1	4
17	6	constrained	b	4	Sandpapr.wav	surface	rubbing		M	0	4
17	6	constrained	b	5	18339_jppi_Stu_sw_paper_crumple_1.wav	paper	cutting:tearing		N	0	4
17	6	constrained	b	6	22694_Erdie_baby2_short.wav	people	crying/laughing	baby screaming	Y	1	4
17	6	random	c	1	cutting paper01.wav	paper	cutting:tearing		N	0	3
17	6	random	c	2	28113_HerbertBoland_Kukuklok_1_slag_short.wav	machine	shouting/barking	cocuu clock	Y	1	3
17	6	random	c	3	p4.wav	animal	shouting/barking	cat	Y	1	3
17	6	random	c	4	cigarette lighter01.wav	metal	rubbing		N	0	3
17	6	random	c	5	389_plagasul_glass_short.wav	glass	hitting:chiming		N	0	3
17	6	random	c	6	ZIPPER.wav	metal	opening/closing	zipper	Y	1	3
17	3	random	c	1	BreakingGlass.wav	glass	hitting:breaking	glass smashing	Y	1	3
17	3	random	c	2	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	stone dropping in water	Y	1	3
17	3	random	c	3	Woodpckr.wav	animal	singing	hammering	Y	1	3
17	9	constrained	b	1	BLINDS.wav	fabric/cloth	opening/closing		N	0	4
17	9	constrained	b	2	14245_adcbicycle_50.wav	metal	hitting	machine	Y	1	4
17	9	constrained	b	3	Polsiren.wav	machine	shouting/barking	sirens	Y	1	4
17	9	constrained	b	4	Knocking.wav	wood	knocking	knocking	Y	1	4
17	9	constrained	b	5	Chewing.wav	people	chewing		N	0	4
17	9	constrained	b	6	cutting cardboard02.wav	paper	cutting:tearing		N	0	4
17	9	constrained	b	7	24640_dobroide_20061030_pigeon_wings_ms_01_short.wav	animal	flying	bird singing	Y	1	4
17	9	constrained	b	8	B1-56.wav	liquid	flowing		N	0	4
17	9	constrained	b	9	9032_MisterTood_Dog_bark2.wav	plastic	singing		N	0	4
17	9	random	c	1	Dropice.wav	liquid:container	hitting	tablets hitting glass	Y	1	4
17	9	random	c	2	changing setting of micro.wav	machine	rubbing	microwave programming	Y	1	4
17	9	random	c	3	42812_KidsCastTechy_Mixing_food_wooden_spoon.wav	glass:wood	hitting		N	0	4
17	9	random	c	4	25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing		N	0	4
17	9	random	c	5	LION.wav	animal	shouting/barking	lion roar	Y	1	4
17	9	random	c	6	KnockingDoor.wav	wood:surface	knocking		N	0	4
17	9	random	c	7	19951_FreqMan_eating_chips_short.wav	people	eating		N	0	4
17	9	random	c	8	tearing paper 01.wav	paper	cutting:tearing		N	0	4
17	9	random	c	9	1928_RHumphries_rbh_Le_Mans_passby_03.wav	vehicle	driving	plane	C	1	4
17	3	constrained	b	1	15416_pagancow_Zipper4.wav	metal	opening/closing	cloth tearing	C	1	4
17	3	constrained	b	2	15478_elonen_stapler.wav	paper	rubbing	stapler	Y	1	4
17	3	constrained	b	3	Watertiles.wav	liquid:surface	flowing	shower drip	Y	1	4

Table G-34: The results from Participant 17 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
18	3	constrained	a	1	ChainRattle.wav	metal	rattling	Ringing bell small like a dinner bell	Y	1	3
18	3	constrained	a	2	KnockingOnDoor.wav	surface:wood	knocking	person knocking lightly on a door	Y	1	3
18	3	constrained	a	3	GlassWindowBreaking.wav	glass	breaking	Breaking bottle possibly dropped from a height	Y	1	3
18	9	random	c	1	25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing	Bucket of water spilled out	Y	1	4
18	9	random	c	2	50092_sunupi_stone_falling_water_short.wav	liquid:container	flowing		N	0	4
18	9	random	c	3	31377_FreqMan_27_coins.wav	metal	hitting		N	0	4
18	9	random	c	4	Toilet.wav	container:liquid	flowing	Toilet being flushed	Y	1	4
18	9	random	c	5	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	4
18	9	random	c	6	GARGLING.wav	people	crying/laughing	A person gargling a liquid	Y	1	4
18	9	random	c	7	42699_K1m218_Chicough.wav	people	crying/laughing	A girl coughing quickly and sharply	Y	1	4
18	9	random	c	8	Cutpaper.wav	paper	cutting:tearing		N	0	4
18	9	random	c	9	CANCRUSH.wav	metal	crunching		N	0	4
18	6	constrained	a	1	33849_acclivity_NoisyDog_short.wav	animal	shouting/barking	Small dog barking	Y	1	4
18	6	constrained	a	2	Pourwatr.wav	liquid:container	flowing	Somebody filling a glass or something similar with a liquid like water	Y	1	4
18	6	constrained	a	3	changing setting of cooker.wav	machine	rubbing	Four beeps	Y	1	4
18	6	constrained	a	4	shaking large matchbox.wav	wood	cutting:tearing		M	0	4
18	6	constrained	a	5	32939_sagetyrtle_running_short.wav	people	walking/running	Somebody going up a stairs sounding like they are running or bounding up	Y	1	4
18	6	constrained	a	6	shaking coins in palm.wav	metal	rubbing	a jangling sound	Y	1	4
18	3	random	c	1	crumbling tinfoil.wav	metal	rubbing:tearing	Someone emptying items out of a basket like leaves or paper or a rustling of similar items	Y	1	3
18	3	random	c	2	375_TwistedLemon_light_switch.wav	surface	sweeping	A tapping noise like a thin fan on a table not a deep sound but more tinny and sharp	Y	1	3
18	3	random	c	3	51164_rutgermuller_Scissors ... ... _Cutting_Air_www.rutgermuller.nl_.wav	metal	cutting:tearing	The sound of a scissors cutting something light in repeated fashion like hair	Y	1	3
18	6	random	c	1	B1-40.wav	container:liquid	flowing		N	0	4
18	6	random	c	2	389_plagasul_glass_short.wav	glass	hitting		N	0	4
18	6	random	c	3	Woodpckr.wav	animal	singing	Repeated little drumming sound	Y	1	4
18	6	random	c	4	42812_KidsCastTechy_Mixing_food_wooden_spoon.wav	glass:wood	hitting	Tip on glass possibly the end result of the sound above where something lightly chimes on a drinking glass	Y	1	4
18	6	random	c	5	RAIN.wav	liquid:surface	hitting	Rain pour	Y	1	4
18	6	random	c	6	Whip.wav	surface	hitting	Fake gunshot like in an old film	Y	1	4
18	9	constrained	a	1	B2-32.wav	ceramic	dropping		N	0	5
18	9	constrained	a	2	spinning coin on table.wav	paper	rubbing		N	0	5
18	9	constrained	a	3	Laughing.wav	people	crying/laughing	A laugh from a person	Y	1	5
18	9	constrained	a	4	Hammerng.wav	metal	hitting	knocking like on a door	Y	1	5
18	9	constrained	a	5	420_TicTacShutUp_crickets_short.wav	animal	singing	night time crickets	Y	1	5
18	9	constrained	a	6	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing		N	0	5
18	9	constrained	a	7	TELEPHON.wav	machine	singing	An old style phone ring goes off once	Y	1	5
18	9	constrained	a	8	Watrbubb.wav	liquid:container	boiling		N	0	5
18	9	constrained	a	9	Explsn.wav	surface	exploding	A large boom not quite like a bomb but explosive in nature goes off	Y	1	5

Table G-35: The results from Participant 18 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
19	9	random	b	1	9032_MisterTood_Dog_bark2.wav	plastic	singing		N	0	3
19	9	random	b	2	DOORBELL.wav	metal	hitting:chiming	Alarm Sound	C	1	3
19	9	random	b	3	dropping one coin on table1.wav	metal:surface	hitting		N	0	3
19	9	random	b	4	SNORING.wav	people	crying/laughing		N	0	3
19	9	random	b	5	BRUSHTEE.wav	people	rubbing		N	0	3
19	9	random	b	6	57876_dkustic_IkaBird.wav	animal	shouting/barking	Screech Sound	Y	1	3
19	9	random	b	7	LAWNMOWR.wav	machine	sweeping	Engine starting	Y	1	3
19	9	random	b	8	shaking shirt.wav	fabric/cloth	sweeping		N	0	3
19	9	random	b	9	turning pages.wav	paper	rubbing		N	0	3
19	9	constrained	a	1	Frog.wav	animal	shouting/barking	Frog sound	Y	1	5
19	9	constrained	a	2	42706_K1m218_Toilet.wav	liquid:container	flowing	Toilet Flushing	Y	1	5
19	9	constrained	a	3	ChainRattle.wav	metal	rattling		N	0	5
19	9	constrained	a	4	folding paper.wav	paper	rubbing:tearing	Cutterly Sound	C	1	5
19	9	constrained	a	5	14579_bjornredtail_typeing2_short.wav	surface	clicking		N	0	5
19	9	constrained	a	6	GlassesClashing.wav	glass	hitting		N	0	5
19	9	constrained	a	7	BouncingWoodenBall.wav	ball	bouncing		N	0	5
19	9	constrained	a	8	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing	Door been closed	Y	1	5
19	9	constrained	a	9	Laughing.wav	people	crying/laughing	Laughter	Y	1	5
19	6	constrained	a	1	shaking coins in palm.wav	metal	rubbing		N	0	5
19	6	constrained	a	2	Crumpapr.wav	paper	crunching		N	0	5
19	6	constrained	a	3	cork pop.wav	glass	opening/closing	Exploxision	Y	1	5
19	6	constrained	a	4	Monkey.wav	animal	shouting/barking	Monkey stuck in the tree	Y	1	5
19	6	constrained	a	5	Drill.wav	machine	cutting:tearing	Ignition	Y	1	5
19	6	constrained	a	6	50758_rutgermuller_Snapping_Fingers_www_rutgermuller_nl_short.wav	people	clicking	Finger click	Y	1	5
19	3	random	b	1	32247_ERH_robin_8_short.wav	animal:countryside	singing	Bird stranded	Y	1	2
19	3	random	b	2	Watertiles.wav	liquid:surface	flowing	Wind Storm effect	M	1	2
19	3	random	b	3	Teakettl.wav	container:liquid	boiling		M	0	2
19	3	constrained	a	1	Chicough.wav	people	shouting/barking	Child Coughing	Y	1	3
19	3	constrained	a	2	SONAR.wav	machine	hitting:chiming	Machine beeping in hazardous conditions	Y	1	3
19	3	constrained	a	3	420_TicTacShutUp_crickets_short.wav	animal	singing		N	0	3
19	6	random	b	1	BABYCRY.wav	people	crying/laughing	Baby Cring in distress	Y	1	4
19	6	random	b	2	18339_jppi_Stu_sw_paper_crumple_1.wav	paper	cutting:tearing		N	0	4
19	6	random	b	3	Chickens.wav	animal:countryside	shouting/barking	Hens crowing	Y	1	4
19	6	random	b	4	SawingWood.wav	wood	cutting		N	0	4
19	6	random	b	5	COW.wav	animal	shouting/barking	Cow mooing	Y	1	4
19	6	random	b	6	TRAIN.wav	vehicle	driving	Train coming on the tracks	Y	1	4

Table G-36: The results from Participant 19 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
20	6	constrained	b	1	Sawing.wav	wood	cutting:tearing	violin	Y	1	3
20	6	constrained	b	2	21687_gbling_horn.wav	vehicle	shouting/barking		N	0	3
20	6	constrained	b	3	B1-56.wav	liquid	flowing		N	0	3
20	6	constrained	b	4	Coindrop.wav	metal	hitting	coins falling	Y	1	3
20	6	constrained	b	5	Birds.wav	animal:countryside	singing	bird	Y	1	3
20	6	constrained	b	6	turning pages.wav	paper	rubbing		N	0	3
20	3	constrained	b	1	BABYCRY.wav	people	crying/laughing	baby crying	Y	1	1
20	3	constrained	b	2	SawingWood.wav	wood	cutting		N	0	1
20	3	constrained	b	3	20732_megamart_mouse_click_short.wav	plastic	clicking		N	0	1
20	6	random	a	1	ELEPHANT.wav	animal	shouting/barking		N	0	4
20	6	random	a	2	GlassesClashing.wav	glass	hitting	pipes	Y	1	4
20	6	random	a	3	Fryfood.wav	container:liquid	boiling	waterfall	C	1	4
20	6	random	a	4	shaking large matchbox.wav	wood	cutting:tearing		N	0	4
20	6	random	a	5	Explsn.wav	surface	exploding	thunder	Y	1	4
20	6	random	a	6	shaking coins in palm.wav	metal	rubbing	metal clinging	Y	1	4
20	9	random	a	1	SONAR.wav	machine	hitting:chiming		N	0	4
20	9	random	a	2	B2-66.wav	fabric/cloth	opening/closing		N	0	4
20	9	random	a	3	BouncingWoodenBall.wav	ball	bouncing		N	0	4
20	9	random	a	4	B1-47.wav	machine	rattling	whistle	C	0	4
20	9	random	a	5	B2-32.wav	ceramic	dropping		N	0	4
20	9	random	a	6	Laughing.wav	people	crying/laughing	laughter	Y	1	4
20	9	random	a	7	Turnpage.wav	paper	sweeping	paper	Y	1	4
20	9	random	a	8	Monkey.wav	animal	shouting/barking		N	0	4
20	9	random	a	9	Gong.wav	metal	hitting:chiming	gong	Y	1	4
20	9	constrained	b	1	Chrbell.wav	metal	hitting:chiming	church bell	Y	1	5
20	9	constrained	b	2	BouncingHeavyBall.wav	ball	bouncing		N	0	5
20	9	constrained	b	3	CARHORN.wav	vehicle	blowing	car horn	Y	1	5
20	9	constrained	b	4	Polsiren.wav	machine	shouting/barking		N	0	5
20	9	constrained	b	5	cutting cardboard02.wav	paper	cutting:tearing		N	0	5
20	9	constrained	b	6	24640_dobroide_20061030_pigeon_wings_ms_01_short.wav	animal	flying		N	0	5
20	9	constrained	b	7	B1-78.wav	fabric/cloth	opening/closing		N	0	5
20	9	constrained	b	8	River.wav	liquid:surface	flowing	water running or rain	Y	1	5
20	9	constrained	b	9	22694_Erdie_baby2_short.wav	people	crying/laughing	baby crying	Y	1	5
20	3	random	a	1	Drill.wav	machine	cutting:tearing	engine	Y	1	3
20	3	random	a	2	32939_sagetyrtle_running_short.wav	people	walking/running	loud steps	Y	1	3
20	3	random	a	3	spinning coin on table.wav	paper	rubbing:tearing		N	0	3

Table G-37*:* The results from Participant 20 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
21	3	random	a	1	cork screw.wav	wood	rubbing		N	0	1
21	3	random	a	2	17502_Jace_Coin_dropping.wav	metal:surface	hitting	money dropping or falling	Y	1	1
21	3	random	a	3	ChainRattle.wav	metal	rattling		N	0	1
21	6	constrained	b	1	Chrbell.wav	metal	hitting:chiming	bell	Y	1	4
21	6	constrained	b	2	River.wav	liquid:surface	flowing	water	Y	1	4
21	6	constrained	b	3	9032_MisterTood_Dog_bark2.wav	plastic	singing		N	0	4
21	6	constrained	b	4	Knocking.wav	wood	knocking	knocking	Y	1	4
21	6	constrained	b	5	15416_pagancow_Zipper4.wav	metal	opening/closing		N	0	4
21	6	constrained	b	6	CARHORN.wav	vehicle	blowing	car horn	Y	1	4
21	3	constrained	b	1	flicking pages.wav	paper	rubbing	wings flapping	Y	1	3
21	3	constrained	b	2	filling metal bowl.wav	liquid:container	flowing	water	Y	1	3
21	3	constrained	b	3	57876_dkustic_IkaBird.wav	animal	shouting/barking	bird chirping	Y	1	3
21	9	random	a	1	MOSQUITO.wav	animal	flying		N	0	2
21	9	random	a	2	22877_Corsica_S_backup_truck_short.wav	vehicle	driving		N	0	2
21	9	random	a	3	SONAR.wav	machine	hitting:chiming		N	0	2
21	9	random	a	4	32939_sagetyrtle_running_short.wav	people	walking/running		N	0	2
21	9	random	a	5	420_TicTacShutUp_crickets_short.wav	animal	singing		N	0	2
21	9	random	a	6	ROOSTER.wav	animal:countryside	shouting/barking	jarring sound bird sound	Y	1	2
21	9	random	a	7	Frog.wav	animal	shouting/barking		N	0	2
21	9	random	a	8	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	2
21	9	random	a	9	42706_K1m218_Toilet.wav	liquid:container	flowing	water	Y	1	2
21	6	random	a	1	Cashreg.wav	metal	hitting:chiming	cash registrar	Y	1	2
21	6	random	a	2	14579_bjornredtail_typeing2_short.wav	surface	clicking		N	0	2
21	6	random	a	3	28303_HerbertBoland_Scissors_short.wav	metal	cutting:tearing		N	0	2
21	6	random	a	4	shaking matchbox.wav	wood	rubbing	knocking sound	Y	1	2
21	6	random	a	5	tearing paper 02.wav	paper	cutting:tearing		N	0	2
21	6	random	a	6	spinning coin on table.wav	paper	rubbing:tearing		N	0	2
21	9	constrained	b	1	Baskball.wav	plastic	hitting	knock	Y	1	4
21	9	constrained	b	2	Polsiren.wav	machine	shouting/barking	siren	Y	1	4
21	9	constrained	b	3	SawingWood.wav	wood	cutting		N	0	4
21	9	constrained	b	4	Sandpapr.wav	surface	rubbing		N	0	4
21	9	constrained	b	5	Chewing.wav	people	chewing		N	0	4
21	9	constrained	b	6	shaking shirt.wav	fabric/cloth	sweeping		N	0	4
21	9	constrained	b	7	B1-56.wav	liquid	flowing		N	0	4
21	9	constrained	b	8	7803_hanstimm_dieselB.wav	vehicle	driving	car	Y	1	4
21	9	constrained	b	9	CAMERA.wav	machine	clicking	cash registrar	Y	1	4

Table G-38*:* The results from Participant 21 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
22	9	random	b	1	Chrbell.wav	metal	hitting:chiming	church bells	Y	1	2
22	9	random	b	2	B1-56.wav	liquid	flowing		N	0	2
22	9	random	b	3	LAWNMOWR.wav	machine	sweeping		N	0	2
22	9	random	b	4	Coindrop.wav	metal	hitting		N	0	2
22	9	random	b	5	Baskball.wav	plastic	hitting		N	0	2
22	9	random	b	6	crumbling paper.wav	paper	rubbing:tearing		N	0	2
22	9	random	b	7	CAMERA.wav	machine	clicking		N	0	2
22	9	random	b	8	Birds.wav	animal:countryside	singing		N	0	2
22	9	random	b	9	Polsiren.wav	machine	shouting/barking	police cars	Y	1	2
22	6	constrained	a	1	Hammerng.wav	metal	hitting	hammering	Y	1	4
22	6	constrained	a	2	Turnpage.wav	paper	sweeping	paper	Y	1	4
22	6	constrained	a	3	ROOSTER.wav	animal:countryside	shouting/barking	chickens	Y	1	4
22	6	constrained	a	4	BouncingWoodenBall.wav	ball	bouncing		N	0	4
22	6	constrained	a	5	Pourwatr.wav	liquid:container	flowing	water	Y	1	4
22	6	constrained	a	6	14579_bjornredtail_typeing2_short.wav	surface	clicking		N	0	4
22	3	random	b	1	237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	kid laughing	Y	1	2
22	3	random	b	2	cutting cardboard03.wav	paper	cutting:tearing		N	0	2
22	3	random	b	3	Sawing.wav	wood	cutting:tearing	saw	Y	1	2
22	6	random	b	1	cutting cardboard02.wav	paper	cutting:tearing		N	0	3
22	6	random	b	2	Knocking.wav	wood	knocking	knocking	Y	1	3
22	6	random	b	3	CUCKOOCL.wav	machine	shouting/barking	clock	Y	1	3
22	6	random	b	4	SNORING.wav	people	crying/laughing		N	0	3
22	6	random	b	5	7803_hanstimm_dieselB.wav	vehicle	driving	engine	Y	1	3
22	6	random	b	6	BRUSHTEE.wav	people	rubbing		N	0	3
22	9	constrained	a	1	ELEPHANT.wav	animal	shouting/barking	elephants	Y	1	5
22	9	constrained	a	2	34855_jackstrebor_Clock_Ticking_short.wav	wood	hitting		N	0	5
22	9	constrained	a	3	22877_Corsica_S_backup_truck_short.wav	vehicle	driving	machine reversing	Y	1	5
22	9	constrained	a	4	Laughing.wav	people	crying/laughing	laughing	Y	1	5
22	9	constrained	a	5	28303_HerbertBoland_Scissors_short.wav	metal	cutting:tearing		N	0	5
22	9	constrained	a	6	changing setting of cooker.wav	machine	rubbing		N	0	5
22	9	constrained	a	7	cork pop.wav	glass	opening/closing	hiss	C	1	5
22	9	constrained	a	8	Watrbubb.wav	liquid:container	flowing	oil boiling	Y	1	5
22	9	constrained	a	9	B2-32.wav	ceramic	dropping		N	0	5
22	3	constrained	a	1	TELEPHON.wav	machine	singing	phone ringing	Y	1	4
22	3	constrained	a	2	33849_acclivity_NoisyDog_short.wav	animal	shouting/barking	dog barking	Y	1	4
22	3	constrained	a	3	Glassbrk.wav	glass	breaking	glass breaking	Y	1	4

Table G-39: The results from Participant 22 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
23	6	random	c	1	B1-40.wav	container:liquid	flowing		N	0	2
23	6	random	c	2	crushing empty drink can.wav	metal	breaking		N	0	2
23	6	random	c	3	KnockingDoor.wav	wood:surface	hitting		N	0	2
23	6	random	c	4	Toilet.wav	container:liquid	flowing	toilet	Y	1	2
23	6	random	c	5	42812_KidsCastTechy_Mixing_food_wooden_spoon.wav	glass:wood	hitting	pin drop	C	0	2
23	6	random	c	6	RAIN.wav	liquid:surface	hitting		N	0	2
23	3	constrained	a	1	SONAR.wav	machine	hitting:chiming	submarine noise	Y	1	2
23	3	constrained	a	2	420_TicTacShutUp_crickets_short.wav	animal	singing	crickets	Y	1	2
23	3	constrained	a	3	shaking large matchbox.wav	wood	cutting:tearing		N	0	2
23	9	constrained	a	1	B2-32.wav	ceramic	dropping		N	0	5
23	9	constrained	a	2	14579_bjornredtail_typeing2_short.wav	surface	clicking		N	0	5
23	9	constrained	a	3	Laughing.wav	people	crying/laughing	laughing	Y	1	5
23	9	constrained	a	4	Monkey.wav	animal	shouting/barking	monkey	Y	1	5
23	9	constrained	a	5	singlebellchime.wav	metal	hitting:chiming	bell jingling	Y	1	5
23	9	constrained	a	6	2502_sdfalk_Car_door_slam_short.wav	vehicle	opening/closing		N	0	5
23	9	constrained	a	7	TELEPHON.wav	machine	singing	telephone ringing	Y	1	5
23	9	constrained	a	8	Watrbubb.wav	liquid:container	boiling	bubbles	Y	1	5
23	9	constrained	a	9	ChainRattle.wav	metal	rattling		N	0	5
23	3	random	c	1	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting	dunking something into water	Y	1	2
23	3	random	c	2	389_plagasul_glass_short.wav	glass	hitting:chiming	bell or chim noise	Y	1	2
23	3	random	c	3	31377_FreqMan_27_coins.wav	metal	hitting:rubbing		N	0	2
23	6	constrained	a	1	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	4
23	6	constrained	a	2	spinning coin on table.wav	paper	rubbing:tearing		N	0	4
23	6	constrained	a	3	33849_acclivity_NoisyDog_short.wav	animal	shouting/barking	dog barking	Y	1	4
23	6	constrained	a	4	Pourwatr.wav	liquid:container	flowing	filling glass water	Y	1	4
23	6	constrained	a	5	B2-66.wav	fabric/cloth	opening/closing	a zip noise	Y	1	4
23	6	constrained	a	6	32939_sagetyrtle_running_short.wav	people	walking/running	climbing stairs	Y	1	4
23	9	random	c	1	Whip.wav	surface	hitting	whip cracking	N	0	5
23	9	random	c	2	Woodpckr.wav	animal	singing		N	0	5
23	9	random	c	3	Clearthr.wav	people	shouting/barking	someone clearing throat	Y	1	5
23	9	random	c	4	CANCRUSH.wav	metal	crunching		N	0	5
23	9	random	c	5	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	dog bark	Y	1	5
23	9	random	c	6	52226_mookie182_Gum_Chewing_short.wav	people	eating		N	0	5
23	9	random	c	7	7893_schluppipuppie_bird001.wav	animal	singing	bird singing	Y	1	5
23	9	random	c	8	LION.wav	animal	shouting/barking	lion groaning	Y	1	5
23	9	random	c	9	42699_K1m218_Chicough.wav	people	crying/laughing	child noise	Y	1	5

Table G-40: The results from Participant 23 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
24	6	constrained	c	1	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating		N	0	3
24	6	constrained	c	2	B2-27.wav	liquid	flowing	water at start	Y	1	3
24	6	constrained	c	3	Airplane.wav	vehicle	flying		N	0	3
24	6	constrained	c	4	changing setting of micro.wav	machine	rubbing	beep	Y	1	3
24	6	constrained	c	5	42812_KidsCastTechy_Mixing_food_wooden_spoon.wav	glass:wood	hitting	something falling	Y	1	3
24	6	constrained	c	6	FIRECRAK.wav	surface	exploding		N	0	3
24	6	random	b	1	16383_JonathanJansen_Metaal_19.wav	metal	hitting	coins	Y	1	3
24	6	random	b	2	14245_adcbicycle_50.wav	metal	hitting		N	0	3
24	6	random	b	3	cutting cardboard01.wav	paper	cutting:tearing		N	0	3
24	6	random	b	4	Baskball.wav	plastic	hitting	ball bounding	Y	1	3
24	6	random	b	5	17918_WIM_shovel10.wav	surface	cutting:tearing		N	0	3
24	6	random	b	6	Chewing.wav	people	chewing	eating apple	Y	1	3
24	3	constrained	c	1	RunningUpstairs.wav	surface:room	walking/running	beat from synth music	C	1	2
24	3	constrained	c	2	Crickets.wav	animal:countryside	singing	bird	C	1	2
24	3	constrained	c	3	crumbling tinfoil.wav	metal	rubbing:tearing		N	0	2
24	9	random	b	1	Wolf.wav	animal	shouting/barking	hyena call	Y	1	3
24	9	random	b	2	Sandpapr.wav	surface	rubbing		N	0	3
24	9	random	b	3	LAWNMOWR.wav	machine	sweeping		M	0	3
24	9	random	b	4	BRUSHTEE.wav	people	rubbing		N	0	3
24	9	random	b	5	Drums.wav	wood	hitting		N	0	3
24	9	random	b	6	CARHORN.wav	vehicle	blowing	car	M	1	3
24	9	random	b	7	B1-56.wav	liquid	flowing		N	0	3
24	9	random	b	8	20732_megamart_mouse_click_short.wav	plastic	clicking		N	0	3
24	9	random	b	9	CAMERA.wav	machine	clicking	camera photo	Y	1	3
24	3	random	b	1	CUCKOOCL.wav	machine	shouting/barking	cuckoo clock with clock hour chime	Y	1	2
24	3	random	b	2	flicking pages.wav	paper	rubbing		N	0	2
24	3	random	b	3	22694_Erdie_baby2_short.wav	people	crying/laughing	baby cry	Y	1	2
24	9	constrained	c	1	TearingPaper.wav	paper	cutting:tearing		N	0	3
24	9	constrained	c	2	Toilet.wav	container:liquid	flowing	toilet flush	Y	1	3
24	9	constrained	c	3	rolling wheelie bin.wav	machine	opening/closing		N	0	3
24	9	constrained	c	4	Woodpckr.wav	animal	singing		N	0	3
24	9	constrained	c	5	389_plagasul_glass_short.wav	glass	hitting:chiming		N	0	3
24	9	constrained	c	6	GARGLING.wav	people	crying/laughing	gargle	Y	1	3
24	9	constrained	c	7	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	3
24	9	constrained	c	8	1928_RHumphries_rbh_Le_Mans_passby_03.wav	vehicle	driving	plane flying past	Y	1	3
24	9	constrained	c	9	cigarette lighter01.wav	metal	rubbing		N	0	3

Table G-41: The results from Participant 24 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
25	6	random	a	1	Pingpong.wav	wood	hitting		N	0	3
25	6	random	a	2	50758_rutgermuller_Snapping_Fingers_www_rutgermuller_nl_short.wav	people	clicking		N	0	3
25	6	random	a	3	Hammerng.wav	metal	hitting	hammering	Y	1	3
25	6	random	a	4	Gong.wav	metal	hitting:chiming	gong	Y	1	3
25	6	random	a	5	Crumpapr.wav	paper	crunching		N	0	3
25	6	random	a	6	B2-32.wav	ceramic	dropping	breaking glass	Y	1	3
25	9	constrained	c	1	Crickets.wav	animal:countryside	singing	bird chirping	Y	1	3
25	9	constrained	c	2	B2-27.wav	liquid	flowing		N	0	3
25	9	constrained	c	3	52226_mookie182_Gum_Chewing_short.wav	people	eating		N	0	3
25	9	constrained	c	4	Toilet.wav	container:liquid	flowing	flushing water	Y	1	3
25	9	constrained	c	5	Motorcyc.wav	machine	driving	motorbike starting	Y	1	3
25	9	constrained	c	6	writing on paper on table.wav	paper	rubbing		N	0	3
25	9	constrained	c	7	18655_Hell_s_Sound_Guy_MOUSE_CLICKS_short.wav	plastic	clicking		N	0	3
25	9	constrained	c	8	cigarette lighter02.wav	metal	opening/closing		N	0	3
25	9	constrained	c	9	KnockingDoor.wav	wood:surface	hitting		N	0	4
25	9	random	a	1	Cashreg.wav	metal	hitting:chiming	typewriter bell	Y	1	4
25	9	random	a	2	Glassbrk.wav	glass	breaking		N	0	4
25	9	random	a	3	KnockingOnDoor.wav	surface:wood	knocking	knocking on door	Y	1	4
25	9	random	a	4	GlassWindowBreaking.wav	glass	breaking		N	0	4
25	9	random	a	5	43745_gelo_papas_Lighter_Ignition.wav	metal:surface	sweeping		N	0	4
25	9	random	a	6	Drill.wav	machine	cutting:tearing	modem noise	C	0	4
25	9	random	a	7	singlebellchime.wav	metal	hitting:chiming		N	0	4
25	9	random	a	8	tearing paper 02.wav	paper	cutting:tearing		N	0	4
25	9	random	a	9	cork screw.wav	wood	rubbing	sawing	Y	1	4
25	3	constrained	c	1	p4.wav	animal	shouting/barking	cat going meow	Y	1	2
25	3	constrained	c	2	50092_sunupi_stone_falling_water_short.wav	liquid:container	flowing	something falling in water	Y	1	2
25	3	constrained	c	3	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating		N	0	2
25	6	constrained	c	1	34855_jackstrebor_Clock_Ticking_short.wav	wood	hitting		N	0	3
25	6	constrained	c	2	14579_bjornredtail_typeing2_short.wav	surface	clicking		N	0	3
25	6	constrained	c	3	ZIPPER.wav	metal	opening/closing	zipper	Y	1	3
25	6	constrained	c	4	7893_schluppipuppie_bird001.wav	animal	singing	dawn chorus	Y	1	3
25	6	constrained	c	5	28113_HerbertBoland_Kukuklok_1_slag_short.wav	machine	shouting/barking		N	0	3
25	6	constrained	c	6	BIKEBELL.wav	vehicle	hitting	bicycle bell	Y	1	3
25	3	random	a	1	Explsn.wav	surface	exploding	caps /fireworks going off	Y	1	1
25	3	random	a	2	Turnpage.wav	paper	sweeping		N	0	1
25	3	random	a	3	shaking matchbox.wav	wood	rubbing		N	0	1

Table G-42*:* The results from Participant 25 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

Participant	NumStimuli	Condition	Pool	Order	Sound	Object	Action	Description	Annotation	Identified	Descriptors
26	6	constrained	b	1	14245_adcbicycle_50.wav	metal	hitting	marble in a jar	Y	1	4
26	6	constrained	b	2	Knocking.wav	wood	knocking	knocking on door	Y	1	4
26	6	constrained	b	3	Birds.wav	animal:countryside	singing		N	0	4
26	6	constrained	b	4	4237_NoiseCollector_soopastarlaugh.wav	people	crying/laughing	Annoying Child laughing	Y	1	4
26	6	constrained	b	5	B1-56.wav	liquid	flowing		N	0	4
26	6	constrained	b	6	Sawing.wav	wood	cutting:tearing	sawing	Y	1	4
26	3	random	c	1	Dropice.wav	liquid:container	hitting	ice cubes in glass	Y	1	2
26	3	random	c	2	BreakingGlass.wav	glass	hitting:breaking		N	0	2
26	3	random	c	3	BIKEBELL.wav	vehicle	hitting:chiming	bell from a bike	Y	1	2
26	3	constrained	b	1	crumbling paper.wav	paper	rubbing:tearing	rice crispies	Y	1	3
26	3	constrained	b	2	Polsiren.wav	machine	shouting/barking	siren	Y	1	3
26	3	constrained	b	3	9032_MisterTood_Dog_bark2.wav	plastic	singing	dog barking	Y	1	3
26	9	random	c	1	Whip.wav	surface	hitting		N	0	4
26	9	random	c	2	50092_sunupi_stone_falling_water_short.wav	liquid:container	hitting		N	0	4
26	9	random	c	3	25819_FreqMan_Splash_1_short.wav	liquid:surface	splashing	splash	Y	1	4
26	9	random	c	4	24965_mich3d_BigDogBarking_02.wav	animal	shouting/barking	dog	Y	1	4
26	9	random	c	5	Crickets.wav	animal:countryside	singing		N	0	4
26	9	random	c	6	Seal.wav	animal	shouting/barking	seal	Y	1	4
26	9	random	c	7	p4.wav	animal	shouting/barking	cat	Y	1	4
26	9	random	c	8	35032_cognito_perceptu_smacking_and_popping_bubblegum_3_short.wav	people	eating		N	0	4
26	9	random	c	9	crushing empty drink can.wav	metal	breaking	symbols	Y	1	4
26	9	constrained	b	1	18339_jppi_Stu_sw_paper_crumple_1.wav	paper	cutting:tearing		N	0	4
26	9	constrained	b	2	BouncingHeavyBall.wav	ball	bouncing	knocking	M	1	4
26	9	constrained	b	3	LAWNMOWR.wav	machine	sweeping		N	0	4
26	9	constrained	b	4	33657_Corsica_S_Meow.wav	animal	shouting/barking	cat	Y	1	4
26	9	constrained	b	5	Teakettl.wav	container:liquid	boiling	kettle	Y	1	4
26	9	constrained	b	6	Sandpapr.wav	surface	rubbing		N	0	4
26	9	constrained	b	7	Baskball.wav	plastic	hitting		M	0	4
26	9	constrained	b	8	7803_hanstimm_dieselB.wav	vehicle	driving	car	Y	1	4
26	9	constrained	b	9	Chewing.wav	people	chewing		N	0	4
26	6	random	c	1	PINBALL.wav	machine	hitting:chiming	ships bell	Y	1	4
26	6	random	c	2	31377_FreqMan_27_coins.wav	metal	hitting:rubbing	change rattling	Y	1	4
26	6	random	c	3	389_plagasul_glass_short.wav	glass	hitting:chiming	knocking	C	0	4
26	6	random	c	4	Cutpaper.wav	paper	cutting:tearing	scrubbing	Y	1	4
26	6	random	c	5	cutting paper02.wav	paper	cutting:tearing		N	0	4
26	6	random	c	6	ZIPPER.wav	metal	opening/closing	zip	Y	1	4

Table G-43: The results from Participant 26 for the 3, 6, and 9 concurrent Auditory Icons conditions with and without prior classification.

H.1 Appendix H - Chapter 6 - First study - repertory grid technique analysis

This appendix holds the grid data, the task list, the figures, and the charts for the Chapter 6, where the participants' tacit classifications of sounds were explored using the repertory grid technique (RGT). This method elicits their constructs and their associations, with regard to the set of sounds presented to them. Understanding the participants' tacit knowledge with regard to the meaning they give to a sound and to sounds in an auditory scene gives a better understanding of an everyday sound scene.

Participant ID	Pairs similarity	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	Singles description
1	flowing sounds	1	4	4	3	1*	4	4	5*	5	1*	1	1	rubbing sounds
1	water sounds	3	5	5*	5	4	3	3	3	2	1	1*	1*	machine sounds
1	mechanical/electrical sounds	2	1*	1	1*	1	3	5*	4	4	3	3	3	paper sounds
1	kitchen sounds	1*	3	1	3	1	5*	2	2	1*	3	1	1	knocking sounds
2	Progressive / Flowing	1*	3	1*	5*	1	5	4	2	2	1	4	1	Repetitive
2	Pointed / Focused	4	2	3	1	5	1	1*	1*	4	5	1	5*	Continuous / Unbroken
2	Tapping / Dripping / Nagging	5	2	2	1	5*	1*	1	3	4	4	1*	5	Hissy / Steamy / Flowing
2	In motion / Bubbly / Bobbing	2	1*	2	4	3	4	5	5	5*	1*	4	1	Touching / Tangential
3	Artifical Watering	3	2	2	5*	1*	5	4	4	1*	2	5	1	Timekeeping
3	Water	5*	3	4	3	1	3	3	3	4	1*	2	1*	Nordic Food
3	Notetaking	3	2	2	5	4	5*	1*	1*	4	3	5	4	Visiting someones office
3	Workshop	1	1*	1*	5	3	2	2	2	3	4	5*	3	Sounds from an empty home
4	Constructed Sounds	1	2	3	5	1*	5	5	1*	4	5*	2	5	Easy Recognisable
4	Concrete Actions	2	2	1*	5	5	1	1	1	1*	5	5*	5	Background / General sound
4	Real	5*	3	2	1	5	1*	1*	2	1	1	5	1	Imaginary
4	Noise	3	1*	1	1*	1	2	2	3	2	5	3	5*	Pleasant sound
5	Shaky sounds	2	4	1*	5	1	5*	3	3	1*	3	5	4	Percussive
5	Danger	2	1*	2	1*	3	3	5*	5	5	5	5	5	No Harm
5	Water / Nature	4	5	5	5	4	5	5	5*	5	1	1*	1*	Humankind thing
5	Air	1*	3	2	3	1*	3	3	3	3	5*	5	5	Water

Fig. H-1*:* Participants Construct Pole and Ranking Results (* indicates sound is part of the original triplet played to the participant, which they used to form the constructs).

Task List

Task List for Evaluators - Auditory Icons for User Identification
Evaluators Instructions:
Using the iTunes application, please perform the tasks that are listed below, please do so in the order presented. The observer will be present and be available to answer any questions you may have or to record any comments that you would like to contribute. Please remember this is an evaluation of the sounds, not an evaluation of you! Do not feel afraid to comment on the application or sounds either positively or negatively, as the goal of the experiment is to gather a user's descriptions of the sounds.

Scenario:
You've been asked to listen to a set of sounds and write your own descriptions/labels for the sounds. These sounds are intended for use as user identifier sounds, where a single sound will be used as an auditory/sonic identifier for a particular person. When writing your descriptions or labels, try to keep them short, a word or phrase rather than a long sentence is preferable.

In this experiment, you will be asked to:

- Listen to triplets of sounds and group two of the three together to form one category and group the third sound into a second but distinct category.
- Rank each of the sounds within the categories you will have created whilst listening to the triplets.
- Rank the sounds in order of your favourites, where the chosen favourite sound would be used as your own auditory/sonic identifier.

Number of tasks to complete: 9 (nine)

Task List:

START OF PILOT STAGE

1. In iTunes, Select the folder "Experiment Pilot" and then select & play the sounds in the playlist "Pilot Comparisons: 1".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

2. In iTunes, Select the folder "Experiment Pilot" and then select the playlist "Alt Pilot Comparisons: 1".

Listen to the sounds presented and rank them within the two categories you created in the previous task.

END OF PILOT STAGE

(continued on next page)

START OF EXPERIMENT - PART A

3. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 1".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

4. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 2".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

5. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 3".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

6. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 4".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

END OF EXPERIMENT - PART A
If you wish to take a break, please feel free to do so.

(continued on next page)

START OF EXPERIMENT - PART B

7. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 1".

Listen to the sounds presented and rank them within the two categories you created in the task 3, which dealt with "Experiment Part A" and the playlist "Comparisons: 1".

8. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 2".

Listen to the sounds presented and rank them within the two categories you created in the task 4, which dealt with "Experiment Part A" and the playlist "Comparisons: 2".

9. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 3".

Listen to the sounds presented and rank them within the two categories you created in the task 5, which dealt with "Experiment Part A" and the playlist "Comparisons: 3".

10. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 4".

Listen to the sounds presented and rank them within the two categories you created in the task 6, which dealt with "Experiment Part A" and the playlist "Comparisons: 4".

END OF EXPERIMENT - PART B

Thank you for your cooperation and assistance in running this experiment. The data and results will be available as soon as the study is finished. No personal, private or confidential information will be included in the results or any published works derived from this study. If you have not already returned your consent form to the observer, please do so. If you have any other comments or observations you wish to share with the observer, please do so, otherwise your time and effort has been appreciated and will be a big help in making this study a success, thank you again.

The RGT method combines the elicitation of descriptors and the rating of the descriptors to produce a large number of descriptors. Principal-component analysis (PCA) (Jolliffe, 2002), multidimensional scaling (MDS) (Borg and Groenen, 1996), and cluster analysis (CA) (Everitt and Hothorn, 2006) are used to analyse these descriptors and corresponding ratings. MDS, CA, and PCA are briefly introduced with specific reference to the analysis of repertory grid data.

H.1.1 Multidimensional scaling applied to the repertory grid data

A MDS analysis of listeners' descriptors can identify regions, or *facets* (Borg and Shye, 1995), in the resulting MDS plot where sounds possess similar values based upon analysed attributes. The participants' responses for each set or triad of stimuli presented can potentially yield a set of bipolar descriptors that have the following properties.

- The descriptors are produced by the listeners in response to the stimuli presented, rather than being selected from a pre-defined list.
- The descriptors were chosen by the participants in order to identify what they perceived were the most salient attributes in how the stimuli differed.
- The descriptors were chosen to provide for each descriptor, a contrary identification where one of the three stimuli was selected as being the 'odd one out'.

MDS is used to visualise the dissimilarity matrix[1] of the participants grid results in a two dimensional plot where the constructs and the elements of participants can be easily visualised. In this experiment the pairwise distinction was taken from each listener's similarity structure which participants created during the rating task stage of the experiment. MDS maintains the data's relative positions, so it can also be used for cluster identification. It can be combined with a minimum spanning tree (MST) to identify any inaccuracies present in the MDS representation or influential observations (Jolliffe, 2002). A MST overlaid with a scatterplot can provide a visual representation of the structure of clusters in the data without any relation to any possible hierarchical structure. This visualises the underlying structure of relations between elements. The analyse of an MDS plot typically looks for common patterns in the plot.

Simplex, circumplex and radex patterns or shapes are commonly found in data with a lower number of dimensional solutions such as the data being analysed in this experiment. When exploring the results of an MDS plot the visual evaluation of these types of patterns

[1] A dissimilarity matrix holds the pairwise distinction between N object, in a NxN matrix format where the xy^{th} element equals the measure of distinction between the x^{th} and the y^{th} object.

can tell us about the number of dimensions of the phenomena under investigation. A simplex pattern is where the relationship between the stimuli represent as data points on a plot can be transformed onto a single line, it often looks like a horseshoe. A simplex pattern would mean that there is a single dimension that these constructs were based upon. In the case of elements or sound stimuli and the participant's MDS plot, a resulting simplex pattern would indicate a lack of variety within the stimulus set depending on the context the sounds are to be used in. A circumplex pattern is where the available data points (stimuli) from the plot look similar to a lumpy circles or ellipses but often this pattern looks closer to a four sided polygon. Finding a circumplex pattern in the case of the elements would mean a two dimensional basis was used by the participants in their classification of the stimuli. The radex pattern is where the data points look like the spokes of a wheel on the plot. The interpretation of these patterns with relation to the results of this experiment is covered in more detail in Section 6.2.2. These types of patterns allow for a rapid examination of the results of any MDS plot and allows an understanding of the dimensionality attributed to the sounds and to the constructs by the participant.

H.1.2 Cluster analysis applied to the repertory grid data

CA is used to group the objects that are represented in the dissimilarity matrix of the participants grids. CA partitions the data and presents it as a tree structure. This visualisation can be helpful in determining the potential groups of objects. CA was performed on the participant's ratings associated with each descriptor as proposed by Berg and Rumsey (1999) and a matrix of distances between each of the scales was calculated. Cluster analysis uses these distances to create a *dendrogram* or branching tree style of diagram, where the descriptors/scales are leaves and where the nodes are clusters. The more similar two scales are, the closer to the bottom their two respective leaves will be connected. MDS and CA offer the ability to represent deep structures from within the data in a visual and understandable manner. The difference between these methods is that MDS results in a non-hierarchical view of the underlying structure of relations between elements in the data set whereas CA presents results in a hierarchical structure of the elements in the data set.

H.1.3 Principal component analysis applied to the repertory grid data

PCA is used to find new variables or trends of interest from the data, called principal components. These principal components are derived from a linear combination of the original grid data variables in a manner that accounts for the majority of the variation found in these variables. PCA is also used for providing a mechanism to simplify the analysis of the correlation matrices[2], derived from the correlations between constructs. The similarity structures which the participants generated in the rating stage were used as the samples that generated the Pearson's r values. These are the correlation coefficient values in the correlation matrix used for the PCA.

In order to provide some simplification for the interpretation of the PCA, the PCA simplification method suggested by Jolliffe (2002, p. 65-67) was used and is indicated in the tables under the simplified columns. The plus + and the minus - signs indicate a coefficient whose absolute value is greater than half the maximum coefficient (absolute value) and the bracketed plus *(+)* and minus *(-)* signs indicate a coefficient whose absolute value is between a quarter and a half of the largest absolute value for the particular principal component. Coefficients whose values are below the cut-off rate of less than a quarter of the largest absolute value for the particular PC are omitted from the tables, leaving blank spaces. This simplified representation is used in conjunction with the coefficients rounded to two decimal places to aid in the interpretation of the principal components.

Participant 1 - PCA - Constructs

The results of the elicitation task for Participant 1 are shown below in Table H-1. It is useful to note which of the stimuli were used in the particular triadic comparison (see Table 6.3 for the full stimuli list with descriptions) which created the poles for this participant's derived scale. The rating task used these poles for its four sub-tasks.

	Constructs **Emergent Pole — Implicit Pole**	**Similar — Different Sounds**
1	flowing sounds–rubbing sounds	*s5, s10 — s8*
2	water sounds–machine sounds	*s11, s12 — s3*
3	mechanical/electrical sounds–paper sounds	*s2, s4 — s7*
4	kitchen sounds–knocking sounds	*s1, s9 — s6*

Table H-1*:* Descriptions of sounds by Participant 1 for the construct poles.

Examining the correlations among the constructs for Participant 1, shown in Table H-2, we

[2] A correlation matrix holds the correlation values between N variables, in a NxN matrix format where the xy^{th} element equals the correlation coefficient r_{xy} of the x^{th} and the y^{th} variable. The correlation coefficient is the degree of linear relationship between the x^{th} and the y^{th} variable where the range is from -1 to +1 representing the strength and the direction of the relationship between the two random variables. A result of -1 means there is a perfect negative relation between the variables, a +1 means a perfect positive relation, and 0 would indicate that there is no relation between the variables. Cohen (1992) presented guidelines for the social sciences with regard to interpreting this value stating that: r=0.1, small effect size; r=0.3, medium effect size; r=0.5, large effect size.

can say that "*water sounds – machine sounds*" (construct 2) are not associated with "*mechanical/electrical sounds – paper sounds*" (construct 3, −0.64). Each construct can be considered in a similar manner to the second construct with regard to the other constructs, rather than consider each of the constructs individually we will present the significant correlations as these are the items we can make the strongest statements about, based on the participant's data. In the case of this participant, there were no other constructs with significant correlations. These types of correlations can help in giving an insight into the participants world view of the sounds presented. In order to reduce the amount of information present in Table H-2, we can use the root mean square correlation among constructs as shown in Table H-3. The "*water sounds – machine sounds*" is the construct most closely associated with the other constructs as shown in Table H-3.

	1	2	3	4
1 flowing sounds–rubbing sounds	1.00	0.42	0.28	0.29
2 water sounds–machine sounds		1.00	−**0.64**	0.18
3 mechanical/electrical sounds–paper sounds			1.00	0.00
4 kitchen sounds–knocking sounds				1.00

Table H-2*:* Correlation Analysis for Participant 1 Constructs.

Construct	Root-mean-square correlation
flowing sounds–rubbing sounds	0.58
water sounds–machine sounds	**0.64**
mechanical/electrical sounds–paper sounds	0.61
kitchen sounds–knocking sounds	0.53
Average of statistic	0.59
Standard deviation of statistic	0.05

Table H-3*:* Root-mean-square (average) correlation among constructs analysis for Participant 1

The interpretation of the first principal component in Table H-4 measures the overall 'naturalness' of the sources of the sounds. The second principal component in Table H-4 contrasts the type of interaction which occurred within the sound, implying that after the 'naturalness' of the sources of the sounds is taken into account, the main source of variation was between sounds with discrete interactions and those with continuous interactions. The first two components is shown in Figure H-2, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-3. These figures can help in illustrating the interpretation of the principal components.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
flowing sounds–rubbing sounds	−0.20	(-)	0.74	+
water sounds–machine sounds	−0.67	-	0.15	
mechanical/electrical sounds–paper sounds	0.67	+	0.12	
kitchen sounds–knocking sounds	−0.25	(-)	−0.65	-
Standard deviation	1.4462		1.0969	
Proportion of Variance	0.5229		0.3008	
Cumulative Proportion	0.5229		0.8237	

Table H-4*:* Principal-Components Analysis for Participant 1 Constructs.

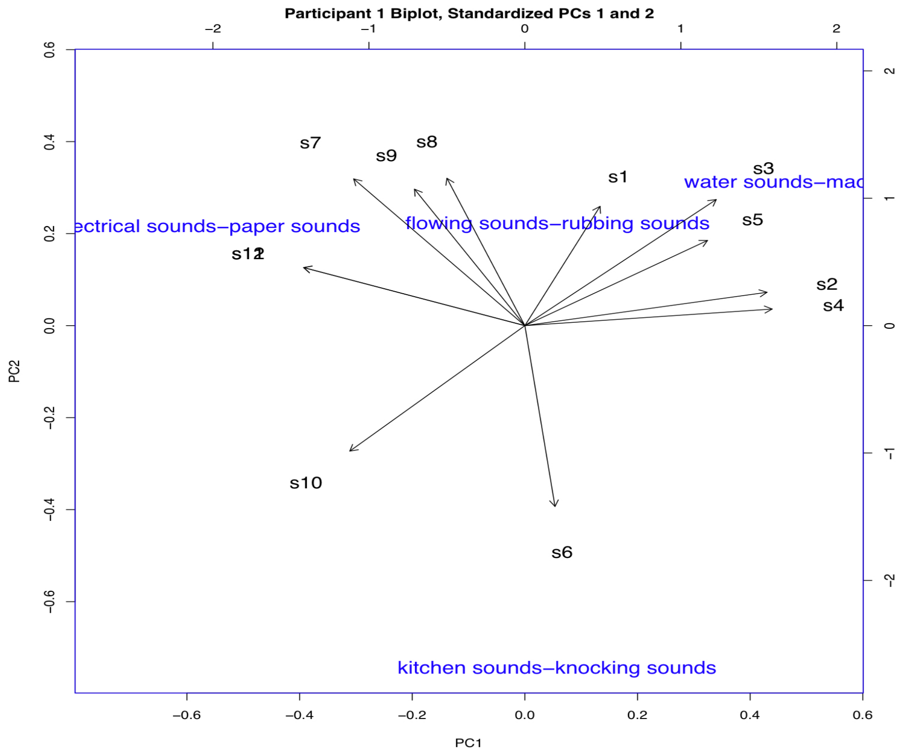


Fig. H-2: *The first two components of the principal–components analysis for Participant 1.*

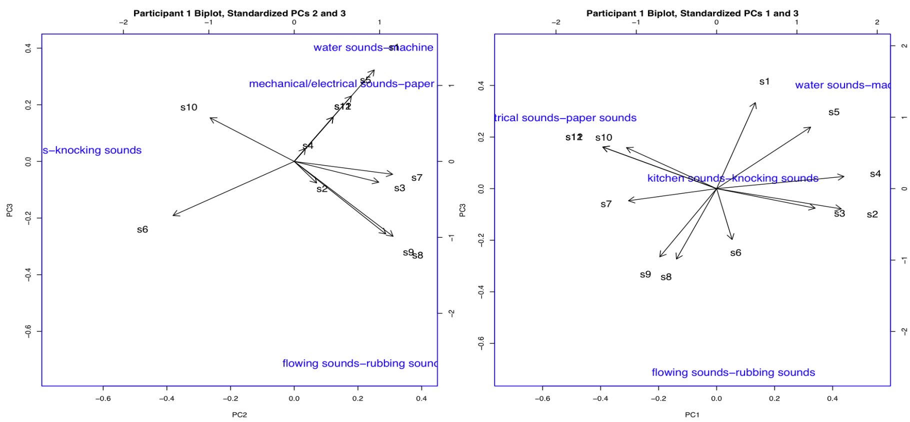


Fig. H-3: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 1.*

Participant 1 - PCA - Elements

The results of the elicitation task for Participant 1 are shown below in Table H-5. Participants responded in free-text format to what they thought each sound was as we can see from Table H-5, these text descriptions were often highly descriptive and described the events or actions. The accuracy of the descriptions and any confusion with their identification can be addressed by the methods previously discussed in Chapter 5. Here we are more interested in the rating, descriptors, and any possible metaphors from the participants than with the identification issues of the sounds.

ID	Sound Description	Participant's Description
s1	gas stove	*gas hob on cooker with flames flickering*
s2	bottling machinery	*machine - possibly life or pressing machine - could be loud vending machine*
s3	cutting machinery	*coffee grinding machine*
s4	electronic alarm clock	*alarm clock ringing*
s5	gas expelling	*sound of hissing gas*
s6	knocking on door	*knocking on door and then opening of door*
s7	turning paper	*reading a magazine and flicking pages*
s8	rubbing and writing	*rubbing surface or drawing on surface*
s9	rubbing sandpaper	*scrubbing of tiles*
s10	stream, water flowing	*river flowing - medium to small - grade 2 or less*
s11	water dripping	*water dripping sound*
s12	water pouring, bath	*water pouring sound*

Table H-5*:* Descriptions by Participant 1 for the elements.

Examining the correlations among the elements for Participant 1, shown in Table H-6, we can say that element *s1* (the sound of gas ring hissing) was definitely associated with *s5 (0.87)* (the sound of gas being expelled from a large canister or cylinder) while being definitely not associated with *s6* (-0.82) (the sound of somebody knocking on a door and then being let in). As previously stated, we will only consider the one element or sound and the elements with the most significant correlations to this element, in order to highlight the points of interest from the participant's data. The *s4* element (the sound of an electric alarm clocking buzzing) is the element most closely associated with the other elements as shown in Table H-7. The elements *s11* (the sound of water drops dripping into a container with water) and *s12* (water flowing from a bath or a tap) are the next most closely associated elements.

The interpretation of the first principal component in Table H-8 contrasts the type of source which produced the sound between those which were more mechanical versus naturalistic. The second principal component in Table H-8 suggests that the next main source

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12
s1	1.00	0.25	0.46	0.43	**0.87**	−**0.82**	0.13	−0.13	−0.19	−0.30	0.17	0.17
s2		1.00	**0.88**	**0.96**	0.68	0.05	−0.53	−0.08	−0.21	−**0.85**	−**0.88**	−**0.88**
s3			1.00	0.79	0.73	−0.38	−0.06	0.31	0.18	−**0.98**	−0.57	−0.57
s4				1.00	**0.82**	0.00	−0.63	−0.32	−0.45	−0.71	−**0.82**	−**0.82**
s5					1.00	−0.52	−0.26	−0.26	−0.37	−0.58	−0.33	−0.33
s6						1.00	−0.67	−0.40	−0.38	0.30	−0.52	−0.52
s7							1.00	**0.80**	**0.85**	0.00	0.77	0.77
s8								1.00	**0.99**	−0.45	0.26	0.26
s9									1.00	−0.32	0.37	0.37
s10										1.00	0.58	0.58
s11											1.00	**1.00**
s12												1.00

Table H-6: Correlation Analysis for Participant 1 Elements.

Construct	Root-mean-square correlation
s1	0.41
s2	0.60
s3	0.58
s4	**0.65**
s5	0.56
s6	0.46
s7	0.54
s8	0.44
s9	0.47
s10	0.55
s11	0.61
s12	0.61
Average of statistic	0.54
Standard deviation of statistic	0.08

Table H-7*:* Root-mean-square (average) correlation among elements analysis for Participant 1

of variation was between the 'friction' and the 'impact' types of sound events. The first two components is in Figure H-4, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-5. These figures can help in illustrating the interpretation of the principal components.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
s1	−0.17	(-)	0.35	+
s2	−0.36	-	0.09	
s3	−0.31	-	0.29	+
s4	−0.37	-	0.05	
s5	−0.32	-	0.21	(+)
s6	−0.07	(-)	−0.51	-
s7	0.30	+	0.32	+
s8	0.18	(+)	0.39	+
s9	0.24	+	0.34	+
s10	0.29	+	−0.30	-
s11	0.35	+	0.10	
s12	0.35	+	0.10	
Standard deviation	2.7094		1.7815	
Proportion of Variance	0.6118		0.2645	
Cumulative Proportion	0.6118		0.8762	

Table H-8*:* Principal-Components Analysis for Participant 1 Elements.

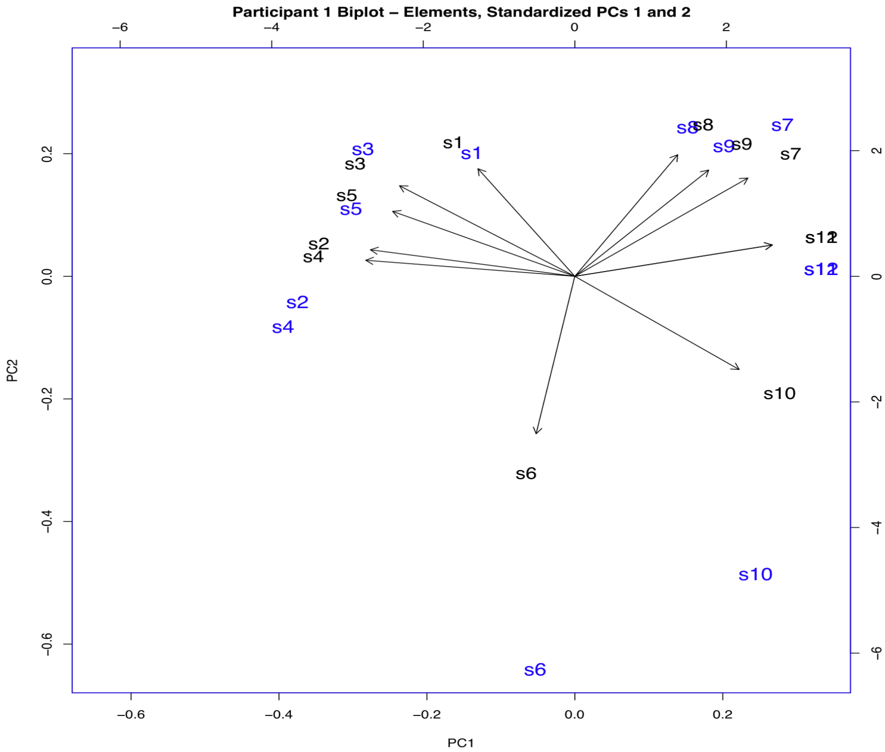


Fig. H-4: *The first two components of the principal–components analysis for Participant 1 Elements.*

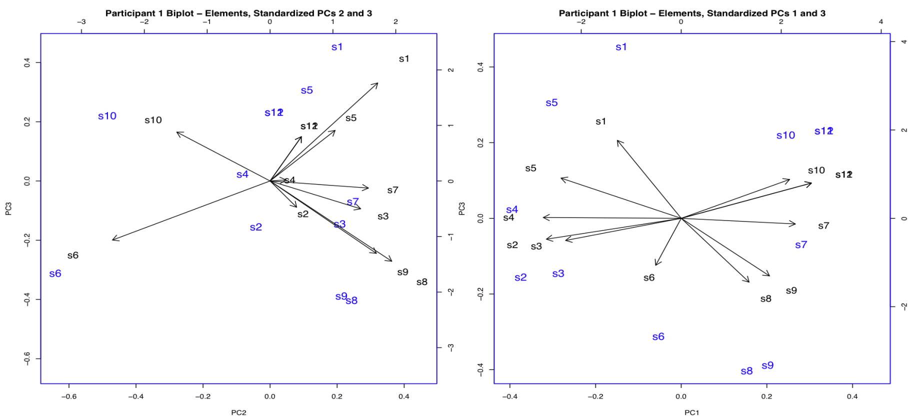


Fig. H-5: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 1 Elements.*

Participant 1 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 1 is shown in Figure H-6 with a single cutoff level taken as 5.9 which results in 3 clusters. These clusters are shown in Table H-9. A minimum spanning tree (MST) is a used to highlight possible distortions produced by the scaling solutions. These distortions are indicated by nearby points on the MDS plot not being linked by an edge of the tree. The MST multidimensional scaling analysis for Participant 1 is shown in Figure H-7, with its related Shepard diagram in Figure H-8. The MST multidimensional scaling analysis has a simplex type pattern meaning that a single dimension can describe the data. A flowing water to machine / impact continuum can be seen in Figure H-7. A Shepard diagram is used to highlight discrepancies between the original dissimilarities (shown in diagram as X's) and the multidimensional scaling solution (the line in the diagram). This type of diagram is used to judge the quality of the multidimensional scaling solution, in an ideal solution all the points would fall on the bisecting line. This type of solution can be seen in the Shepard diagram for Participant 1 in Figure H-8.

Cluster 1 contained sounds that were seen as either liquid flowing sounds or mechanical / rubbing sounds. The shortest distance match in this cluster was at 5.9 between "*flowing sounds – rubbing sounds*" and "*water sounds – machine sounds*". The next sub-cluster joined at a distance of 6.25 and was "*mechanical/electrical sounds – paper sounds*". The last sub-cluster joined at a distance of 6.5 and was "*kitchen sounds – knocking sounds*". These constructs suggest the sounds are clustered based on the type of interactions occurring within the sounds with continuous type sounds either rubbing or flowing being clustered together and then broadening out into discrete impact type sounds such as knocking.

These sounds and the derived constructs were construed in terms of their subject matter and the relevance to location. In some cases, while the constructs may appear to be similar, if they do not distinguish between elements in the same manner and from a similar or the same context then they are unlikely to be referring to the same concept. The ordering of the presentation and choice of stimuli within the triads could be an issue, further explorations using the same stimuli set but where different triads of sounds are presented to the other participants.This provides a better view of the constructs and sounds.

Each cluster can be considered in a similar manner to the first cluster and rather than consider each of the clusters individually, we will present the cluster with the shortest significant distances from the set of clusters. This is the cluster that we can make the strongest statements about, based on the participant's data.

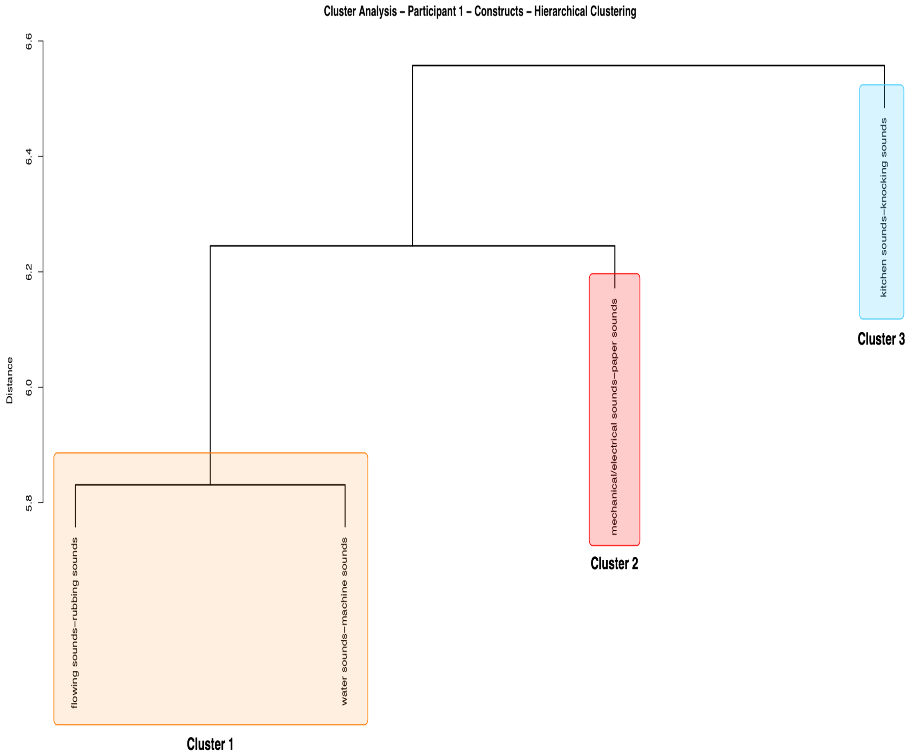


Fig. H-6: *The cluster analysis of the RGT constructs for Participant 1.*

	Constructs *Emergent Pole* - *Implicit Pole*
1	flowing sounds — rubbing sounds water sounds — machine sounds
2	mechanical/electrical sounds — paper sounds
3	kitchen sounds — knocking sounds

Table H-9: Clusters obtained for Participant 1.

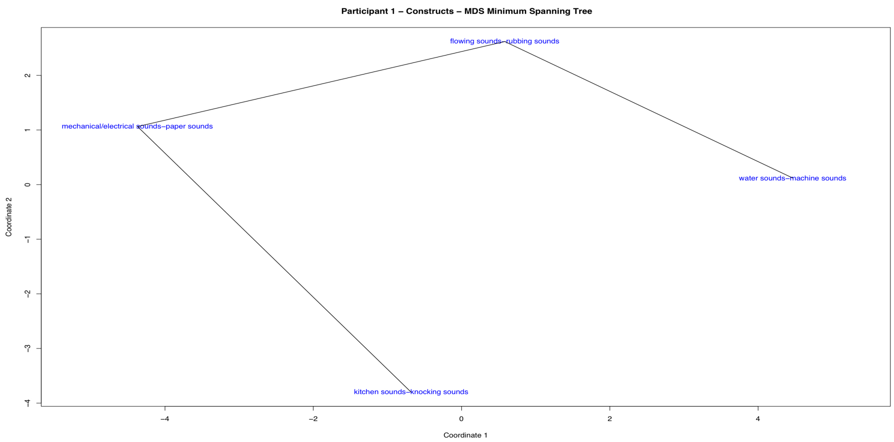


Fig. H-7: *MDS minimum spanning tree analysis of the RGT constructs for Participant 1.*

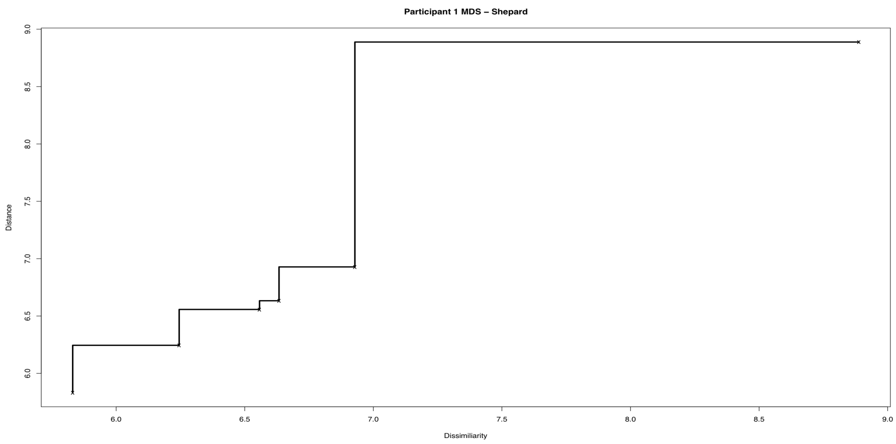


Fig. H-8: *Shepard diagram of the MDS analysis of the RGT constructs for Participant 1.*

Participant 1 - Cluster Analysis and MDS - Elements

The cluster analysis of the elements from Participant 1 is shown in Figure H-9 which resulted in 4 clusters as shown in Table H-10. The distances of the clusters were Cluster 1 (Gaseous & Electro-mechanical) had a distance of 0.87, Cluster 2 (Impact) had a distance of 1.33, Cluster 3 (Friction) had a distance of 0.6, and Cluster 4 (Liquids) was found at a distance of 0.71. Cluster 1 contained sounds that were seen as gaseous or electro-mechanical sounds. The shortest distance match in this cluster between s2 (the sound of a bottling plant machine in operation) and s4 (an electronic alarm clock bleeping). Cluster 2 contained a single sound, s6 (the sound of somebody knocking on a door) which was seen as an impact type sound. Cluster 3 contained three sounds, s7 (the sound of paper sheets or a book being turned), s8 (the sound of rubbing and writing), and s9 (rubbing using sandpaper). These were seen as friction type sounds. Cluster 4 contained sounds which were seen as liquid type sounds.

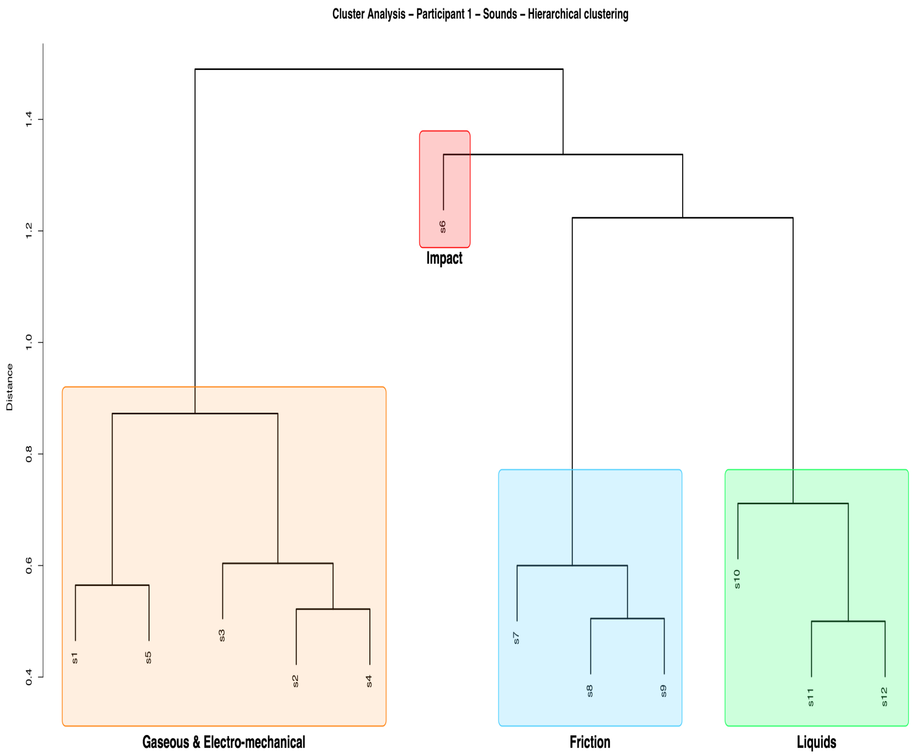


Fig. H-9: *The cluster analysis of the RGT elements (sounds) for Participant 1.*

	Cluster "label" *Stimuli Number*
1	Gaseous & Electro-mechanical *s1, s5, s3, s2, s4*
2	Impact *s6*
3	Friction *s7, s8, s9*
4	Liquids *s10, s11, s12*

Table H-10: Clusters obtained for the elements of Participant 1.

The MST multidimensional scaling analysis for the elements from Participant 1 is shown in Figure H-10, with its related Shepard diagram in Figure H-11. The MST multidimensional scaling analysis has a simplex type pattern meaning that a single dimension can describe the data. A gaseous to impact continuum can be seen in Figure H-10. The Shepard diagram in Figure H-11, highlights some discrepancies and indicates that the current scaling solution is not the most ideal solution as all of the points did not fall on the bisecting line in the figure. However, for the case of this research, it is adequate for the exploratory purposes for this study and attempting to fit or massage the data to fall on a particular line will not provide any additional clarity for the purposes of this research.

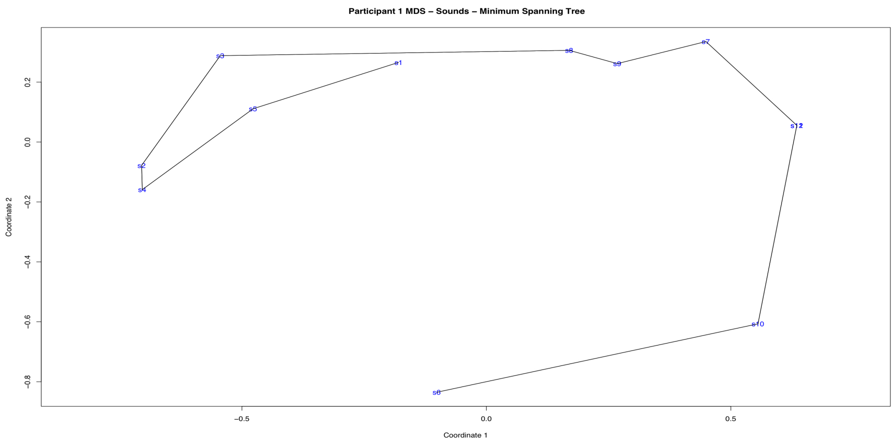


Fig. H-10: *MDS minimum spanning tree analysis of the RGT elements for Participant 1.*

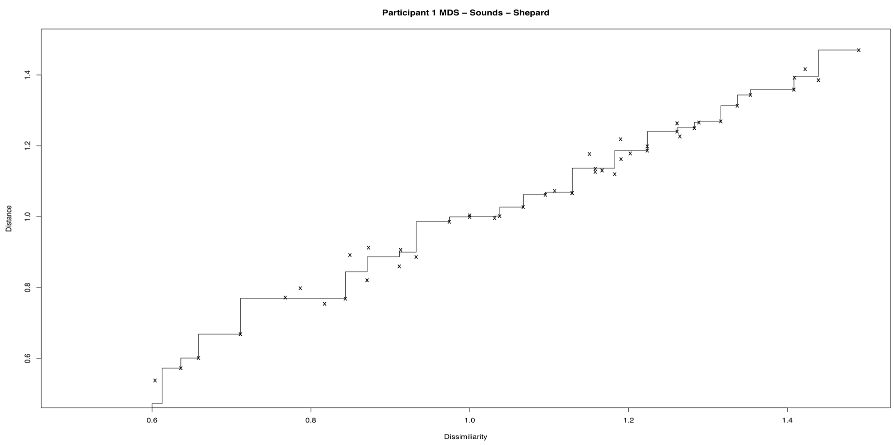


Fig. H-11: *Shepard diagram of the MDS analysis of the RGT elements for Participant 1.*

H.1.4 Results and Observations for Participant 1

The results of the analysis of the constructs from Participant 1 from the PCA, CA, and MDS analysis suggest that the ‘naturalness’ of the sound sources followed by the type of interaction, whether discrete or continuous are factors which used by this participant. The results of the analysis of the elements (sounds) from Participant 1 from the PCA, CA, and MDS analysis suggest that a mechanical–natural scale followed by the type of action on a friction–impact scale were used by this participant as factors. These potential scales are further strengthened by mapping the free text descriptors onto the classification from the CLOSED project as shown in Chapter 6.

Participant 2 - PCA - Constructs

The results of the elicitation task for Participant 2 are shown below in Table H-11. It is useful to note which of the stimuli were used in the particular triadic comparison (see Table 6.3 for the full stimuli list with descriptions) which created the poles for this participant's derived scale. The rating task used these poles for its four sub-tasks.

	Constructs **Emergent Pole — Implicit Pole**	**Similar — Different Sounds**
1	progressive/flowing–repetitive	*s1, s3 — s4*
2	pointed/focused–continuous/unbroken	*s7, s8 — s12*
3	tapping/dripping/nagging–hissy/steamy/flowing	*s6, s11 — s5*
4	in motion/bubbly/bobbing–touching/tangential	*s2, s10 — s9*

Table H-11: Descriptions of sounds by Participant 2 for the construct poles.

Examining the correlations among the constructs for Participant 2, shown in Table H-12, we can say that "*progressive/flowing – repetitive*" (construct 1) were definitely not associated with "*pointed/focused – continuous/unbroken*" and "*tapping/dripping/nagging – hissy/steamy/flowing*", which were constructs 2 (−0.84) and 3 (−0.86) respectively. Each construct can be considered in a similar manner to the second construct with regard to the other constructs, rather than consider each of the constructs individually, we will present the significant correlations as these are the items we can make the strongest statements about, based on the participant's data. In the case of this participant there were no other constructs with significant correlations. These types of correlations can help in giving an insight into the participants world view of the sounds presented. In order to reduce the amount of information present in Table H-12, we can use the root mean square correlation among constructs as shown in Table H-13. The "*pointed/focused – continuous/unbroken*" was the construct which was most closely associated with the other constructs as shown in Table H-13.

	1	2	3	4
1 progressive/flowing–repetitive	1.00	−**0.84**	−**0.86**	0.54
2 pointed/focused–continuous/unbroken		1.00	**0.90**	−0.60
3 tapping/dripping/nagging–hissy/steamy/flowing			1.00	−0.42
4 in motion/bubbly/bobbing–touching/tangential				1.00

Table H-12: Correlation Analysis for Participant 2 Constructs.

Construct	Root-mean-square correlation
progressive/flowing–repetitive	0.83
pointed/focused–continuous/unbroken	**0.85**
tapping/dripping/nagging–hissy/steamy/flowing	0.82
in motion/bubbly/bobbing–touching/tangential	0.67
Average of statistic	0.79
Standard deviation of statistic	0.08

Table H-13*:* Root-mean-square (average) correlation among constructs analysis for Participant 2

The interpretation of the first principal component in Table H-14 contrasts the temporal nature of the sound whether it is a flowing continuous type of sound or a repetitive sequence type sound. The second principal component in Table H-14 contrasts the type of interaction which occurred within the sounds, implying that after the discrete or continuous nature of the sounds is taken into account, the main source of variation between the sounds was between those sounds with a friction type interaction and those with an impact type interaction occurring. The first two components are shown in Figure H-12, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-13. These figures can help in illustrating the interpretation of the principal components.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
progressive/flowing–repetitive	−0.50	-	−0.37	(-)
pointed/focused–continuous/unbroken	0.51	+	0.08	
tapping/dripping/nagging–hissy/steamy/flowing	0.50	+	0.37	(+)
in motion/bubbly/bobbing–touching/tangential	−0.48	-	0.85	+
Standard deviation	1.9640		0.3746	
Proportion of Variance	0.9644		0.0351	
Cumulative Proportion	0.9644		0.9994	

Table H-14*:* Principal-Components Analysis for Participant 2 Constructs.

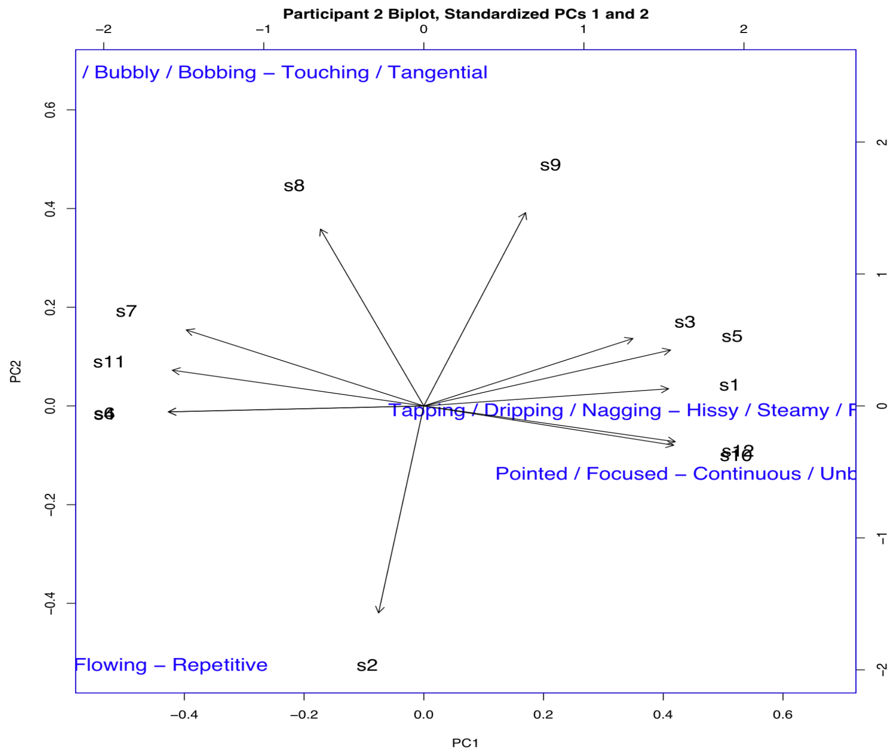


Fig. H-12*: The first two components of the principal–components analysis for Participant 2.*

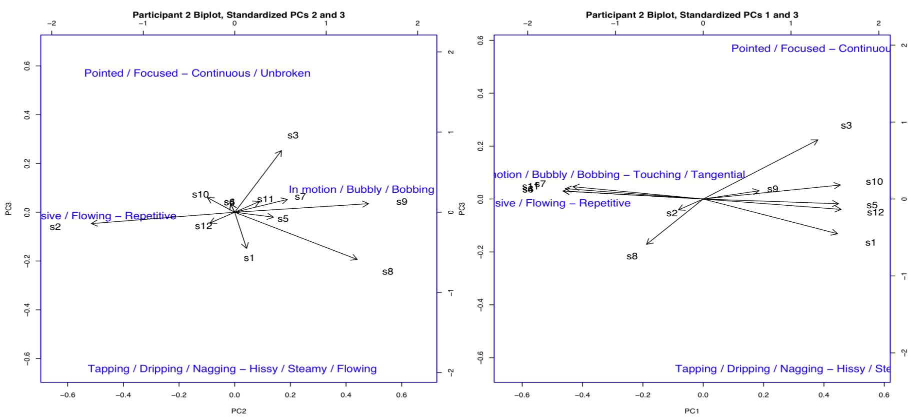


Fig. H-13*: The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 2.*

Participant 2 - PCA - Elements

The results of the elicitation task for Participant 2 are shown below in Table H-15. Participants responded in free-text format to what they thought each sound was as we can see from Table H-15, these text descriptions were often highly descriptive and described the events or actions. The accuracy of the descriptions and any confusion with their identification can be addressed by the methods previously discussed in Chapter 5, but in this section we are more interested in the rating, descriptors, and any possible metaphors from the participants than with the identification issues of the sounds.

ID	Sound Description	Participant's Description
s1	gas stove	*beginning / progression / flame*
s2	bottling machinery	*jarring action*
s3	cutting machinery	*howling / etching*
s4	electronic alarm clock	*repeating / needling*
s5	gas expelling	*steamy / continuous*
s6	knocking on door	*rapping / tapping / completing*
s7	turning paper	*sharp / punctuated*
s8	rubbing and writing	*brushing / light darting*
s9	rubbing sandpaper	*shuffling / touching*
s10	stream, water flowing	*bubbling / in motion*
s11	water dripping	*dripping / plunking*
s12	water pouring, bath	*full / flowing / continuous*

Table H-15*:* Descriptions by Participant 2 for the elements.

Examining the correlations among the elements for Participant 2, shown in Table H-16, we can say that element *s4* (the sound of electronic alarm clock buzzing) was definitely associated with *s6 (1.00)* (a knocking on door sound), *s7 (0.92)* (the sound of paper being turned), and *s11 (0.98)* (the sound of water dripping) while being definitely not associated with *s5 (−0.97)* (the sound of gas being expelled from a large canister or cylinder), *s10 (−0.96)* (the sound of a stream with water flowing), and *s12 (−0.98)* (the sound of water pouring). As previously stated, we will only consider the one element or sound and the elements with the most significant correlations to this element, in order to highlight the points of interest from the participant's data.

The elements *s4* (the sound of an electric alarm clocking buzzing) and *s6* (a knocking on door sound) are the joint elements for being most closely associated with the other elements as shown in Table H-17. The element *s5* (the sound of gas being expelled from a large canister or cylinder) was the next most closely associated elements.

The interpretation of the first principal component in Table H-18 contrasts the 'friction' and the 'impact' types of sound events. The second principal component in Table H-18 sug-

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12
s1	1.00	−0.22	0.67	**−0.97**	**0.95**	**−0.97**	**−0.89**	−0.21	0.44	**0.89**	**−0.95**	**0.95**
s2		1.00	−0.50	0.20	−0.43	0.20	−0.20	−0.72	**−0.97**	0.00	0.00	0.00
s3			1.00	−0.79	**0.85**	−0.79	−0.59	−0.24	0.65	0.79	−0.71	0.71
s4				1.00	**−0.97**	**1.00**	**0.92**	0.36	−0.42	**−0.96**	**0.98**	**−0.98**
s5					1.00	**−0.97**	**−0.80**	−0.15	0.62	**0.89**	**−0.90**	**0.90**
s6						1.00	**0.92**	0.36	−0.42	**−0.96**	**0.98**	**−0.98**
s7							1.00	0.64	−0.03	**−0.96**	**0.98**	**−0.98**
s8								1.00	0.58	−0.59	0.51	−0.51
s9									1.00	0.22	−0.23	0.23
s10										1.00	**−0.98**	**0.98**
s11											1.00	**−1.00**
s12												1.00

Table H-16*:* Correlation Analysis for Participant 2 Elements.

Construct	Root-mean-square correlation
s1	**0.76**
s2	0.37
s3	**0.69**
s4	**0.80**
s5	**0.79**
s6	**0.80**
s7	**0.74**
s8	0.49
s9	0.48
s10	**0.77**
s11	**0.77**
s12	**0.77**
Average of statistic	0.68
Standard deviation of statistic	0.15

Table H-17*:* Root-mean-square (average) correlation among elements analysis for Participant 2

gests that the next main source of variation was due to the type of source which produced the sound between those which were more mechanical versus naturalistic. The first two components are shown in Figure H-14, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-15.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
s1	-0.32	-	-0.01	
s2	0.10	(+)	0.63	+
s3	-0.31	-	-0.12	
s4	0.32	+	-0.01	
s5	-0.32	-	-0.08	
s6	0.32	+	-0.01	
s7	0.31	+	-0.16	(-)
s8	0.20	+	-0.51	-
s9	-0.20	-	-0.52	-
s10	-0.32	-	0.09	
s11	0.32	+	-0.08	
s12	-0.32	+	0.08	
Standard deviation	3.1154		1.4988	
Proportion of Variance	0.8088		0.1872	
Cumulative Proportion	0.8088		0.9960	

Table H-18*:* Principal-Components Analysis for Participant 2 Elements.

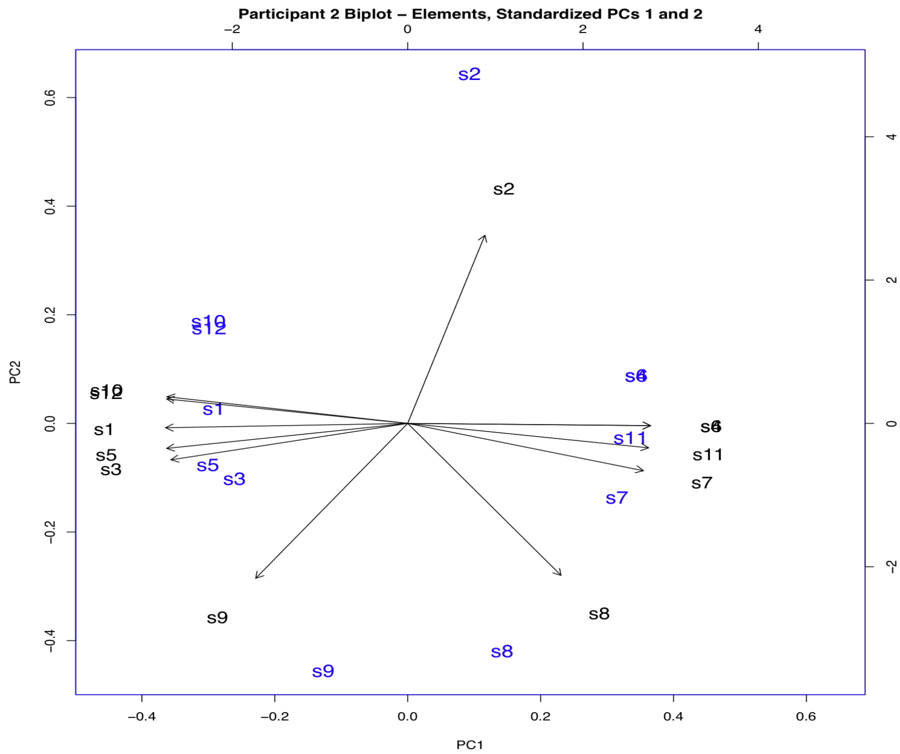


Fig. H-14: *The first two components of the principal–components analysis for Participant 2 Elements.*

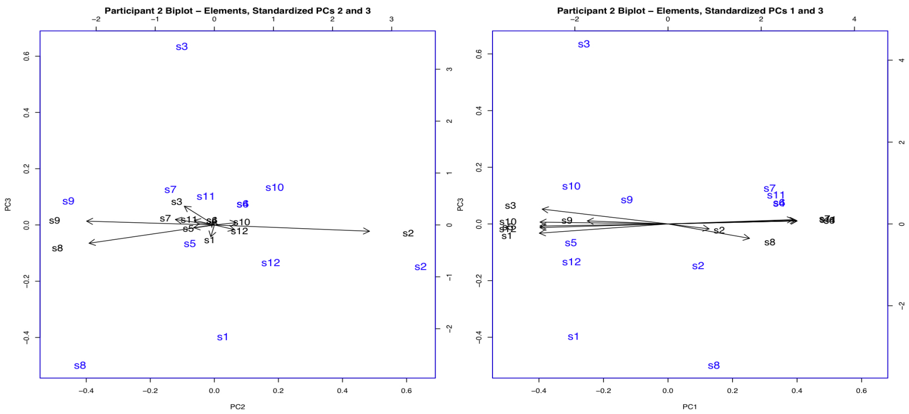


Fig. H-15: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 2 Elements.*

Participant 2 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 2 is shown in Figure H-16 which resulted in 2 clusters as shown in Table H-19. Cluster 1 was found at a distance of 5.57 and Cluster 2 was found at a distance of 2.65. Cluster 1 contained sounds that were seen as either flowing sounds or repetitive sounds. The shortest distance match in this cluster was at 5.57 between "*pro-*

gressive/flowing – repetitive" and "*in motion/bubbly/bobbing – touching/tangential*". Cluster 2 contained sounds that were seen as either brief impact sounds or gaseous yet continuous sounds. The shortest distance match in this cluster was at 2.65 between "*progressive/flowing – repetitive*" and "*in motion/bubbly/bobbing – touching/tangential*".

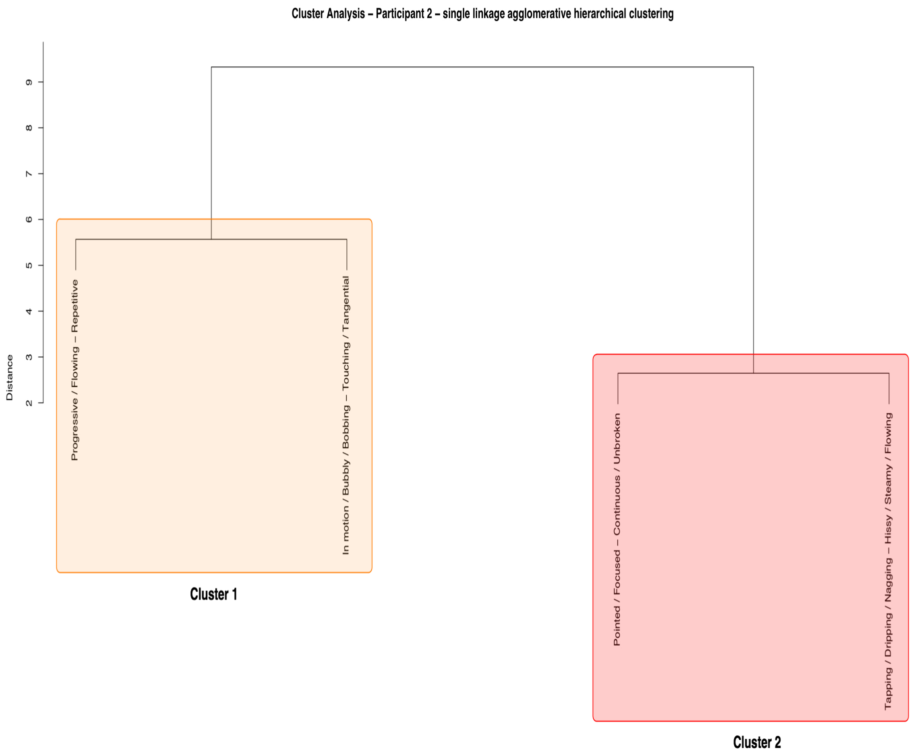


Fig. H-16*: The cluster analysis of the RGT constructs for Participant 2.*

	Constructs *Emergent Pole* - *Implicit Pole*
1	progressive/flowing–repetitive in motion/bubbly/bobbing–touching/tangential
2	pointed/focused–continuous/unbroken tapping/dripping/nagging–hissy/steamy/flowing

Table H-19*:* Clusters obtained for Participant 2.

These constructs suggest the sounds are clustered based on the type of interactions occurring within the sounds with continuous type sounds either gaseous or flowing or sounds of a discrete impact type such as knocking. Both of the clusters for this participant are quite similar and further support the idea that there is a continuum based on flowing or impact sound categorisation which was used to organise the sounds. The differences between the clusters are important as the shorter the distance between clusters, the more significant and stronger the statements that can be made, based on the participant's data.

The MST multidimensional scaling analysis for Participant 2 is shown in Figure H-17, with its related Shepard diagram in Figure H-18. The MST multidimensional scaling analysis shows a simplex type pattern meaning that only a single dimension is needed to account for the constructs. It is likely for Participant 2 that this can be seen as continuous to discrete event continuum. The Shepard diagram is an ideal solution as all the points fall on the bisecting line. In the case of Participant 2, the current results are interesting as they show the potential for an individualised continuum but to gain further detail, more detailed explorations are required.

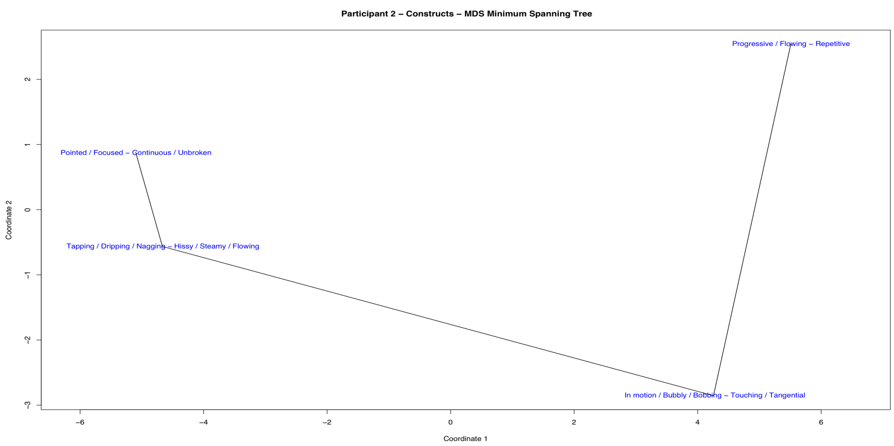


Fig. H-17: *MDS minimum spanning tree analysis of the RGT constructs for Participant 2.*

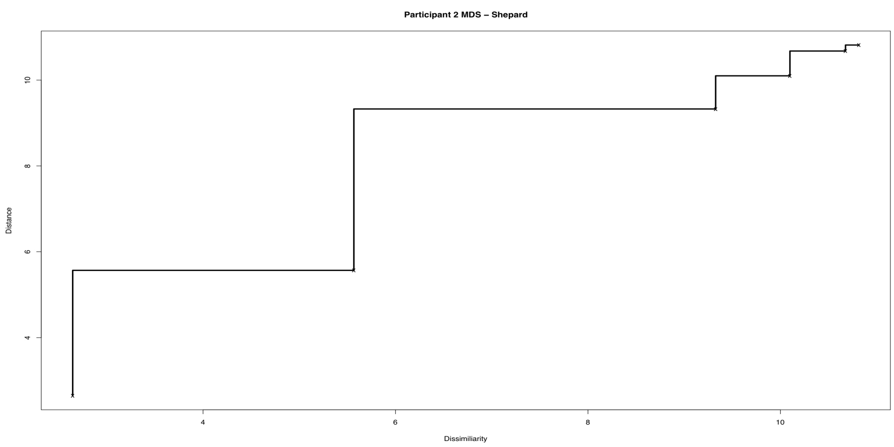


Fig. H-18: *Shepard diagram of the MDS analysis of the RGT constructs for Participant 2.*

Participant 2 - Cluster Analysis and MDS - Elements

The cluster analysis of the elements from Participant 2 is shown in Figure H-19 which resulted in 4 clusters. The resulting clusters are shown in Table H-20. The distances of the clusters were Cluster 1 had a distance of 1.1, Cluster 2 had a distance of 0.66, Cluster 3 had a distance of 0.52, and Cluster 4 was found at a distance of 0.71. The resulting clusters are shown in Table H-20. Cluster 1 contained sounds that were seen as either flowing sounds or repetitive sounds. The shortest distance match in this cluster between s4 (an electronic alarm clock bleeping) and s6 (The sound of knocking on a door). Cluster 2 contained a single sound, s3 (the sound of cutting machinery) which was seen as an impact type sound. Cluster 3 contained sounds where were seen as either gaseous or liquid and that were flowing or expelling type interactions. Cluster 4 contained two sounds, s8 (the sound of rubbing and writing) and s9 (rubbing using sandpaper). These were seen as friction sounds, in particular, using a rubbing type of interaction.

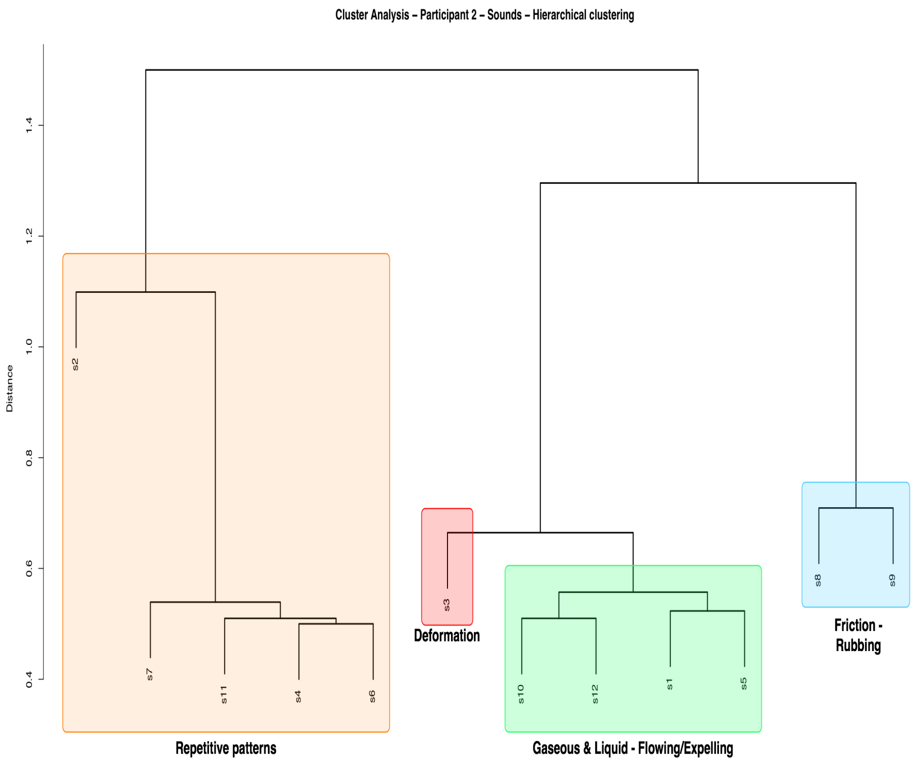


Fig. H-19: *The cluster analysis of the RGT elements (sounds) for Participant 2.*

The MST multidimensional scaling analysis for Participant 2's elements is shown in Figure H-20, with its related Shepard diagram in Figure H-21. The horseshoe like shape pattern in Figure H-20 indicates that is a friction – impact continuum, which Participant 2 used for their classification of the elements or sounds presented.

	Cluster "label" *Stimuli Number*
1	Repetitive *s2, s7, s11, s4, s6*
2	Deformation *s3*
3	Gaseous Liquid – Flowing/Expelling *s10, s12, s1, s5*
4	Friction – Rubbing *s8, s9*

Table H-20*:* Clusters obtained for the elements of Participant 2.

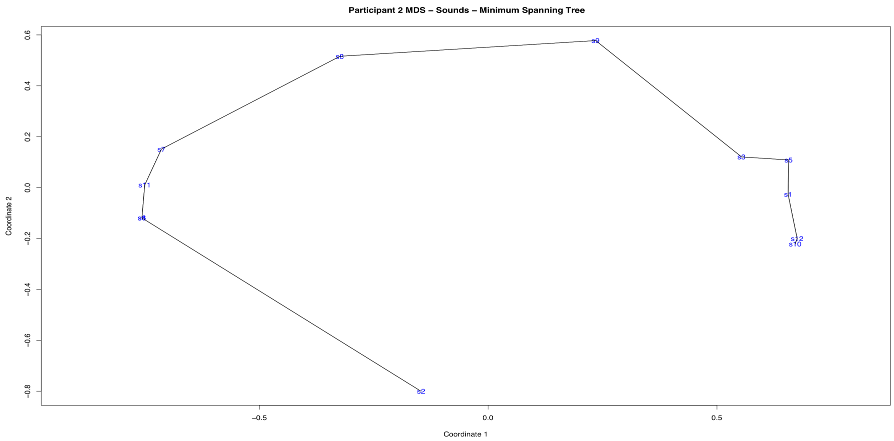


Fig. H-20*: MDS minimum spanning tree analysis of the RGT elements for Participant 2.*

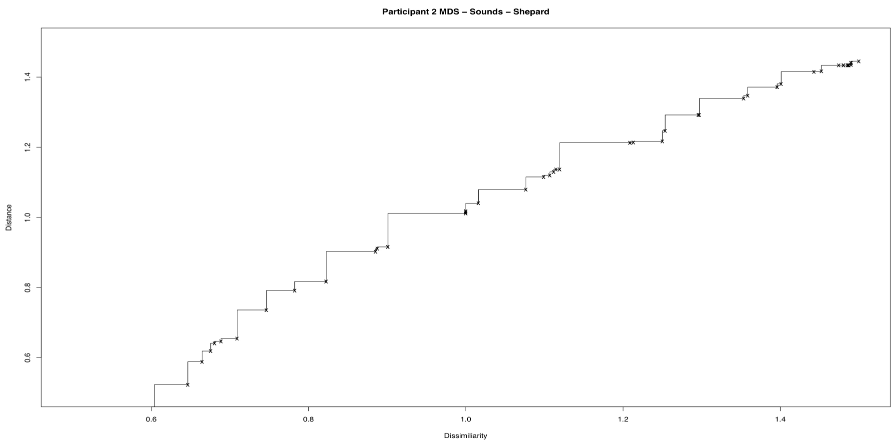


Fig. H-21*: Shepard diagram of the MDS analysis of the RGT elements for Participant 2.*

H.1.5 Results and Observations for Participant 2

The results of the analysis of the constructs from Participant 2 from the PCA, CA, and MDS analysis seem to support that after the discrete or continuous nature of the sounds is taken into account, the main source of variation between the sounds was between those sounds with a friction type interaction and those with an impact type interaction occurring. The results of the analysis of the elements (sounds) from Participant 2 from the PCA, CA, and MDS analysis seem to support that a friction–impact scale followed by the type of action on a mechanical–natural scale were used by this participant as factors, in a similar but reverse fashion to Participant 1.

Participant 3 - PCA - Constructs

The results of the elicitation task for Participant 3 are shown below in Table H-21. It is useful to note which of the stimuli were used in the particular triadic comparison (see Table 6.3 for the full stimuli list with descriptions) which created the poles for this participant's derived scale. The rating task used these poles for its four sub-tasks.

	Constructs **Emergent Pole – Implicit Pole**	**Similar — Different Sounds**
1	artificial watering–timekeeping	*s5, s9 — s4*
2	water–nordic food	*s10, s12 — s1*
3	notetaking–visiting someones office	*s7, s8 — s6*
4	workshop–sounds from an empty home	*s2, s3 — s11*

Table H-21*:* Descriptions of sounds by Participant 3 for the construct poles.

Examining the correlations among the constructs for Participant 3, shown in Table H-22, it is difficult to make strong assumptions due to the relatively weak correlations between constructs. In the case of Participant 3, we can say that "*notetaking–visiting someones office*" (construct 3) is associated with "*workshop–sounds from an empty home*" (construct 4) (0.64) and that this is the strongest association within the constructs for this participant. Each construct can be considered in a similar manner to the third construct with regard to the other constructs. In the case of this participant there were no other constructs with significant correlations. These types of correlations can help in giving an insight into the participants world view of the sounds presented. In order to reduce the amount of information present in Table H-22, we can use the root mean square correlation among constructs as shown in Table H-23. The "*artificial watering – timekeeping*" was the construct which was most closely associated with the other constructs as shown in Table H-23.

The interpretation of the first principal component in Table H-24 contrasts the nature of the source of the sounds whether it is a natural or a manmade or manmade environmental type sound. The second principal component in Table H-24 contrasts the type of interaction

	1	2	3	4
1 artificial watering–timekeeping	1.00	0.21	0.16	0.26
2 water–nordic food		1.00	−0.25	−0.54
3 notetaking–visiting someones office			1.00	**0.64**
4 workshop–sounds from an empty home				1.00

Table H-22*:* Correlation Analysis for Participant 3 Constructs.

Construct	Root-mean-square correlation
artificial watering–timekeeping	0.53
water–nordic food	0.59
notetaking–visiting someones office	0.61
workshop–sounds from an empty home	**0.67**
Average of statistic	0.60
Standard deviation of statistic	0.05

Table H-23*:* Root-mean-square (average) correlation among constructs analysis for Participant 3

which occurred within the sounds, implying that after the natural or a manmade aspect of the sounds is taken as the next source of variation between the sounds after the type of the source of the sound is taken into account. The first two components are in Figure H-22, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-23.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
artificial watering–timekeeping	−0.13		0.97	+
water–nordic food	−0.58	-	−0.15	
notetaking–visiting someones office	0.56	+	−0.11	
workshop–sounds from an empty home	0.57	+	0.18	
Standard deviation	1.6914		1.0066	
Proportion of Variance	0.7152		0.2533	
Cumulative Proportion	0.7152		0.9685	

Table H-24*:* Principal-Components Analysis for Participant 3 Constructs.

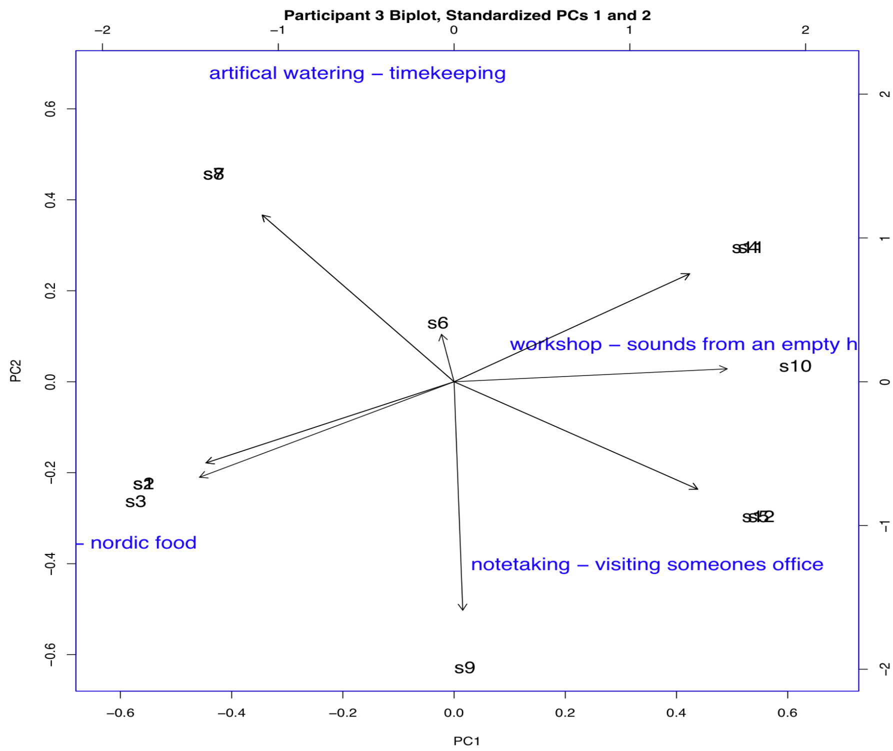


Fig. H-22: *The first two components of the principal–components analysis for Participant 3.*

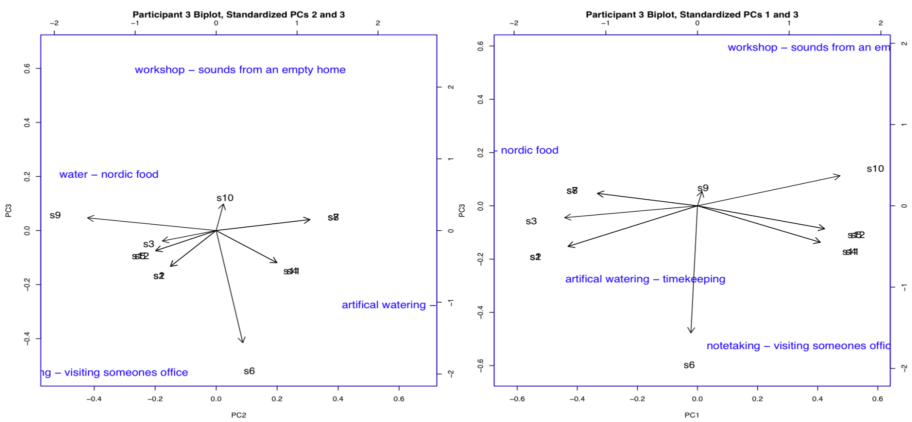


Fig. H-23: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 3.*

Participant 3 - PCA - Elements

The results of the elicitation task for Participant 3 are shown below in Table H-25. The free-text responses from participants is shown in Table H-25.

ID	Sound Description	Participant's Description
s1	gas stove	*someone taking a bite of cracker bread*
s2	bottling machinery	*could be either electric machine or mechanical machine*
s3	cutting machinery	*woodcutting machine (bandsaw)*
s4	electronic alarm clock	*alarm clock*
s5	gas expelling	*water in shower or a sink*
s6	knocking on door	*someone knocking on door and eventually being let in*
s7	turning paper	*someone leafing through his/her notes in a notebook*
s8	rubbing and writing	*someone using an eraser and brushing off the rubber bits*
s9	rubbing sandpaper	*sprinkler watering lawn*
s10	stream, water flowing	*rain*
s11	water dripping	*clock ticking*
s12	water pouring, bath	*filling up the bath*

Table H-25*:* Descriptions by Participant 3 for the elements.

Examining the correlations among the elements for Participant 3, shown in Table H-26, we can say that element *s1* (the sound of gas ring hissing) was definitely associated with *s2 (1.00)* (the sounds of a bottling plant and its machinery), *s3 (0.97)* (the cutting machinery in operation), while being definitely not associated with *s10 (−0.95)* (the sound of a stream with water flowing) and *s11 (−0.82)* (the sound of water dripping). The element *s10* (the sound of a stream with water flowing) was the most closely associated element with the other elements as shown in Table H-27. The elements *s3* (the cutting machinery in operation), *s5* (the sound of gas being expelled from a large canister or cylinder), and *s12* (the sound of water being poured into a bath) were jointly the next most closely associated elements.

The interpretation of the first principal component in Table H-28 contrasts the sounds which are flowing or have a repetitive pattern with non-reptitive or non-continuous sounds. The second principal component in Table H-28 suggests that the next main source of variation was due to type of interaction between objects in the sound with the contrast being between scraping / impact type sounds and friction type sounds. The first two components were already

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12
s1	1.00	**1.00**	**0.97**	**−0.82**	−0.54	0.27	0.32	0.32	0.29	**−0.95**	**−0.82**	−0.54
s2		1.00	**0.97**	**−0.82**	−0.54	0.27	0.32	0.32	0.29	**−0.95**	**−0.82**	−0.54
s3			1.00	**−0.93**	−0.57	0.04	0.31	0.31	0.37	**−0.92**	**−0.93**	−0.57
s4				1.00	0.56	0.33	−0.26	−0.26	−0.47	0.77	**1.00**	0.56
s5					1.00	0.04	**−0.95**	**−0.95**	0.47	0.77	0.56	**1.00**
s6						1.00	0.09	0.09	−0.31	−0.26	0.33	0.04
s7							1.00	**1.00**	−0.73	−0.60	−0.26	**−0.95**
s8								1.00	−0.73	−0.60	−0.26	**−0.95**
s9									1.00	0.00	−0.47	0.47
s10										1.00	0.77	0.77
s11											1.00	0.56
s12												1.00

Table H-26: Correlation Analysis for Participant 3 Elements.

Construct	Root-mean-square correlation
s1	0.65
s2	0.65
s3	0.66
s4	0.65
s5	0.66
s6	0.26
s7	0.57
s8	0.57
s9	0.47
s10	**0.70**
s11	0.65
s12	0.66
Average of statistic	0.59
Standard deviation of statistic	0.12

Table H-27*:* Root-mean-square (average) correlation among elements analysis for Participant 3

shown in Figure H-24, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-25. These figures can help in illustrating the interpretation of the principal components.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
s1	−0.33	-	−0.16	(-)
s2	−0.33	-	−0.16	(-)
s3	−0.33	-	−0.20	(-)
s4	0.31	+	0.25	(+)
s5	0.32	+	−0.18	(-)
s6	−0.05		0.33	+
s7	−0.28	-	0.34	+
s8	−0.28	-	0.34	+
s9	0.03		−0.61	-
s10	0.34	+	0.03	
s11	0.31	+	0.25	(+)
s12	0.32	+	−0.18	(-)
Standard deviation	2.9192		1.6106	
Proportion of Variance	0.7101		0.2162	
Cumulative Proportion	0.7101		0.9263	

Table H-28*:* Principal-Components Analysis for Participant 3 Elements.

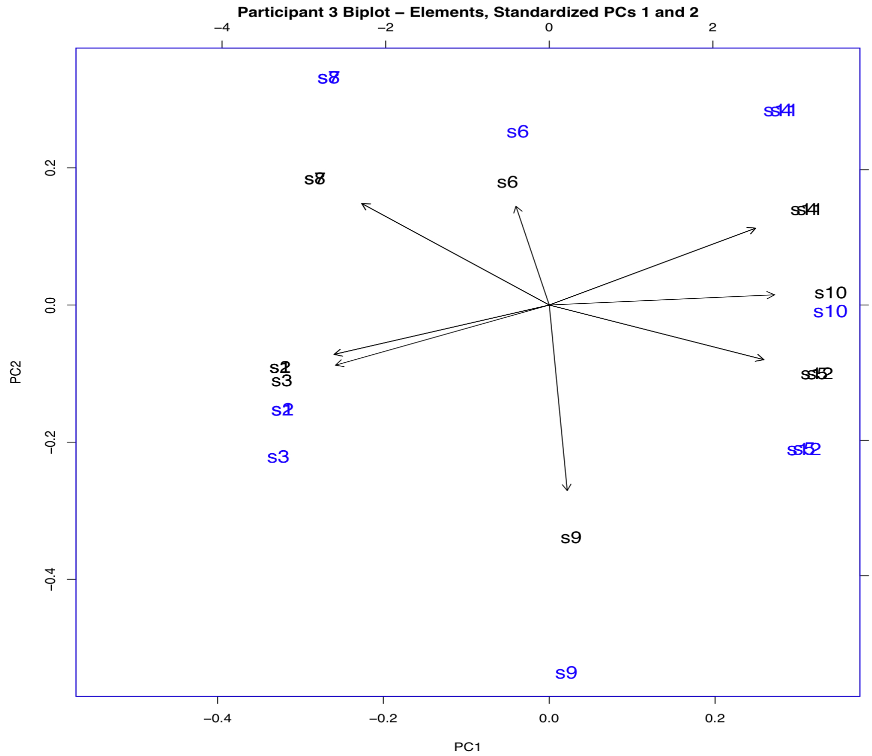


Fig. H-24*: The first two components of the principal–components analysis for Participant 3 Elements.*

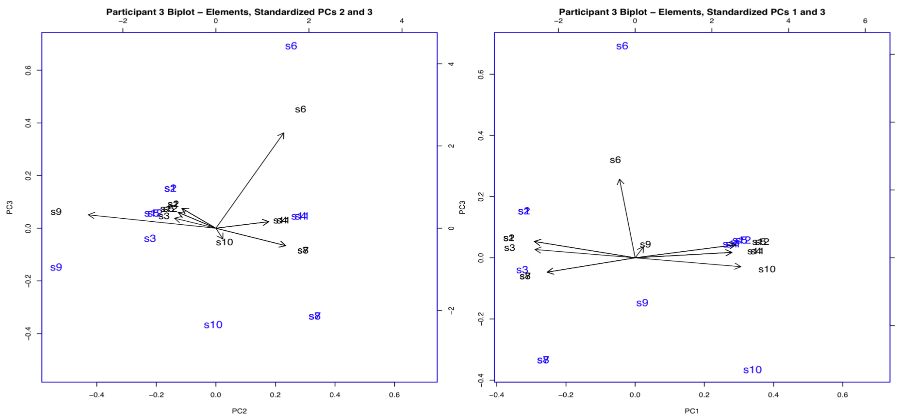


Fig. H-25*: The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 3 Elements.*

Participant 3 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 3 is shown in Figure H-26 which resulted in 2 clusters and shown in Table H-29, Cluster 1 was found at a distance of 6.16 and Cluster 2 was found at a distance of 4.58. Cluster 1 contained sounds that were seen as either water sounds or human activity type sounds. The shortest distance match in this cluster was at 6.16 between "*artificial* *watering–timekeeping*” and “*water–nordic food*”. Cluster 2 contained sounds that were seen as occurring in particular spaces such as offices or homes but with additional information related to the context being also associated with the sounds. The shortest distance match in this cluster was at 4.58 between “*notetaking–visiting someones office*” and “*workshop–sounds from an empty home*”. These constructs suggest the sounds are clustered somewhat based on the location or type of activities occurring and how the sounds fit into the particular context. The results for this participant have shown that both of the clusters for this participant are quite similar and would indicate that there is a continuum based on activity and locational context which was used to organise the sounds but to gain further detail and to verify this hypothesis, more detailed explorations are required.

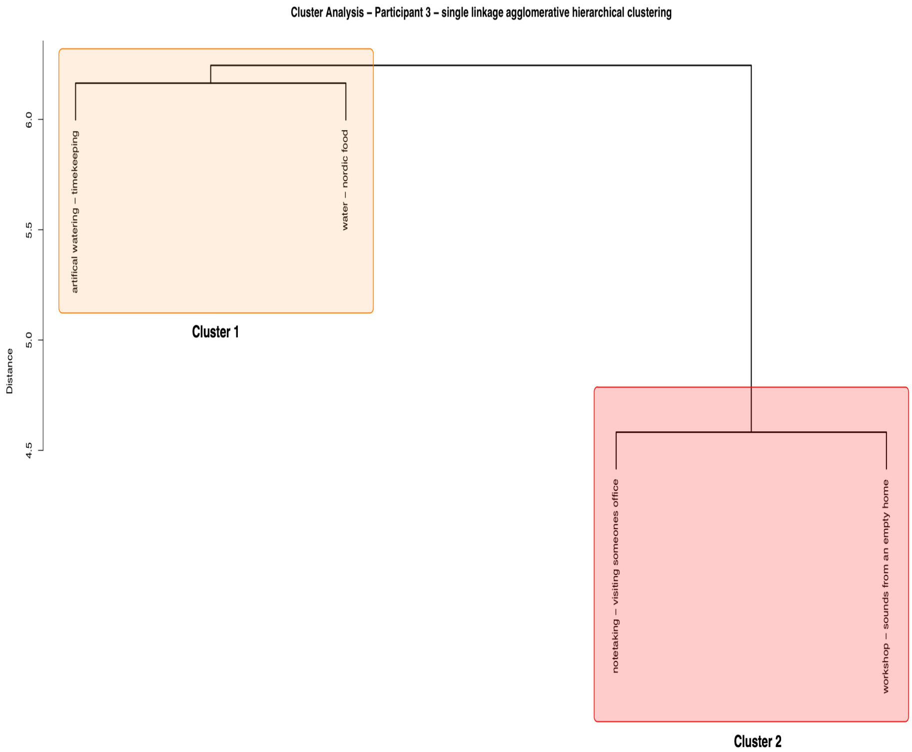


Fig. H-26*: The cluster analysis of the RGT constructs for Participant 3.*

	Constructs *Emergent Pole* - *Implicit Pole*
1	artificial watering–timekeeping water–nordic food
2	notetaking–visiting someones office workshop–sounds from an empty home

Table H-29*:* Clusters obtained for Participant 3.

Both of the clusters for this participant are quite similar and would indicate that there is a continuum based on activity and locational context which was used to organise the sounds. This is supported by Figure H-27 and MST MDS analysis. The differences between the clusters are important as the shorter the distance between clusters, the more significant and stronger the statements that can be made, based on the participant's data.

The MST multidimensional scaling analysis for Participant 3 is shown in Figure H-27, with its related Shepard diagram in Figure H-28. The horseshoe like shape pattern in Figure H-27 indicates that is a activity – object / location continuum which Participant 3 used for their classification of constructs. In the case of Participant 3, the current results are interesting as they show the potential for an individualised continuum based on the sound but to gain further detail, more detailed explorations are required.

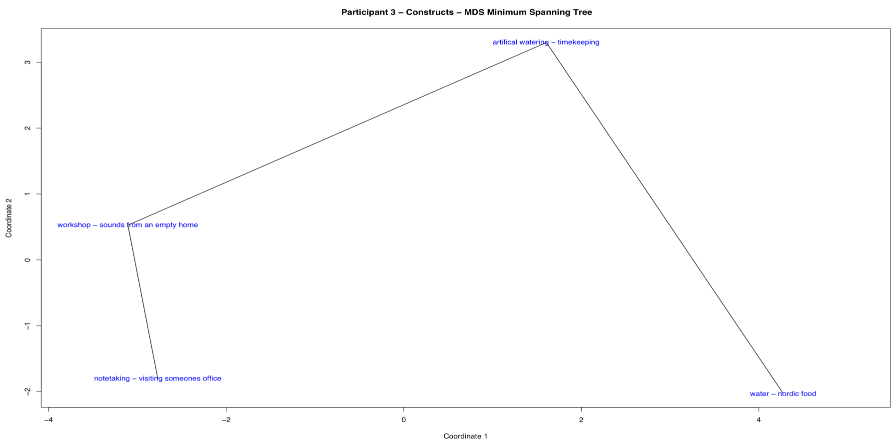


Fig. H-27: *MDS minimum spanning tree analysis of the RGT constructs for Participant 3.*

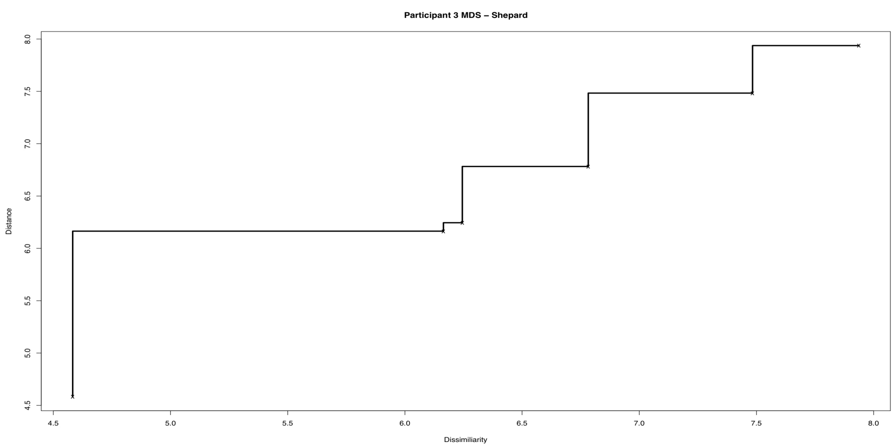


Fig. H-28: *Shepard diagram of the MDS analysis of the RGT constructs for Participant 3.*

Participant 3 - Cluster Analysis and MDS - Elements

The cluster analysis of the elements from Participant 3 is shown in Figure H-29 which resulted in 3 clusters. The resulting clusters are shown in Table H-30. The distances of the clusters were Cluster 1 had a distance of 0.98, Cluster 2 had a distance of 1.24, and Cluster 3 was found at a distance of 0.72. Cluster 1 contained sounds that were seen as either impact, scraping, or deformation sounds. The shortest distance match in this cluster between s1 (the sound of a gas stove being started and idling) and s2 (The sound of a bottling plant's machinery in operation). Cluster 2 contained a single sound, s9 (the sound of sandpaper being rubbed) which was seen as an a friction or rubbing type sound. Cluster 3 contained sounds where were seen as either being continuous or repetitive patterns.

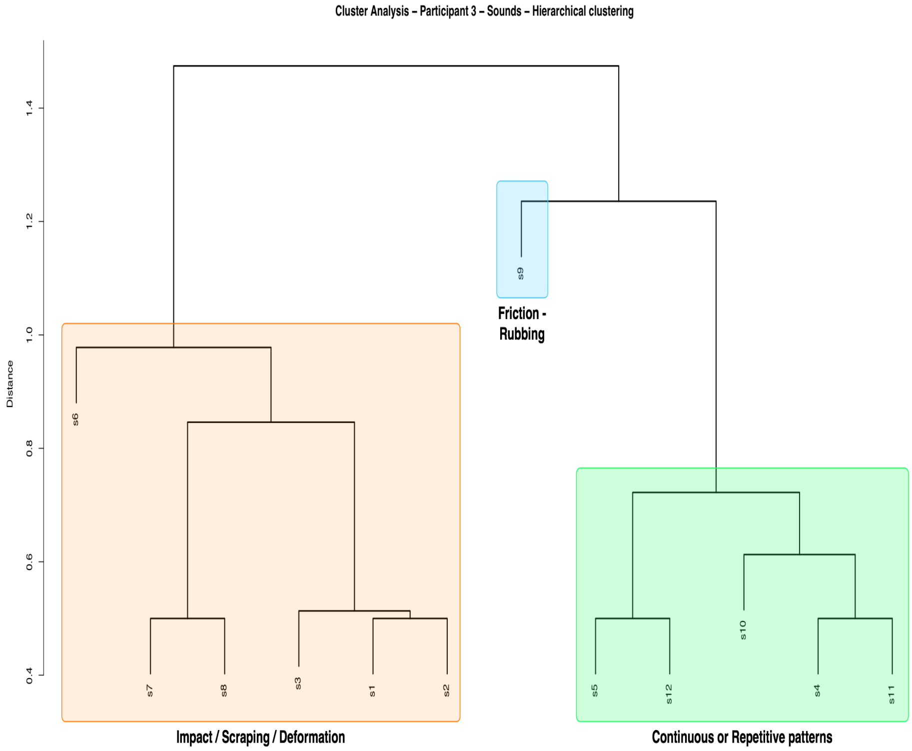


Fig. H-29: *The cluster analysis of the RGT elements (sounds) for Participant 3.*

	Cluster "label" *Stimuli Number*
1	Impact/Scraping/Deformation *s6, s7, s8, s3, s1, s2*
2	Friction/Rubbing *s9*
3	Continuous or Repetitive patterns *s5, s12, s10, s4, s11*

Table H-30: Clusters obtained for the elements of Participant 3.

The MST multidimensional scaling analysis for Participant 3's elements is shown in Figure H-30, with its related Shepard diagram in Figure H-31. The simplex or horseshoe like shape pattern in Figure H-30 indicates that is a friction – impact versus continuous sound continuum which Participant 3 used for their classification of the elements or sounds presented.

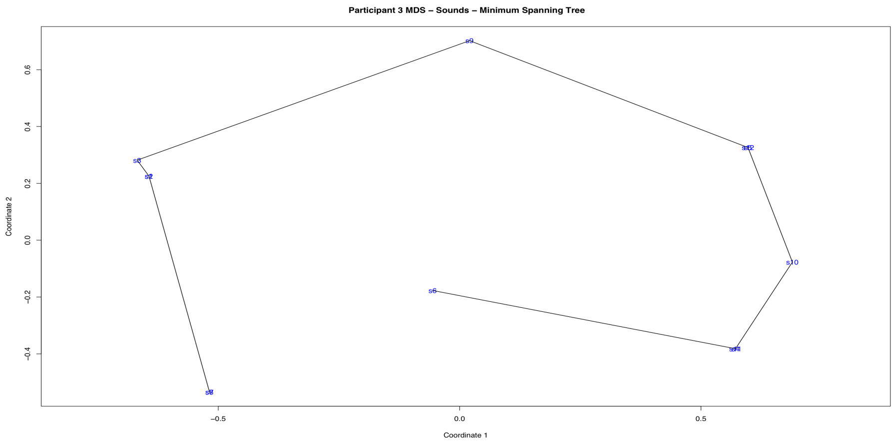


Fig. H-30: *MDS minimum spanning tree analysis of the RGT elements for Participant 3.*

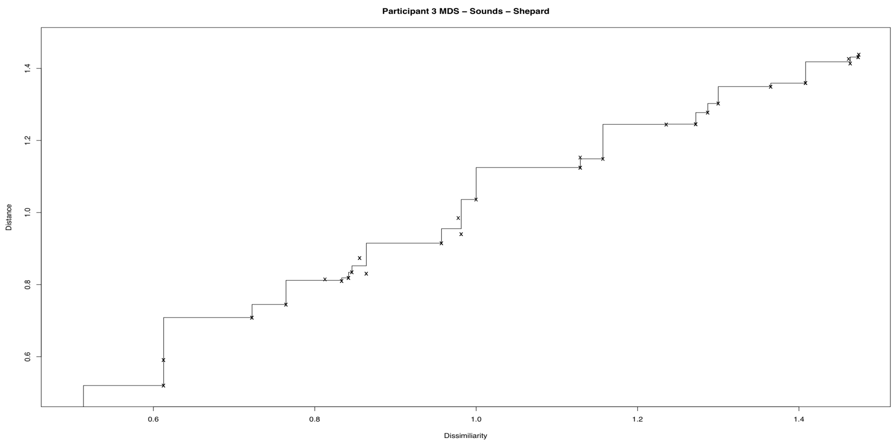


Fig. H-31: *Shepard diagram of the MDS analysis of the RGT elements for Participant 3.*

H.1.6 Results and Observations for Participant 3

The results of the analysis of the constructs from Participant 3 from the PCA, CA, and MDS analysis suggest that after the natural or manmade nature of a sounds is taken into account, the sound's source is the next most important factor. This is supplemented by the clustering results where clusters were somewhat based on the location or type of activities occurring and how the sounds fitted into the particular context. The results of the analysis of the elements (sounds) from Participant 3 from the PCA, CA, and MDS analysis indicate that it the flowing or repetitive pattern versus non-reptitive or non-continuous nature of sounds which suggested a repetitive–non-repetitive scale were also judged on a further impact–friction axis or scale.

Participant 4 - PCA - Constructs

The results of the elicitation task for Participant 4 are shown below in Table H-31. It is useful to note which of the stimuli were used in the particular triadic comparison (see Table 6.3 for the full stimuli list with descriptions) which created the poles for this participant's derived scale. The rating task used these poles for its four sub-tasks.

	Constructs **Emergent Pole – Implicit Pole**	**Similar — Different Sounds**
1	constructed sounds–easy recognisable	*s5, s8 — s10*
2	concrete actions–background / general sound	*s3, s9 — s11*
3	real–imaginary	*s6, s7 — s1*
4	noise–pleasant sound	*s2, s4 — s12*

Table H-31*:* Descriptions of sounds by Participant 4 for the construct poles.

Examining the correlations among the constructs for Participant 4, shown in Table H-32, we can say that "*constructed sounds–easy recognisable*" (construct 1) was definitely not associated with "*real–imaginary*", which was constructs 3 (−0.84). In the case of this participant there were no other constructs with significant correlations. These types of correlations can help in giving an insight into the participants world view of the sounds presented. In order to reduce the amount of information present in Table H-32, we can use the root mean square correlation among constructs as shown in Table H-33. The third construct "*real–imaginary*", was the construct which, was most closely associated with the other constructs as shown in Table H-33.

	1	2	3	4
1 constructed sounds–easy recognisable	1.00	0.09	−**0.84**	0.24
2 concrete actions–background / general sound		1.00	0.23	0.35
3 real–imaginary			1.00	−0.17
4 noise–pleasant sound				1.00

Table H-32*:* Correlation Analysis for Participant 4 Constructs.

Construct	Root-mean-square correlation
constructed sounds–easy recognisable	0.66
concrete actions–background / general sound	0.54
real–imaginary	**0.67**
noise–pleasant sound	0.55
Average of statistic	0.61
Standard deviation of statistic	0.07

Table H-33*:* Root-mean-square (average) correlation among constructs analysis for Participant 4

The interpretation of the first principal component in Table H-34 contrasts the perceived nature of the sounds whether the sound is real world sound or a more abstract imaginary or evocative type of sound. The second principal component in Table H-34 contrasts the density of interactions which occurred within the sounds, implying that after the real or imaginary nature of the sounds is taken into account, the main source of variation between the sounds was between those sounds with more events or interactions occurring and those which were had less occurring in them (these were heard as more pleasant). The first two components are shown in Figure H-32, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-33.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
constructed sounds–easy recognisable	-0.63	-	0.04	
concrete actions–background / general sound	0.15		-0.86	-
real–imaginary	0.64	+	-0.10	
noise–pleasant sound	-0.42	-	-0.51	-
Standard deviation	1.5304		1.0606	
Proportion of Variance	0.5855		0.2812	
Cumulative Proportion	0.5855		0.8667	

Table H-34*:* Principal-Components Analysis for Participant 4 Constructs.

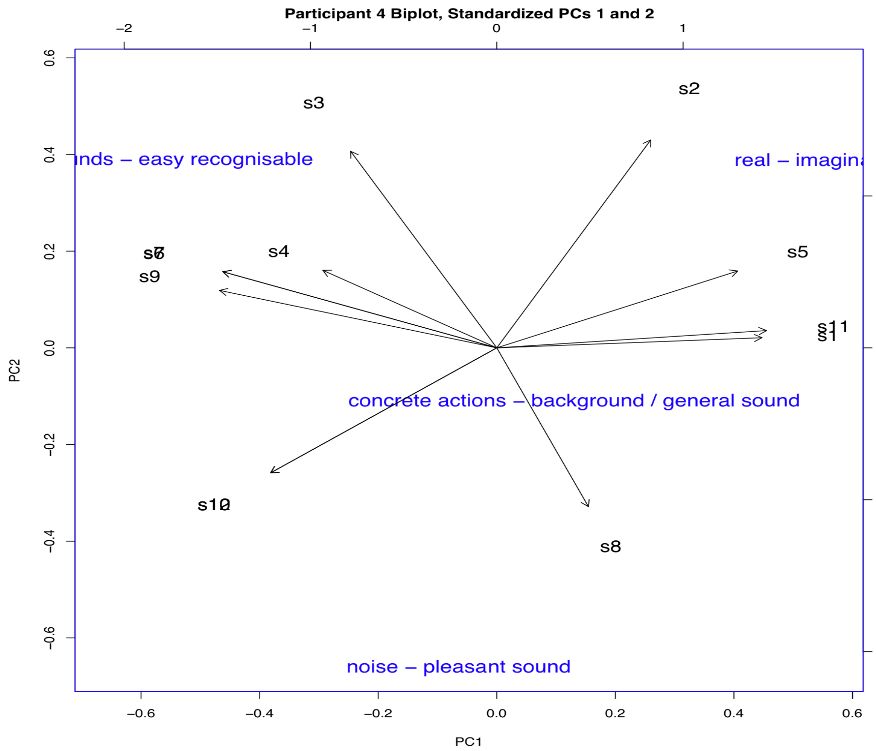


Fig. H-32*: The first two components of the principal–components analysis for Participant 4.*

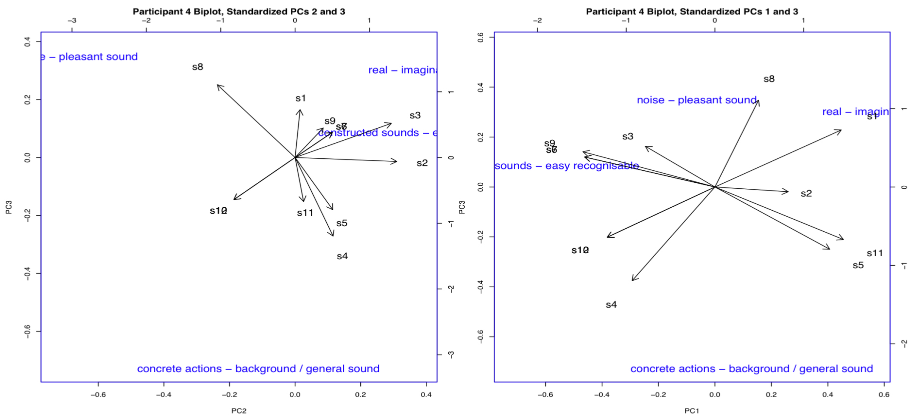


Fig. H-33*: The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 4.*

Participant 4 - PCA - Elements

The results of the elicitation task for Participant 4 are shown below in Table H-35.

ID	Sound Description	Participant's Description
s1	gas stove	*a space walk*
s2	bottling machinery	*cargo door of an old space shuttle*
s3	cutting machinery	*a machine starting and doing something with pebbles*
s4	electronic alarm clock	*alarm clock*
s5	gas expelling	*szzz - sound which could be smoke or possibly air vent*
s6	knocking on door	*knocking and opening of a door*
s7	turning paper	*turning pages of a book with brute force*
s8	rubbing and writing	*sketching on paper*
s9	rubbing sandpaper	*sandpaper*
s10	stream, water flowing	*pouring water - river*
s11	water dripping	*irregular ticking*
s12	water pouring, bath	*pouring water in a small stream*

Table H-35*:* Descriptions by Participant 4 for the elements.

Examining the correlations among the elements for Participant 4 shown in Table H-36, we can say that element *s6* (the sound of somebody knocking on a wooden door) was definitely associated with *s7 (1.00)* (the sound of paper being turned) and *s9 (1.00)* (the sound of sand paper being rubbed) while being definitely not associated with *s11* (-0.91) (the sound of water dripping). As previously stated, we will only consider the one element or sound and the elements with the most significant correlations to this element, in order to highlight the points of interest from the participant's data.

The element *s1* (the sound of gas stove being lit and idling) was the most closely associated element with the other elements as shown in Table H-37. The elements *s6* (the sound of somebody knocking on a door), *s7* (the sound of paper or pages being turned), and *s9* (the sound of sand paper being rubbed) were the joint next most closely associated elements.

The interpretation of the first principal component in Table H-38 contrasts the 'foreground' and the 'background' nature of the sounds and the sound events where the 'foreground' sounds are those which would be familiar to the participant as previously performed human activities while the 'background' sounds are those which do not require a direct human interaction and hence are more in background of a sonic scene. The second principal

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12
s1	1.00	0.48	−0.25	**−0.85**	0.51	−0.70	−0.70	0.56	−0.69	**−0.88**	0.62	**−0.88**
s2		1.00	0.43	0.00	0.71	−0.22	−0.22	−0.43	−0.29	**−0.82**	0.54	**−0.82**
s3			1.00	0.30	−0.30	0.78	0.78	−0.45	0.74	−0.17	−0.52	−0.17
s4				1.00	0.00	0.46	0.46	**−0.90**	0.41	0.58	−0.19	0.58
s5					1.00	−0.76	−0.76	−0.30	**−0.82**	−0.58	**0.96**	−0.58
s6						1.00	**1.00**	−0.32	**1.00**	0.44	**−0.91**	0.44
s7							1.00	−0.32	**1.00**	0.44	**−0.91**	0.44
s8								1.00	−0.25	−0.17	−0.06	−0.17
s9									1.00	0.47	**−0.94**	0.47
s10										1.00	−0.56	**1.00**
s11											1.00	−0.56
s12												1.00

Table H-36: Correlation Analysis for Participant 4 Elements.

Construct	Root-mean-square correlation
s1	**0.68**
s2	0.50
s3	0.49
s4	0.48
s5	0.61
s6	0.67
s7	0.67
s8	0.41
s9	0.67
s10	0.59
s11	0.65
s12	0.59
Average of statistic	0.58
Standard deviation of statistic	0.09

Table H-37*:* Root-mean-square (average) correlation among elements analysis for Participant 4

component in Table H-38 suggests that the next main source of variation was due to the type of interaction involved in the sounds between those which were impact versus those that were deformation type sounds. The first two components are shown in Figure H-34, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-35.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
s1	−0.33	-	0.02	
s2	−0.24	-	−0.50	-
s3	0.22	+	−0.44	-
s4	0.27	+	−0.27	(-)
s5	−0.31	-	−0.22	(-)
s6	0.32	+	−0.11	
s7	0.32	+	−0.11	
s8	−0.16	(-)	0.55	+
s9	0.33	+	−0.07	
s10	0.30	+	0.22	(+)
s11	−0.32	-	−0.09	
s12	0.30	+	0.22	(+)
Standard deviation	2.9664		1.3949	
Proportion of Variance	0.7333		.1622	
Cumulative Proportion	0.7333		0.89543	

Table H-38: Principal-Components Analysis for Participant 4 Elements.

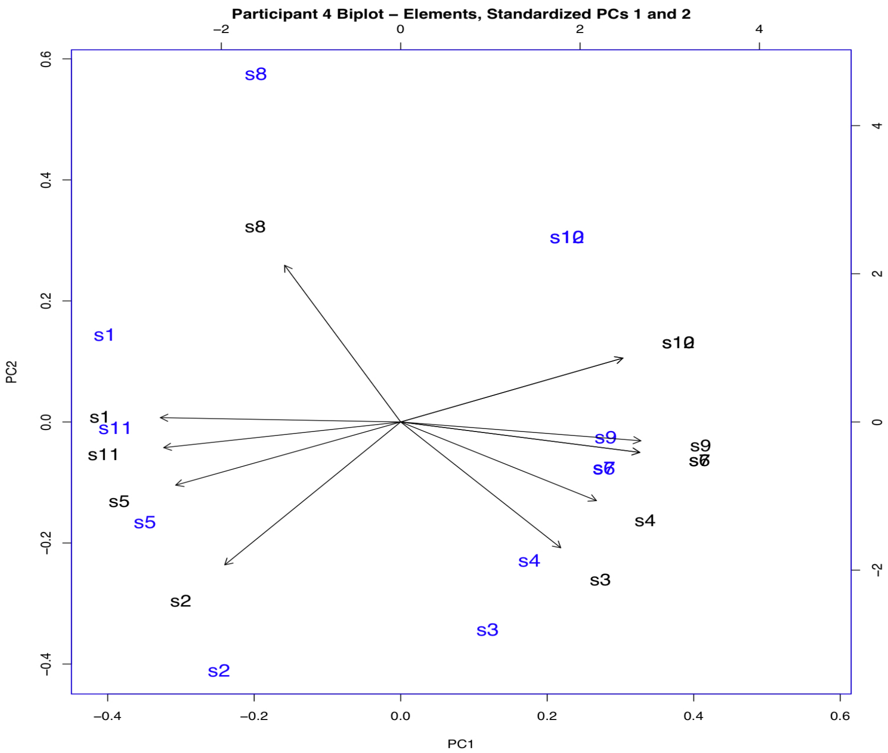


Fig. H-34: *The first two components of the principal–components analysis for Participant 4 Elements.*

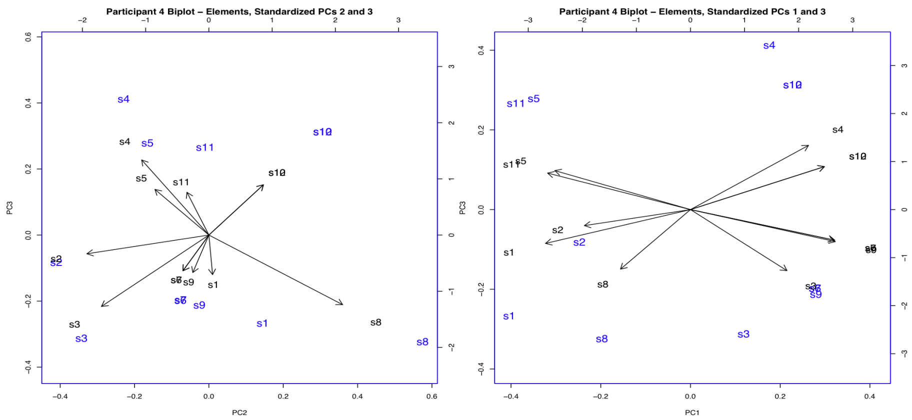


Fig. H-35: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 4 Elements.*

Participant 4 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 4 is shown in Figure H-36 which resulted in 3 clusters. The resulting clusters are shown in Table H-39, Cluster 1 was found at a distance of 7.21, Cluster 2 was found at a distance of 6.71, and Cluster 3 was found at a distance of 7.75. Cluster 1 contained sounds that were seen as either being “constructed” sounds or easily recognisable sounds. The shortest distance match in this cluster was at 7.21 between "*constructed sounds – easy recognisable*". Cluster 2 contained sounds that were seen as either background sounds or as active human activity sounds. The shortest distance match in this cluster was at 6.71 between "*concrete actions – background / general sound*" and "*noise – pleasant sound*". Cluster 3 contained sounds that heard as either realistic or imaginary with the distance match for this cluster between "*real – imaginary*" being 7.75. These constructs suggest the sounds are clustered somewhat based on the type of interactions occurring within the sounds with human activity type sounds versus background sounds and being further judged based on whether the sounds were constructed or imaginary versus their realism or recognisability. In the case of Participant 4, the current results are interesting as they show the potential for an individualised continuum based on the sound but to gain further detail, more detailed explorations are required.

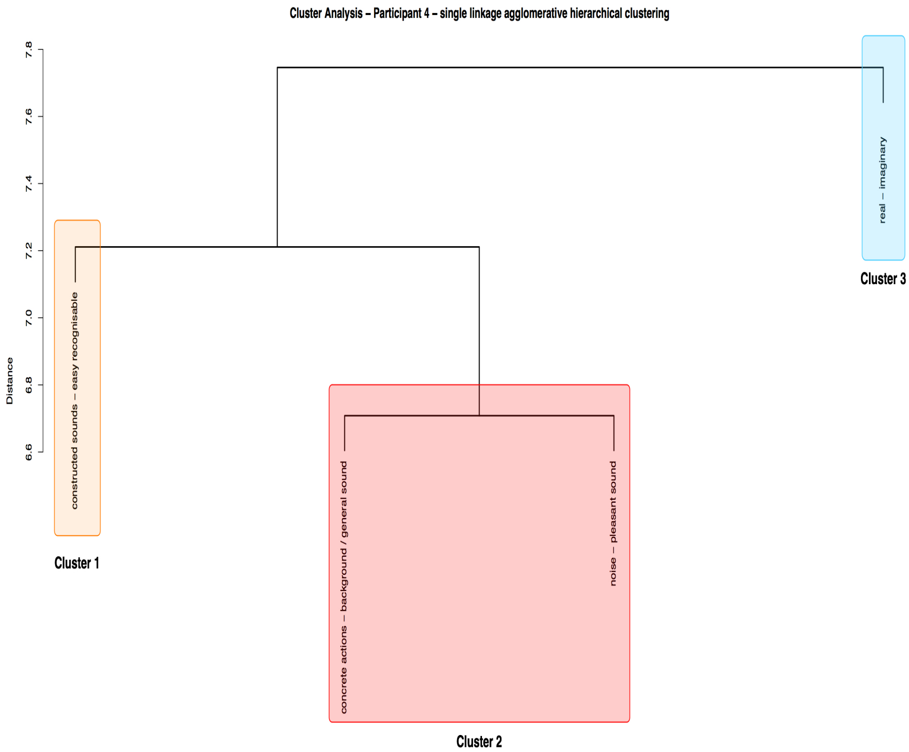


Fig. H-36: *The cluster analysis of the RGT constructs for Participant 4.*

The MST multidimensional scaling analysis for Participant 4 is shown in Figure H-37, with its related Shepard diagram in Figure H-38. The MST multidimensional scaling analysis shows a simplex type pattern meaning that only a single dimension is needed to account for the constructs. It is likely for Participant 4, that this can be seen as real to created type of sound continuum. In the case of Participant 4, the current results are interesting as they show

	Constructs *Emergent Pole* - *Implicit Pole*
1	constructed sounds–easy recognisable
2	concrete actions–background / general sound noise–pleasant sound
3	real–imaginary

Table H-39*:* Clusters obtained for Participant 4.

the potential for an individualised continuum based on the sound but to gain further detail, more detailed explorations are required.

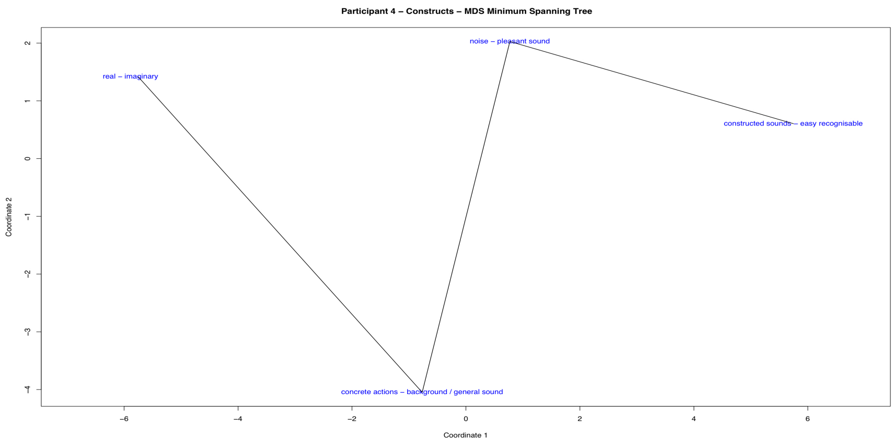


Fig. H-37*: MDS minimum spanning tree analysis of the RGT constructs for Participant 4.*

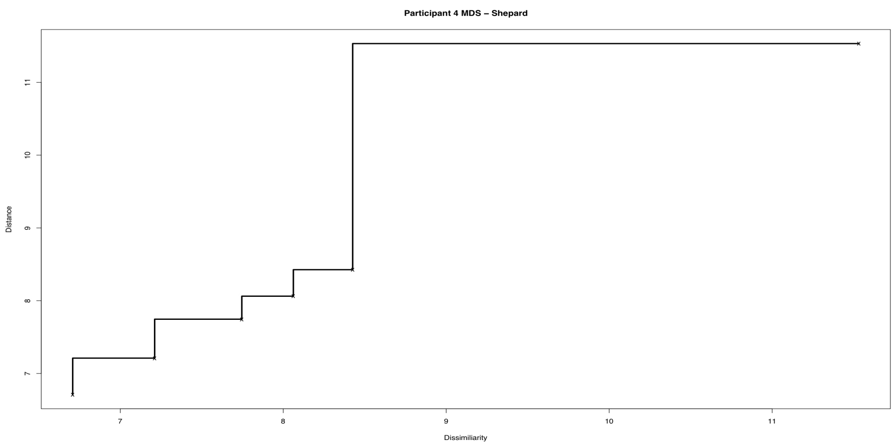


Fig. H-38*: Shepard diagram of the MDS analysis of the RGT constructs for Participant 4.*

Participant 4 - Cluster Analysis and MDS - Elements

The cluster analysis of the elements from Participant 4 is shown in Figure H-39 which resulted in 4 clusters. The resulting clusters are shown in Table H-40. The distances of the clusters were Cluster 1 had a distance of 0.63, Cluster 2 had a distance of 0.71, Cluster 3 had a distance of 0.72, and Cluster 4 was found at a distance of 0.73. The resulting clusters are shown in Table H-40. Cluster 1 contained sounds that were seen as deformation, impact or friction type sounds. The shortest distance match in this cluster between s6 (The sound of knocking on a door) and s7 (The sounds of pages being turned). Cluster 2 contained sounds that were seen as continuous, repetitive patterns. The shortest distance match in this cluster between s10 (The sound of a stream flowing) and s12 (The sound of water being poured into a bath). Cluster 3 contained two sounds which were seen as gaseous continuous type sounds. Cluster 4 contained three sounds, s2 (The sound of a bottling plant's machinery), s5 (The sound of gas being expelled from a gas cylinder), and s11 (The sound of water dripping). These were seen as continuous impact type sounds.

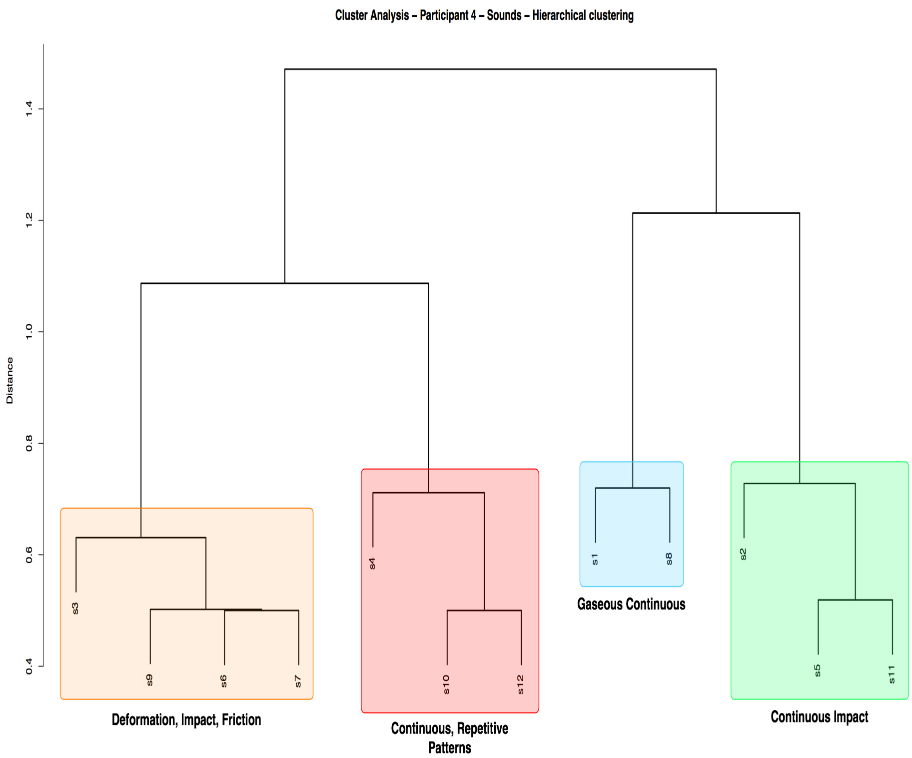


Fig. H-39: *The cluster analysis of the RGT elements (sounds) for Participant 4.*

The MST multidimensional scaling analysis for Participant 4's elements is shown in Figure H-40, with its related Shepard diagram in Figure H-41. The horseshoe like shape pattern in Figure H-40 indicates that is a potential deformation – friction / impact continuum which

	Cluster "label" *Stimuli Number*
1	Deformation, Impact, Friction *s3, s9, s6, s7*
2	Continuous, Repetitive Patterns *s4, s10, s12*
3	Gaseous Continuous *s1, s8*
4	Continuous Impact *s2, s5, s11*

Table H-40*:* Clusters obtained for the elements of Participant 4.

Participant 4 used for their classification of the elements or sounds presented.

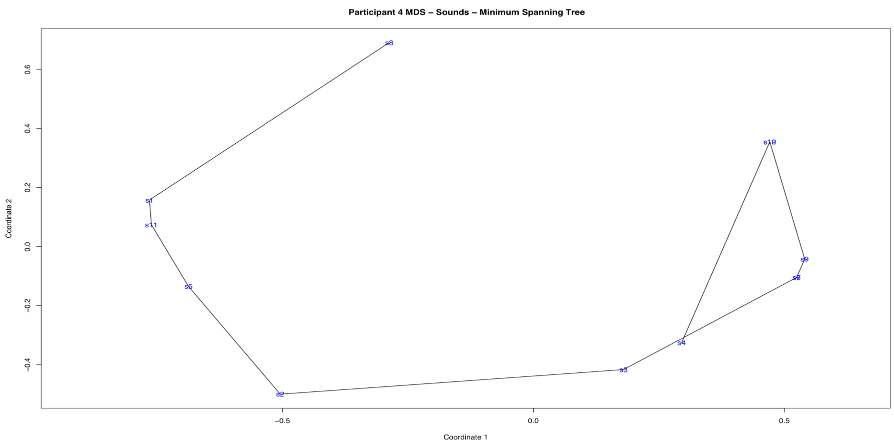


Fig. H-40*: MDS minimum spanning tree analysis of the RGT elements for Participant 4.*

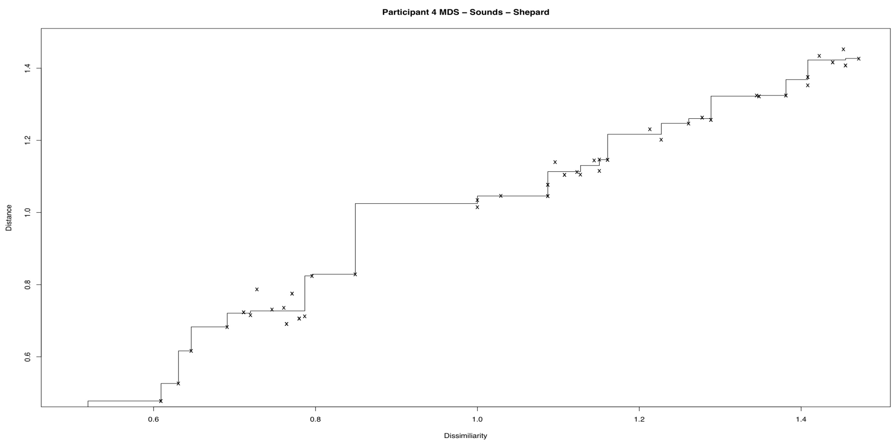


Fig. H-41*: Shepard diagram of the MDS analysis of the RGT elements for Participant 4.*

H.1.7 Results and Observations for Participant 4

The results of the analysis of the constructs from Participant 4 from the PCA, CA, and MDS analysis indicate that after the real or imaginary nature of the sounds is taken into account, the main source of variation between the sounds was between those sounds with more events or interactions occurring and those which were had less occurring in them (these were heard as more pleasant). The results from the constructs also pointed that that the constructed or imaginary nature of a sound versus its realism or recognisability was a factor. The results of the analysis of the elements (sounds) from Participant 4 from the PCA, CA, and MDS analysis suggested a foreground–background scale where those which would be familiar to the participant as previously performed human activities would be the 'foreground' while the 'background' sounds are those which do not require a direct human interaction. This scale had an additional axis or scale, that of the sounds position on impact–deformation continuum.

Participant 5 - PCA - Constructs

The results of the elicitation task for Participant 5 are shown below in Table H-41. It is useful to note which of the stimuli were used in the particular triadic comparison (see Table 6.3 for the full stimuli list with descriptions) which created the poles for this participant's derived scale. The rating task used these poles for its four sub-tasks.

	Constructs **Emergent Pole – Implicit Pole**	**Similar — Different Sounds**
1	shaky sounds–percussive	*s3, s9 — s6*
2	danger–no harm	*s2, s4 — s7*
3	water / nature–humankind thing	*s11, s12 — s8*
4	air–water	*s1, s5 — s10*

Table H-41*:* Descriptions of sounds by Participant 5 for the construct poles.

Examining the correlations among the constructs for Participant 5, shown in Table H-42, we can say that "*water / nature – humankind thing*" (construct 3) were definitely not associated with "*air – water*" construct 4 (-0.71). Each construct can be considered in a similar manner to the third construct with regard to the other constructs, rather than consider each of the constructs individually we will present the significant correlations as these are the items we can make the strongest statements about, based on the participant's data. In the case of this participant there were no other constructs with significant correlations. These types of correlations can help in giving an insight into the participants world view of the sounds presented. In order to reduce the amount of information present in Table H-42, we can use the root mean square correlation among constructs as shown in Table H-43. The "*air – water*" was the construct which was most closely associated with the other constructs as shown in Table H-43.

	1	2	3	4
1 shaky sounds–percussive	1.00	−0.05	−0.26	0.59
2 danger–no harm		1.00	−0.50	0.57
3 water / nature–humankind thing			1.00	−**0.71**
4 air–water				1.00

Table H-42*:* Correlation Analysis for Participant 5 Constructs.

Construct	Root-mean-square correlation
shaky sounds–percussive	0.59
danger–no harm	0.63
water / nature–humankind thing	0.67
air–water	**0.74**
Average of statistic	0.66
Standard deviation of statistic	0.06

Table H-43*:* Root-mean-square (average) correlation among constructs analysis for Participant 5

The interpretation of the first principal component in Table H-44 contrasts a mix of mechanical nature and danger or uncertainty about the sounds with safer more naturalistic sounds. The second principal component in Table H-44 is similar to the first component but with more focus on the type of impact or deformation occurring within the sounds and how this interaction fits within the participant's constructs of safe or dangerous. The first two components are shown in Figure H-42, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-43.

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
shaky sounds–percussive	−0.36	-	0.78	+
danger–no harm	−0.46	-	−0.62	-
water / nature–humankind thing	0.58	+	0.07	
air–water	−0.57	-	0.07	
Standard deviation	1.7336		0.9905	
Proportion of Variance	0.7514		0.2453	
Cumulative Proportion	0.7514		0.9966	

Table H-44*:* Principal-Components Analysis for Participant 5 Constructs.

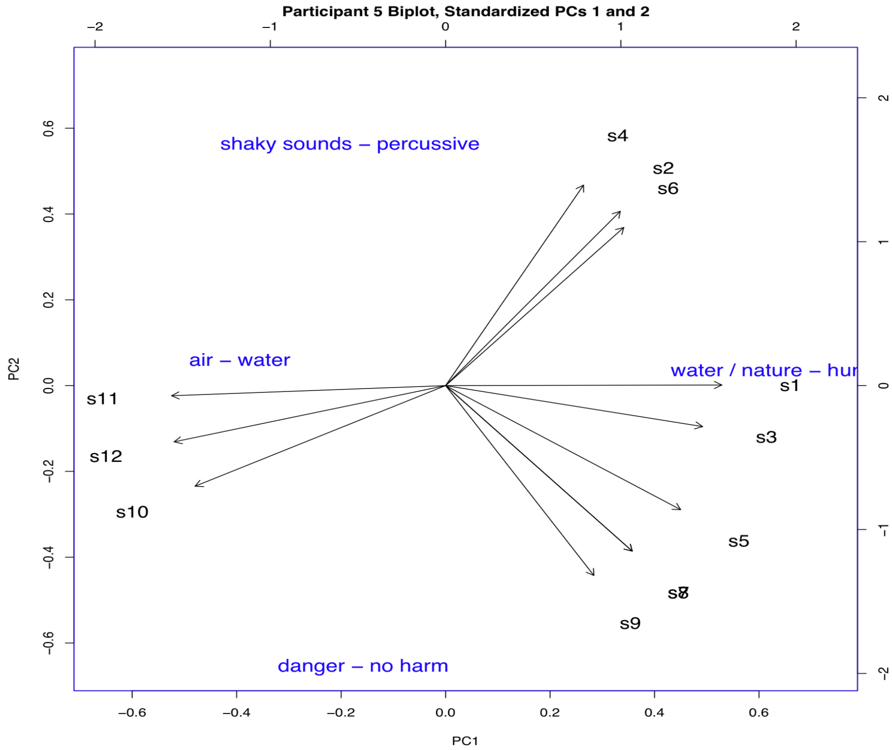


Fig. H-42: *The first two components of the principal–components analysis for Participant 5.*

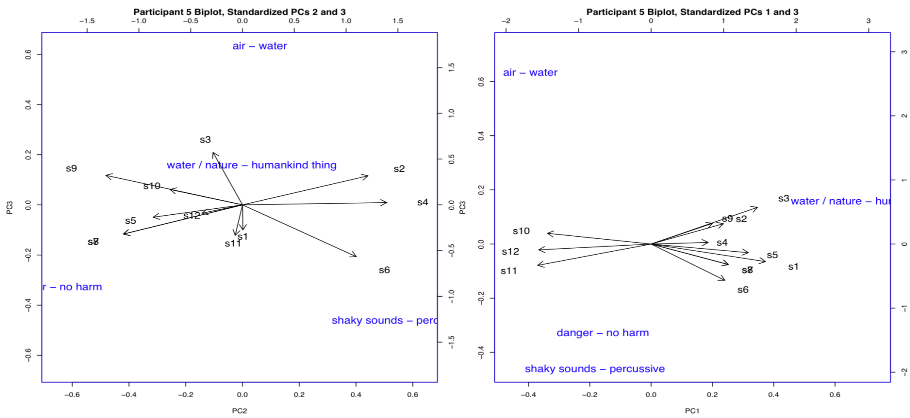


Fig. H-43: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 5.*

Participant 5 - PCA - Elements

The results of the elicitation task for Participant 5 are shown below in Table H-45.

ID	Sound Description	Participant's Description
s1	gas stove	*air flow*
s2	bottling machinery	*machinery activating*
s3	cutting machinery	*some industrial machinery*
s4	electronic alarm clock	*alarm*
s5	gas expelling	*air pressure*
s6	knocking on door	*door*
s7	turning paper	*paper*
s8	rubbing and writing	*scratching*
s9	rubbing sandpaper	*shaken*
s10	stream, water flowing	*water*
s11	water dripping	*drops*
s12	water pouring, bath	*water - river*

Table H-45*:* Descriptions by Participant 5 for the elements.

Examining the correlations among the elements for Participant 5 shown in Table H-46, we can say that element *s2* (the sound of bottling machinery in operation) was definitely associated with *s4 (0.97)* (The sound of an electronic alarm clock bleeping) and *s6 (0.85)* (a knocking on door sound) while being definitely not associated with *s10* (-0.87) (the sound of a stream with water flowing) and *s12* (-0.80) (the sound of water pouring). The element *s1* (the sound of a gas stove being lit and flickering) was the most closely associated element with the other elements as shown in Table H-47. The element *s11* (the sound of gas being expelled from a large canister or cylinder) was the next most closely associated element.

The interpretation of the first principal component in Table H-48 contrasts the source of the sounds with water sounds being contrasted against the other sounds. The second principal component in Table H-48 suggests that the next main source of variation was due to the type of interaction occurring with the sounds where impact repetitive sounds were contrasted with continuous friction sounds. The first two components are shown in Figure H-44, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure H-45.

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12
s1	1.00	0.58	**0.84**	0.48	**0.84**	0.69	0.69	0.69	0.48	−**0.90**	−**0.93**	−**0.94**
s2		1.00	0.51	**0.97**	0.10	**0.85**	−0.17	−0.17	−0.25	−**0.87**	−0.68	−**0.80**
s3			1.00	0.30	**0.83**	0.33	0.67	0.67	0.70	−0.70	−**0.96**	−**0.86**
s4				1.00	−0.06	**0.90**	−0.30	−0.30	−0.45	−**0.82**	−0.52	−0.69
s5					1.00	0.19	**0.96**	**0.96**	**0.87**	−0.52	−0.78	−0.68
s6						1.00	0.00	0.00	−0.30	−**0.90**	−0.58	−0.76
s7							1.00	**1.00**	**0.90**	−0.30	−0.58	−0.46
s8								1.00	**0.90**	−0.30	−0.58	−0.46
s9									1.00	−0.09	−0.52	−0.32
s10										1.00	**0.87**	**0.97**
s11											1.00	**0.97**
s12												1.00

Table H-46: Correlation Analysis for Participant 5 Elements.

Construct	Root-mean-square correlation
s1	**0.76**
s2	0.58
s3	0.70
s4	0.57
s5	0.65
s6	0.54
s7	0.59
s8	0.59
s9	0.56
s10	0.69
s11	0.75
s12	0.74
Average of statistic	0.64
Standard deviation of statistic	0.08

Table H-47*:* Root-mean-square (average) correlation among elements analysis for Participant 5

	PC1	*PC1 Simplified*	PC2	*PC2 Simplified*
s1	-0.34	-	0.00	
s2	-0.26	-	0.36	+
s3	-0.34	-	-0.08	
s4	-0.21	-	0.43	+
s5	-0.30	-	-0.25	-
s6	-0.26	-	0.34	+
s7	-0.25	-	-0.36	-
s8	-0.25	-	-0.36	-
s9	-0.20	-	-0.44	-
s10	0.32	+	-0.19	(-)
s11	0.34	+	-0.02	
s12	0.33	+	-0.11	
Standard deviation	2.9295		1.8355	
Proportion of Variance	0.7151		0.2808	
Cumulative Proportion	0.7151		0.9959	

Table H-48*:* Principal-Components Analysis for Participant 5 Elements.

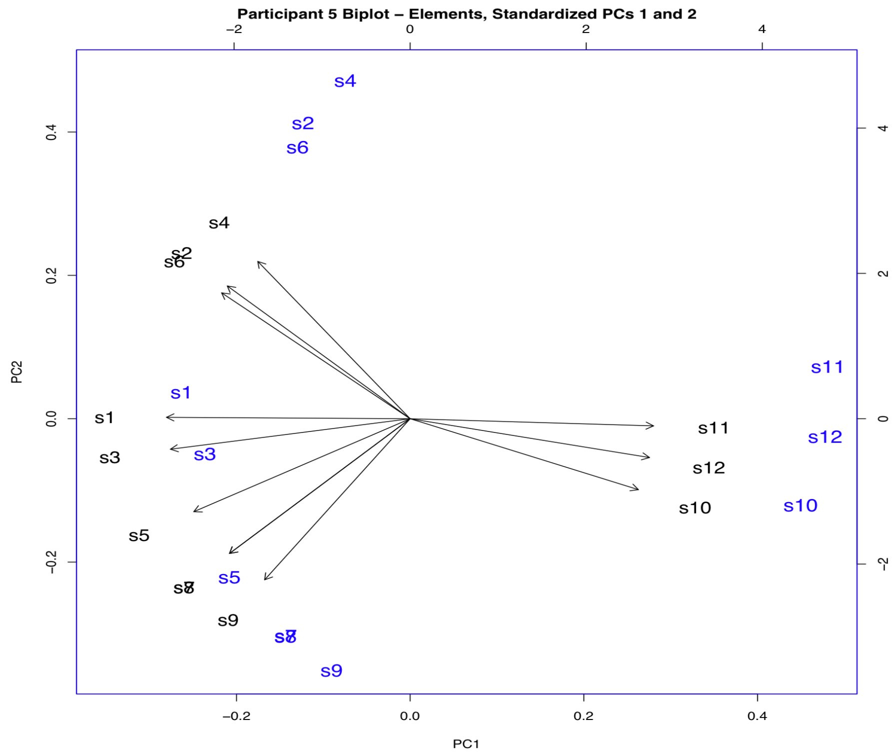


Fig. H-44: *The first two components of the principal–components analysis for Participant 5 Elements.*

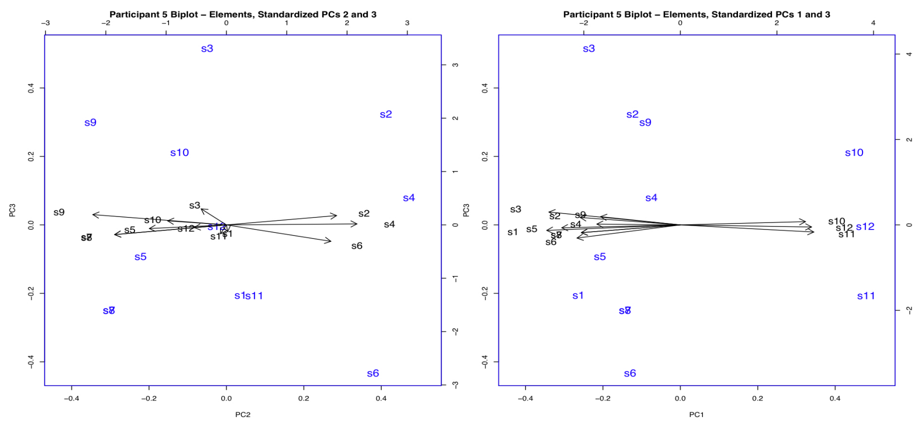


Fig. H-45: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 5 Elements.*

Participant 5 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 5 is shown in Figure H-46 which resulted in 3 clusters. The resulting clusters are shown in Table H-49, Cluster 1 was found at a distance of 4.47, Cluster 2 was found at a distance of 5.00, and Cluster 3 was found at a distance of 9.11. Cluster 1 contained sounds that were seen as either impact sounds or gaseous/liquid sounds.

The shortest distance match in this cluster was at 4.47 between "*shaky sounds – percussive*" and "*air – water*". Cluster 2 contained sounds that were seen as either dangerous or harmless sounds. The shortest distance match in this cluster was at 5.00 between "*danger – no harm*". Cluster 3 contained sounds that were seen as either impact sounds or gaseous/liquid sounds. The shortest distance match in this cluster was at 9.11 between "*water / nature – humankind thing*". These constructs suggest that both the source and the type of interactions occurring within the sounds are used by Participant 5 for clustering. The three clusters for this participant are distinct and would indicate that there are two distinct scales in use by Participant 5 based on the perceived action or result of the action in the sounds or the naturalness of the sounds and their sources.

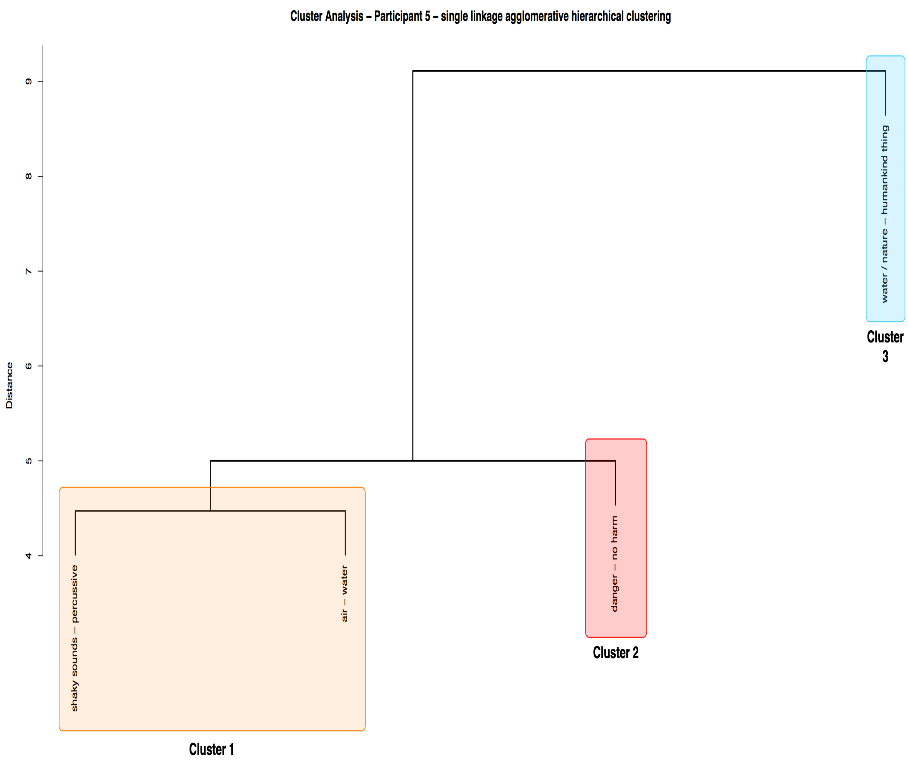


Fig. H-46: *The cluster analysis of the RGT constructs for Participant 5.*

	Constructs *Emergent Pole* - *Implicit Pole*
1	shaky sounds–percussive air–water
2	danger–no harm
3	water / nature–humankind thing

Table H-49: Clusters obtained for Participant 5.

The MST multidimensional scaling analysis for Participant 1 is shown in Figure H-47, with its related Shepard diagram in Figure H-48. The horseshoe like shape pattern in Figure H-50 indicates a single dimension used by Participant 5 for their construct classification, however it is not immediately obvious from the MST graph. At this point, it is useful to look at the cluster analysis to see if it can provide the additional detail required to identify this dimension. In the case of Participant 5, the current results are interesting as they show the potential for two distinct scales but to gain further detail, more detailed explorations are required.

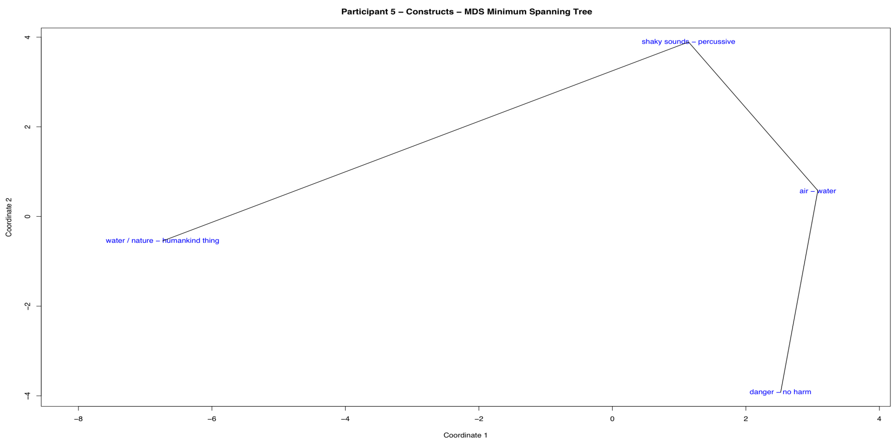


Fig. H-47: *MDS minimum spanning tree analysis of the RGT constructs for Participant 5.*

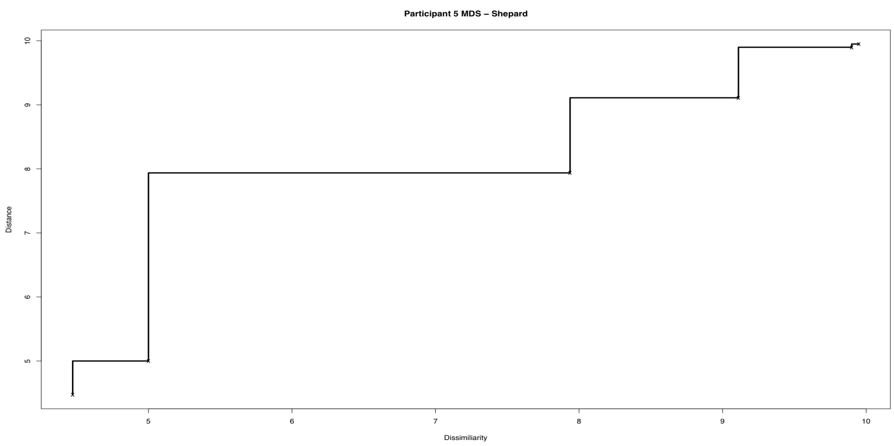


Fig. H-48: *Shepard diagram of the MDS analysis of the RGT constructs for Participant 5.*

Participant 5 - Cluster Analysis and MDS - Elements

The cluster analysis of the elements from Participant 5 is shown in Figure H-49 which resulted in 4 clusters. The resulting clusters are shown in Table H-50. The distances of the clusters were Cluster 1 had a distance of 0.56, Cluster 2 had a distance of 0.58, Cluster 3 had a distance of 0.57, and Cluster 4 was found at a distance of 0.56. The resulting clusters are shown in Table H-50.

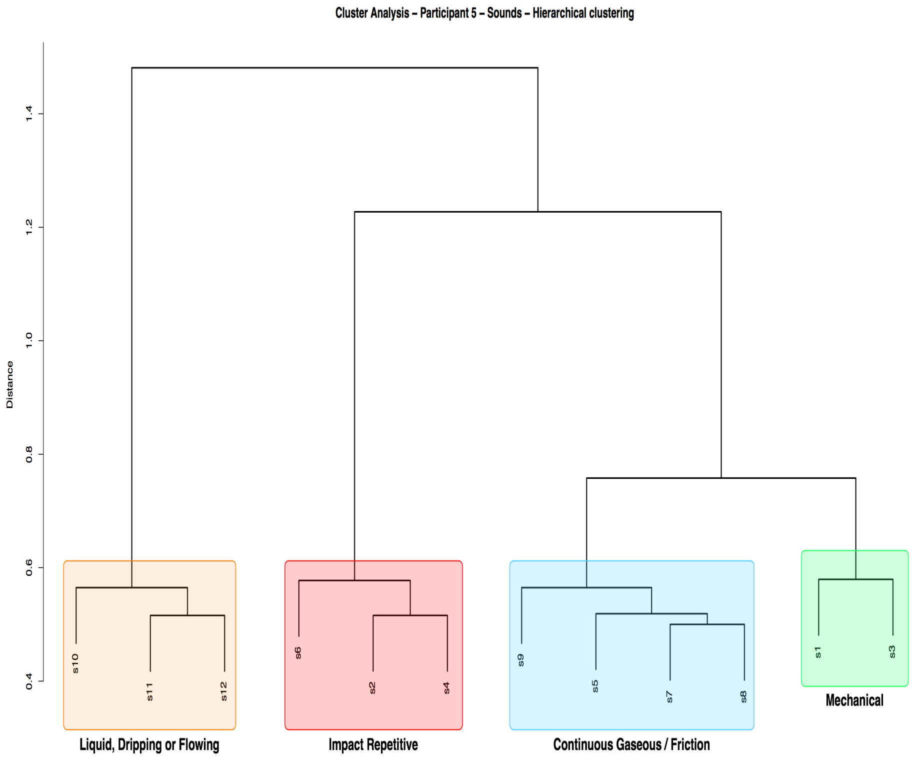


Fig. H-49: *The cluster analysis of the RGT elements (sounds) for Participant 5.*

	Cluster "label" *Stimuli Number*
1	Liquid, Dripping or Flowing *s10, s11, s12*
2	Impact Repetitive *s6, s2, s4*
3	Continuous Gaseous / Friction *s9, s5, s7, s8*
4	Mechanical *s1, s3*

Table H-50: Clusters obtained for the elements of Participant 5.

Cluster 1 contained sounds that were heard as liquid, dripping, or flowing sounds. The shortest distance match in this cluster between s11 (water dripping) and s12 (The sound of water being poured into a bath). Cluster 2 contained sounds where were heard as impact type interactions which were also repetitive sounds. Cluster 3 contained sounds that were heard as having continuous gaseous or friction type interactions. Cluster 4 contained sounds that were of a mechanical nature, the two sounds in this cluster were s1 (the sound of a gas stove starting and idling) and s3 (the sound of cutting machinery).

The MST multidimensional scaling analysis for Participant 5's elements is shown in Figure H-50, with its related Shepard diagram in Figure H-51. The horseshoe like shape pattern in Figure H-50 indicates that is a electromechanical source – naturalistic source continuum, which Participant 5 used for their classification of the elements or sounds presented.

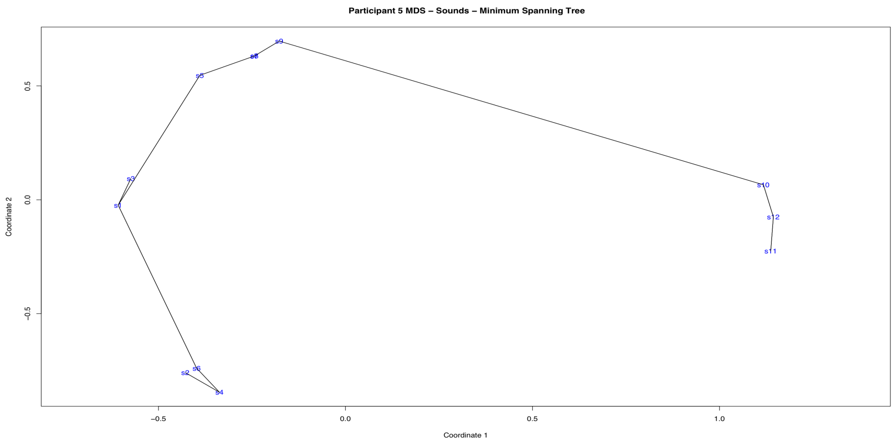


Fig. H-50*: MDS minimum spanning tree analysis of the RGT elements for Participant 5.*

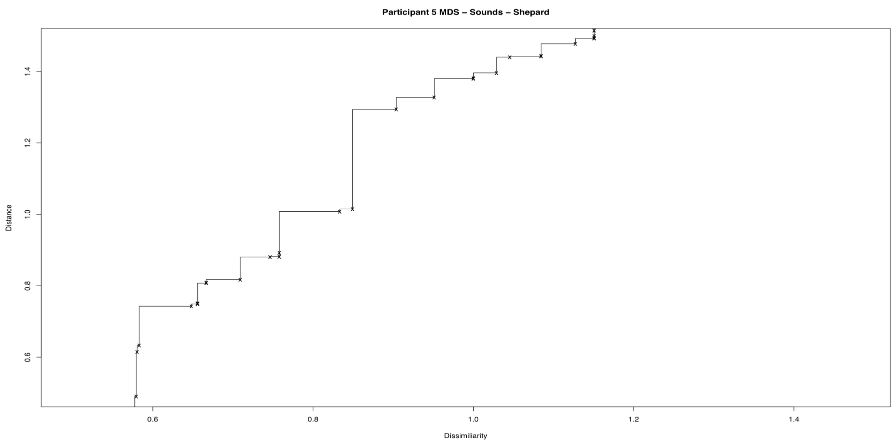


Fig. H-51*: Shepard diagram of the MDS analysis of the RGT elements for Participant 5.*

H.1.8 Results and Observations for Participant 5

The results of the analysis of the constructs from Participant 5 from the PCA, CA, and MDS analysis seem to support that a mix of mechanical nature and danger or uncertainty about the sounds with safer, more naturalistic sounds was one factor, the next factor was between sounds with impacts or deformations occurring within the sounds and how this interaction fits within the participant's constructs of safe or of dangerous. The results of the analysis of the elements (sounds) from Participant 5 from the PCA, CA, and MDS analysis suggested that the source of the sounds with water sounds being contrasted against the other sounds was a factor. An additional factor for the sounds was the type of interaction occurring with the sounds where impact repetitive sounds were contrasted with continuous friction sounds.

I.1 Appendix I - Chapter 6 - An Ambient Auditory Information System for Co-Located Colleagues

This work was influenced by prior systems such as Mauney and Walker (2004) work on stock market monitoring, Peep (Gilfix and Crouch, 2000), WISP (Kilander and Lönnqvist, 2002), Bovermann et al. (2006) work on Ambient Data Displays and Audio Aura (Mynatt et al., 1998). Mauney's system concentrated on dynamically rendering sonifications of real-time stock market data. It used an immersive soundscape consisting of natural sounds with threshold values mapped to trigger certain sounds based on the stock market data. Peep created a network monitoring system using Auditory Icons to represent network events. WISP or the Weakly Intrusive Ambient Soundscape as envisioned by Kilander and Lönnqvist (2002) allows the states and events in the particular computational and physical environment to be presented as subtle and non-intrusive distinct sound cues that conveys information through intuition rather than through interruption. Ambient Data Displays combine ideas from earlier systems with concepts from Tangible Computing (Wisneski et al., 1998). The most influential of these previous systems on the system being developed was the Audio Aura system. This was designed as a serendipitous soundscape for peripheral awareness using background auditory cues.

Presence and place are complex and loaded terms. In this exploration, we have adopted the definition of place as meaning the direct, everyday experience (in the phenomenological) sense of that place and use the group awareness system to investigate one aspect of place, that of *presence - being* in a similar manner as described by Turner et al. (2003). Basso (1996) described how '*places possess a marked capacity for triggering acts of self-reflection, inspiring thoughts about who one presently is, or memories of who one used to be, or musings on who one might become*', that complemented the thinking behind a group awareness system as a mechanism for supporting awareness and lightweight interactions. The exploratory of the system examines the use of auditory representations using Auditory Icons as presence indicators, in order to better understand presence indicators and their overall usefulness for conveying a sense of presence.

The work of Erickson and Kellogg (2000) on *social translucence* and in particular on their concept of a *social proxy* contributed to the design of the new interface. Social translucence is concerned with making social information visible within an interactive system. It has three properties, visibility, awareness, and accountability. It presents social information using an abstract approach which is not tied to an existing analogue, this is the same approach taken in the Out to Lunch system created by Cohen (1994a). A social proxy (Erickson et al., 2002) is a minimalist representation of users that depicts their presence and activities, it is a collective resource aimed as an infrastructure artefact. Social issues are intimately tied with issues regarding privacy and security. Another influence was the work by Pentland (2005) on socially aware platforms which quantify social context in human communication and on collective nature and its implications of communication (Pentland, 2007).

The sound design for the interface used the concept of *ambiguity of information* as defined by Gaver et al. (2003), where expectations are projected onto an interpretation of incomplete information. The system and its sound design exploited this concept to use imprecise auditory representations of users and their activities to emphasise the uncertainty of their availability. The implementation of this concept had an added benefit that it improved users' privacy through the use of a non direct sound mapping to the user and to activities. The mappings between people and their activities are not easily perceived, by encouraging other people to interpret the auditory situation for themselves, it encourages them to grapple with the conceptual sound mapping and to establish a deeper and more personal relation with the meaning offered by the system. Ambiguity is a useful concept for this type of system given how it mines the available information from instant messenger systems and from local network devices to infer a person's particular availability state and location. These types of information are rarely exact and by including an ambiguous design element with regard to the sound mapping, they can be offset somewhat.

Awareness is yet another loaded term from psychology and it is easy to get lost in circular arguments. We have adopted the prevailing concept from the area of ambient / peripheral displays for use within these studies. Awareness has been defined by Dourish and Bellotti (1992) "*an understanding of the activities of others, which provides a context for your own activities*" and expanded by Wisneski et al. (1998) as "*the state of knowing about the environment in which you exist; about your surroundings, and the presence and activities of others*". Lightweight interactions are the type of interactions that are triggered by informal, spontaneous interaction between people. These lightweight interactions or *opportunistic* interactions are the kind that happen when people meet one another when they have something to discuss, such as in the corridor or at the coffee machine. Studies have shown that these informal interactions are useful for getting work done (Isaacs et al., 2002, Kraut et al., 1990, Kraut and Streeter, 1995). Encouraging *awareness moments* as discussed by Nardi et al. (2000) which "*produce a certain feeling in people, rather than accomplishing information exchange ... Awareness moments argue for a richer notion of communication than current media theories allow. Even when no direct information exchange is taking place, people want to maintain connection with others, outside the context of specific events of information exchange*" was one concept that inspired this exploration.

Designing auditory interfaces for awareness and lightweight interactions that function within existing work practices and the existing workplace soundscape (Macaulay and Crerar, 1998) requires new approaches to explore the technologically rich modern work environments, which are burdened with high information and interaction loads. Auditory interfaces are one mechanism for increasing the potential bandwidth for communication in these types of environments.

"Ambient displays", "peripheral systems" or "notification systems" are some of the labels given to the study of systems that "*present information within a space through subtle changes in light, sound, or movement, which can be processed in the background of awareness*" (Wisneski et al., 1998). As the exploratory system we are discussing is an ambient information systems, we will use Pousman and Stasko (2006) definition of ambient information systems characteristics for the behavioral characteristics[3] of the developed system:

- Display information that is **important but not critical**.
- Can **move from the periphery to the focus of attention** and back again.
- Focus on the **tangible**; representations **in the environment**.
- **Provide subtle changes** to reflect updates in information (should not be distracting).
- **Are aesthetically pleasing** and environmentally appropriate.

The taxonomy by Pousman and Stasko (2006) pointed to one pattern, this helped to structure the research and the design of the Auditory Display. The ambient group Auditory Display (see Section I.1.1) can be seen as a **multiple information consolidator** which is concerned with displaying many individual pieces of information about people and their presence in a consolidated manner. It was influenced by the concept of *interstital* information appliances (Wickramasuriya et al., 2007), where it is possible to grab a snippet of information quickly and opportunistically during the interstices between other activities. The existing classifications of these types of systems concentrate on the visual aspects and often neglect or relegate the auditory aspects. This provides opportunities for explorations of ambient information systems using Auditory Displays either as the sole element or as an aspect within a multimodal ambient information system.

The motivation for an ambient group Auditory Display was to provide information about the *presence* and *availability* of co-located colleagues. In particular, we use Huang's and Mynatt's idea of the *Semi-Public Display for Small, Co-located Groups* (Huang and Mynatt,

[3] Their emphasises are shown in bold.

2003) where the information displayed is used to support members of the co-located group within a particular physical space, the space being somewhere not frequented by passerbys. This display can be classified as a *Single Display Groupware* as defined by Stewart et al. (1999) as “*computer programs that enable co-present users to collaborate via a shared computer with a single shared display and simultaneous use of multiple input devices*”.

The ambient group Auditory Display is triggered by sensors and as such the placement of the system was an important factor. The system was positioned inside the door of the laboratory where it was activated by a pressure sensitive floor. This space was chosen on reflection upon the location of the research group, its physical space, its activities, and on comments by Nichols et al. (2002) where doors can “*serve as a medium for communication, where people can broadcast individual messages to passerby’s*” and as “*physical barriers*”. The idea continues the approach of Nichols et al. (2002) whose LabraDoor used a door which was supplemented to function as a mediator and as a medium for communication. The layout of the laboratory includes a small waiting area before widening out into a larger cubicle area, we hoped that the system could provide people entering an *auditory gist* (Zhao et al., 2004) as they enter this waiting area.

This system was aimed at encouraging *awareness moments* and facilitating lightweight interactions or “*opportunistic*” interactions. The display centralised the relevant *presence* and *availability* information about group members from several sources and we hoped this would reduce the effort necessary in gathering such information from various channels such as email or word-of-mouth. This information was taken from several sources including machine presence on the network and instant message activity by the user which were the primary sources. The display was located beside the coffee / dining area used by all the colleagues. It aimed at providing a short auditory gist of who was present and available at that time by using Sound IDs, similar to those in the Hubbub system (Isaacs et al., 2002) but designed with Auditory Icons rather than Earcons.

I.1.1 Technical Details For The Systems

The system was designed for portability, based on the Mac OS X system architecture using Ruby, Python and C and builds upon two existing open source applications, Growl[4] and Boodler[5]. The group Auditory Display is directed toward research group members who are co-located in the same office-space and displays information about the presence and public availability of fellow research group members that it is intended for others to hear. Simple heuristics such as preventing two sounds from having onsets occurring simultaneously (or

[4]http://growl.info/

[5]http://www.eblong.com/zarf/boodler/

nearly simultaneously) can prevent merging of those sounds, while staggering similar event sounds and ignoring multiple alerts based on the same event were used to establish a less cluttered but still informative soundscape.

Architecture and Hardware Details The group display uses an embedded computer, an Arduino[6] shown in Figure I-1 to monitor the sensors fitted inside the floor of the laboratory entrance. The initial sensors used, were pressure sensors or force sensitive resistors (FSR) (Fraden, 2004). Later sensors using capacitive sensing approach, similar in concept to the Theremin (Glinsky, 2000). This replaced the initial FSR approach.

The Arduino was chosen because it is a open source computing platform which in addition to containing the micro-controller board also offers a development environment for writing software for the board. It runs on Windows, Macintosh OS X, and Linux operating systems and can be interfaced with many programming languages including Flash, Processing, PD, Max/MSP, Director, Ruby, and C. This embedded system allows for the group display to be activated when somebody enters or leaves the lab. The group display used a Apple Mac Mini with OS X to run the processes which poll the state of the members in the Interaction Design Centre research group and creates the sounds for display. The group display contains a sound card whose output is activated by the pressure sensors linked to the Mac Mini by the Arduino. This Mac Mini uses rsync (Tridgell, 1999) with a local web-server to facilitate the offering of audio files on-demand so that iTunes, WinAmp, or other music player can request the group display sound files for playback on the particular group member's computer. This allows group members at their desks to determine the presence and availability of other group members.

Software Architecture The design approach used for this system is similarly architecturally to the UNIX philosophy of pipes (Wikipedia, 2007) or pipelines (Project, 2006), as such there are many short, single purpose programs or scripts tasked with one particular function. The components of the system are programmed in Ruby, Python, C, Applescript and there is some shell scripting used. The system was designed to be easily adaptable to change and as such many of the key pipes or programs in the process are code-generated files (Herrington, 2003). This allows for changes to the configuration files to be easily reflected across all components in the pipeline. A high level view of the architecture is shown in Figure I-2. The data flows into the system are from two main sources, network information from local computers or instant messenger status and from the physical sensors. These trigger changes

[6] The Arduino contains a standardized "bootloader", 8 kBytes of Flash program memory, 1 kByte of RAM, runs at 12MHz, has 13 digital input/output pins, and 5 analog input pins. It is based around a ATMEL AVR ATmega8 processor (Catsoluis, 2002) and uses a RISC type architecture.

Fig. I-1: *The Arduino board and its programming interface.*

in the system, updating the relevant user data, calling the playback of the sounds, etc. The outputs of the system are the playback of the sound when triggered by the sensors and the rsync-ing (Tridgell, 1999) of group sound files to a web-server for streamed playback across the local network.

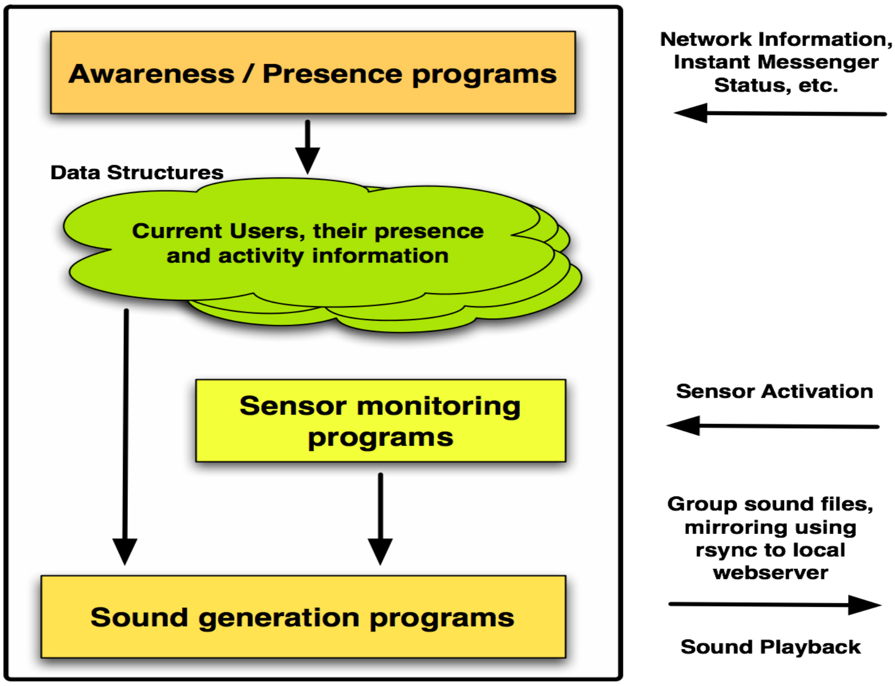


Fig. I-2: *A very abstract and high level outline of the current system.*

A more detailed overview of the system is shown in Figure I-3, which the individual watchers are shown on the left of the diagram and the main group display process in the middle of the diagram. The watchers are logical collections of programs which react upon certain changes such as the triggering of the door sensor (DoorSensorWatcher) or a state change in a user's instant messenger status (IMWatcher). These programs and their source code, as well as scripts used are all available as free software under the GNU General Public License version 3 or greater. This exploration and the system developed was usable and stable for a small group; however testing has not been carried out in larger environments or with large groups. The components break down into approximate 2805 physical source lines of code (SLOC) divided as follows by programming language: Ruby 1571 lines, ANSI C 463 lines, Objective-C 269 lines, Python 243 lines, Shell scripts 139 lines, and Applescript scripts 120 lines. The experimental procedure which was linked into iTunes and facilitating randomisation, triad generation, and the recording of the results was developed using Ruby and contributed another 796 physical source lines of code. The SLOC metric is used to illustrate that the effort in developing an exploratory infrastructure is a non trivial task. As part of this thesis and to assist future research in this area, the source code and scripts developed are and will remain available via the Internet for download and use. This will hopefully ensure that the system, or at least some useful aspects from it will be of benefit to other developers and researchers addressing the issues of group awareness.

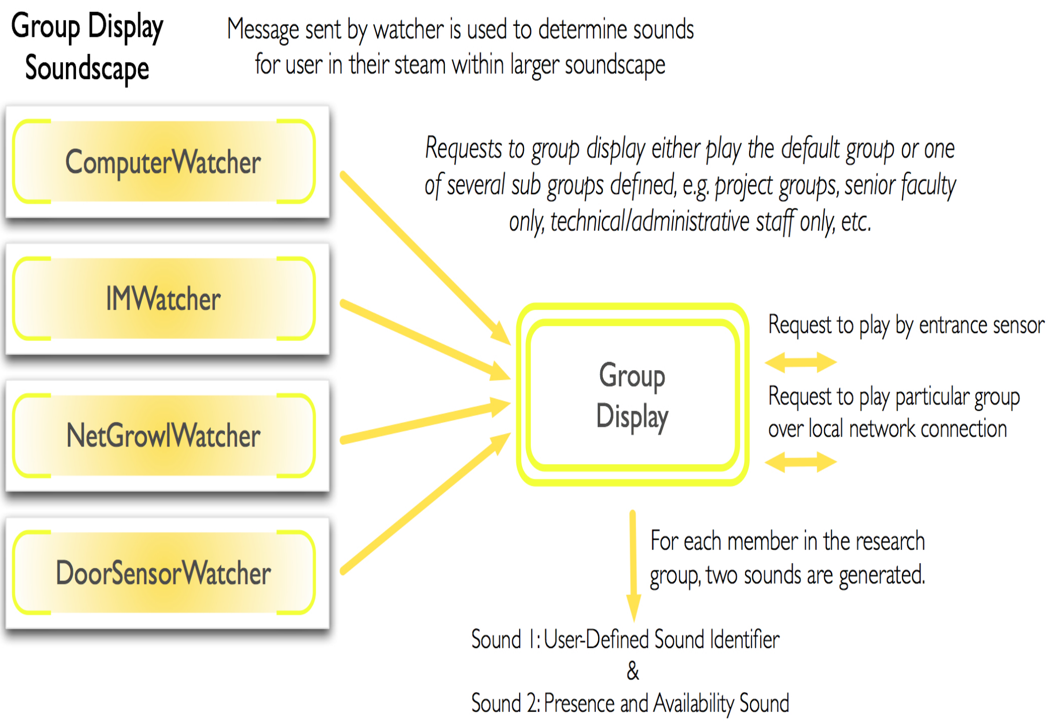


Fig. I-3: *A more detailed high level outline of the current system.*

J.1 Appendix J - Chapter 6 - Second study - repertory grid technique analysis

This appendix holds the grid data, the task list, the figures, and the charts for the second study presented in Chapter 6, where the participants' tacit classifications of sounds were explored using the repertory grid technique (RGT) to help in eliciting what the participant's constructs and their associations were.

Participant ID	Pairs similarity	s1	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s18	s19	s20	s22	s23	s26	s28	Singles description
1	critters	1	1	1	1	2	1*	2	2	3	1*	5	3	2	1	2	4	5*	3	4	unpleasant
1	forest birds	5	1*	3	1*	1	3	4	5	4	5	5	4	5	4	5	5	4	5*	5	everyday noises
1	unnatural sounds	5	5	5	5	5	5	5	5	4	5	1*	2	5	5	5*	1*	1	1	2	animals
1	horse like animals	4	3	5*	3	3	4	1*	1*	2	2	3	3	3	2	1	3	3	3	3	children's pets
1	near shore animals	5*	4	5	4	4	5	5	5	5	5	3	3	1*	1*	5	3	3	3	3	domestic
1	farm life	1	2	4	2	2	1	1	1	1*	1	4	4	1	3	1*	5	4	5	5*	everyday sounds
1	animals	2	1	4	1	1*	1	1	2	2	1*	4	5*	5	4	2	3	3	3	4	sea-faring boat sounds
2	animal sounds	1	1*	1	2	1	1	1	1	1	1	4	5	2	1	1*	1	5	5*	5	sounds from objects
2	animal sounds	1	1	1	1*	1	1	1	1*	1	1	5	5	1	1	1	1	5	5	5*	object sounds
2	land animals	1	1	1	2	1*	1	1	1	1	1*	3	3	4	5*	1	1	3	3	3	water animals
2	object noise	5	4	5	4	5	5	5	5	5*	5	1*	1	4	5	5	5	1*	1	1	animal noise
2	farm animals	5*	2	4	2	2	1*	1	1	1	1	3	3	2	2	1*	2	3	3	3	house pet sounds
2	predominant animals sounds	1	1	1	1	1	1	1	1	1	1	5	5*	1*	1	1	1*	5	5	5	object sounds
2	bird sounds	5	2	1*	2	1	1*	5*	5	5	5	3	3	2	5	5	5	3	3	3	animal sounds
3	time-based night/early	5*	1*	1	1	3	1	1	3	5	3	5	3	1*	3	1	5	5	5	5	anytime - attention grabber
3	nature sounds-waring but not mar	1	3	1*	1*	2	2	1	2	3	2	5*	5	2	1	1	3	5	4	4	manmade-attention grabber
3	more familiar	1	3	4	1	4	1	1	1	5	1	1	2	1	1*	5*	1*	3	1	1	less familiar
3	nature sounds	1	1	2	1	1*	1*	1	1	4	1	5	5	1	1	1	3	5*	5	5	man made
3	non animal sounds	5	5	5	5	5	5	5	5*	4	5	1	1*	5	5	5	5	1	1	1*	animal sounds
3	sheep like sounds	5	5	5	5	5	5	5	5	1*	1*	5	5	5	5	5	5	5	5*	5	non sheep like
3	complete animal sound	1	1	1	1*	1	1	1*	1	5*	1	3	3	1	1	1	1	3	3	3	incomplete animal sounds
4	hard to identify	5	5	4	5	3	5	1	5*	3	5	4	4	5	1*	5	2	3	5	1*	easy to identify
4	annoying	4	5	1*	5*	3	5	5	4	1	4	1*	5	5	1	4	3	4	2	2	relaxing
4	animals alone-single	1*	4	1	4	2	5*	1	1	1	2	3	3	4	1	1*	1	3	3	3	animals in a farm
4	sheep	4	4	3	3	2	2	3	2	1*	1*	4	5*	5	4	2	3	3	4	4	village
4	relaxing open places	3	1*	3	1	2	1	2	2	4	2	4	1	1*	2	5	5*	4	4	4	dark closed spaces
4	man made things	5	4	5	4	5*	5	5	5	5	5	1	1	4	4	5	5	1*	1*	1	animals
4	countryside sounds	2	1	2	1	1	1	5*	1	5	1	2	1*	1	1*	1	4	3	3	3	weird animal sounds
5	welcome greeting recognition	2	5	3	3	3	1*	1	1*	2	1	5	5*	3	2	1	2	5	4	4	night - mechanical
5	comfortable	4	2	2	1*	2	4	2	1	5	2	5*	4	1	2	1*	1	5	4	5	sheep
5	dark dead lonely	4	1*	5	3	3	5	3	4	4	5*	1	2	3	5	5	4	1	2	1*	outdoors life
5	leaving fading	5	2	2	1	4	4	1*	2	5*	5	5	1	1*	2	4	2	2	3	4	cut off
5	concentrating	5	1	3	3	4	5	4	4	5	5	2	1	3	5*	4	4	1*	1*	2	attention seeking
5	relaxing becalmed	5*	2	1*	1	4	5	2	4	5	5	5	4	1	5	4	1*	4	2	4	pleading
5	insistent	1	5	4	4	1*	1	2	2	1	2	5	2	4	2	2	3	5*	4	1	unfeeling detached

Fig. J-1: Participants Construct Pole and Ranking Results (* indicates sound is part of the original triplet played to the participant, which they used to form the constructs).

Task List

Task List for Evaluators - Auditory Icons for User Identification.doc
Evaluators Instructions:
Using the iTunes application, please perform the tasks that are listed below, please do so in the order presented. The observer will be present and be available to answer any questions you may have or to record any comments that you would like to contribute. Please remember this is an evaluation of the sounds, not an evaluation of you! Do not feel afraid to comment on the application or sounds either positively or negatively, as the goal of the experiment is to gather a user's descriptions of the sounds.

Scenario:
You've been asked to listen to a set of sounds and write your own descriptions/labels for the sounds. These sounds are intended for use as user identifier sounds, where a single sound will be used as an auditory/sonic identifier for a particular person. When writing your descriptions or labels, try to keep them short, a word or phrase rather than a long sentence is preferable.

In this experiment, you will be asked to:

- Listen to triplets of sounds and group two of the three together to form one category and group the third sound into a second but distinct category.
- Rank each of the sounds within the categories you will have created whilst listening to the triplets.
- Rank the sounds in order of your favourites, where the chosen favourite sound would be used as your own auditory/sonic identifier.

Number of tasks to complete: 16 (sixteen)

Task List:

START OF PILOT STAGE

1. In iTunes, Select the folder "Experiment Pilot" and then select & play the sounds in the playlist "Pilot Comparisons: 1".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

2. In iTunes, Select the folder "Experiment Pilot" and then select the playlist "Alt Pilot Comparisons: 1".

Listen to the sounds presented and rank them within the two categories you created in the previous task.

END OF PILOT STAGE

(continued on next page)

START OF EXPERIMENT - PART A

1. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 1".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

2. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 2".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

3. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 3".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

4. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 4".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

(continued on next page)

5. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 5".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

6. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 6".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

7. In iTunes, Select the folder "Experiment Part A" and then select & play the sounds in the playlist "Comparisons: 7".

Listen to the sounds presented, write your description for the two sounds you feel most belong together and then write your description for the sound, which you feel, is different from the other pair of sounds.

END OF EXPERIMENT - PART A

If you wish to take a break, please feel free to do so.

START OF EXPERIMENT - PART B

1. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 1".

Listen to the sounds presented and rank them within the two categories you created in the task 3, which dealt with "Experiment Part A" and the playlist "Comparisons: 1".

(continued on next page)

2. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 2".

Listen to the sounds presented and rank them within the two categories you created in the task 4, which dealt with "Experiment Part A" and the playlist "Comparisons: 2".

3. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 3".

Listen to the sounds presented and rank them within the two categories you created in the task 5, which dealt with "Experiment Part A" and the playlist "Comparisons: 3".

4. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 4".

Listen to the sounds presented and rank them within the two categories you created in the task 6, which dealt with "Experiment Part A" and the playlist "Comparisons: 4".

5. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 5".

Listen to the sounds presented and rank them within the two categories you created in the task 7, which dealt with "Experiment Part A" and the playlist "Comparisons: 5".

6. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 6".

Listen to the sounds presented and rank them within the two categories you created in the task 8, which dealt with "Experiment Part A" and the playlist "Comparisons: 7".

(continued on next page)

2. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 2".

Listen to the sounds presented and rank them within the two categories you created in the task 4, which dealt with "Experiment Part A" and the playlist "Comparisons: 2".

3. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 3".

Listen to the sounds presented and rank them within the two categories you created in the task 5, which dealt with "Experiment Part A" and the playlist "Comparisons: 3".

4. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 4".

Listen to the sounds presented and rank them within the two categories you created in the task 6, which dealt with "Experiment Part A" and the playlist "Comparisons: 4".

5. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 5".

Listen to the sounds presented and rank them within the two categories you created in the task 7, which dealt with "Experiment Part A" and the playlist "Comparisons: 5".

6. In iTunes, Select the folder "Experiment Part B" and then select the playlist "Alt Comparisons: 6".

Listen to the sounds presented and rank them within the two categories you created in the task 8, which dealt with "Experiment Part A" and the playlist "Comparisons: 7".

(continued on next page)

Participant 1 - PCA - Constructs

Examining the correlations among the constructs for Participant 1, shown in Table J-1, we can say that *critters* or *small animals* are associated with *farm life* (construct 6, 0.62) and they are definitely not *unnatural sounds* (construct 3, -0.89). Each construct can be considered in a similar manner to the first construct with regard to the other constructs, rather than consider each of the constructs individually we will present the significant correlations as these are the items we can make the strongest statements about, based on the participant's data. There was a significant correlation (0.89) between "*critters-unpleasant*" and "*unnatural sounds-animals*". There was also a significant correlation (0.83) between "*farm life-everyday sounds*" and "*unnatural sounds-animals*". These correlations can help in giving an insight into the participants world view of the sounds presented. In order to reduce the amount of information present in Table J-1, we can use the root mean square correlation among constructs as shown in Table J-2. The "*unnatural sounds – animals*" construct is the construct most closely associated with the other constructs as shown in Table J-2. "*Farm life – everyday sounds*" is the next most closely associated construct.

	1	**2**	**3**	**4**	**5**	**6**	**7**
1 critters — unpleasant	1.00	0.41	**-0.89**	-0.05	-0.33	0.62	0.44
2 forest birds — everyday noises		1.00	-0.38	-0.28	-0.16	0.14	0.46
3 unnatural sounds — animals			1.00	-0.15	0.39	**-0.83**	-0.48
4 horse like animals — childrens pets				1.00	-0.08	0.38	0.26
5 near shore animals — domestic					1.00	-0.47	-0.70
6 farm life — everyday sounds						1.00	0.59
7 animals — sea-faring boat sounds							1.00

Table J-1: Correlation Analysis for Participant 1.

The interpretation of the first principal component in Table J-3 measures the overall 'naturalness' of the sources of the sounds. The second principal component in Table J-3 contrasts animal sounds with everyday sounds, implying that after the 'naturalness' of the sources of the sounds is taken into account, the main source of variation is between sounds from animals and those from everyday sources. The first two components are shown in Figure J-2, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown

Construct	Root-mean-square correlation
critters — unpleasant	0.62
forest birds — everyday noises	0.48
unnatural sounds — animals	0.66
horse like animals — childrens pets	0.43
near shore animals — domestic	0.54
farm life — everyday sounds	0.63
animals — sea-faring boat sounds	0.60
Average of statistic	0.57
Standard deviation of statistic	0.08

Table J-2*:* Root-mean-square (average) correlation among constructs analysis for Participant 1.

in Figure J-3.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
critters - unpleasant	-0.44	+	0.22	
forest birds - everyday noises	-0.24	-	0.59	+
unnatural sounds - animals	0.48	+	-0.05	
horse like animals - childrens pets	-0.08	(-)	-0.72	-
near shore animals - domestic	0.38	+	0.13	
farm life - everyday sounds	-0.45	-	-0.27	(-)
animals/sea - faring boat sounds	-0.41	-	-0.07	
Standard deviation	2.22		1.20	
Proportion of Variance	0.70		0.21	
Cumulative Proportion	0.70		0.91	

Table J-3*:* Principal–Components Analysis for Participant 1.

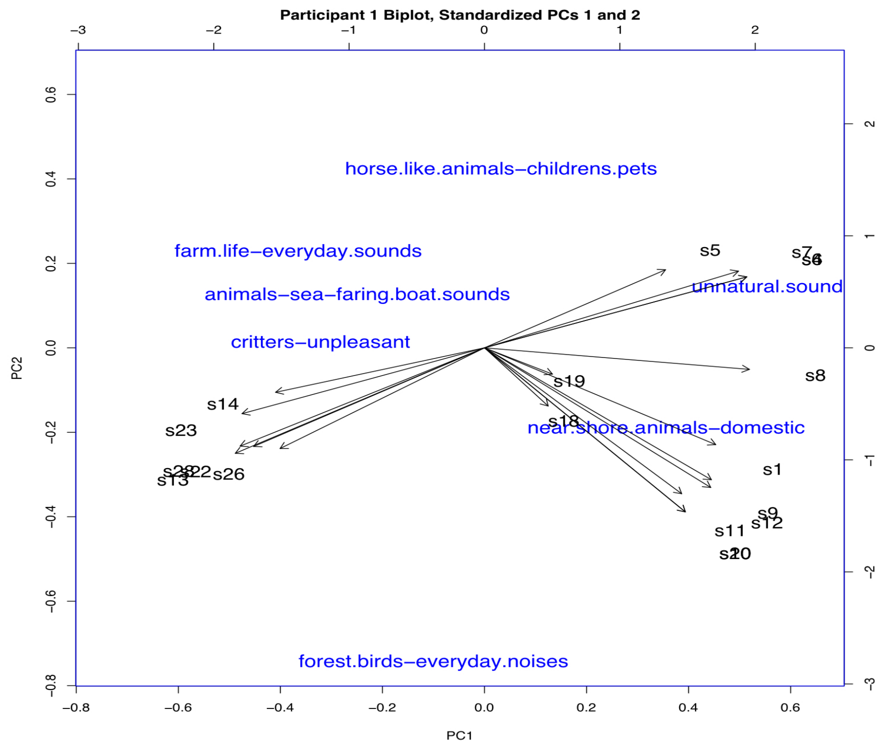


Fig. J-2*: The first two components of the principal–components analysis for Participant 1.*

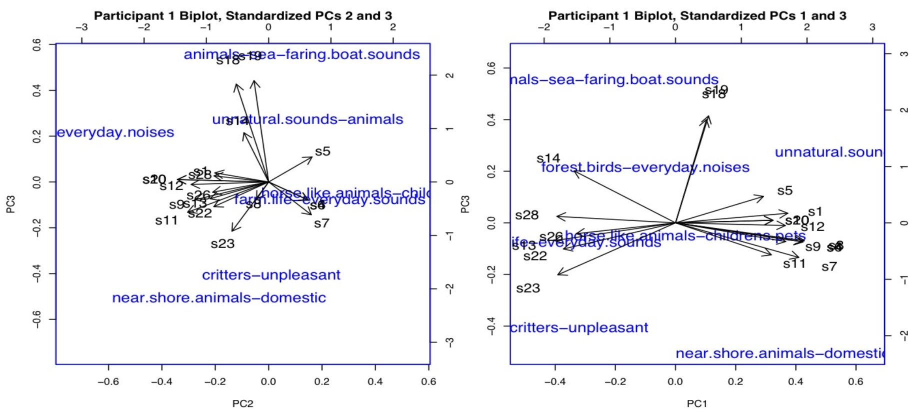


Fig. J-3*: The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 1.*

Participant 1 - PCA - Elements

The results of the elicitation task for Participant 1 are shown below in Table J-4. Participants responded in free-text format to what they thought each sound was as we can see from Table J-4, these text descriptions were often highly descriptive and described the events or actions.

ID	Sound Description	Participant's Description
s1	cat	*cat*
s4	owl	*cuckoo bird*
s5	bird song	*budgerigar (budgie)*
s6	bird song	*unknown bird*
s7	bird song	*forest bird*
s8	rooster	*rooster*
s9	donkey	*donkey*
s10	horse	*horse*
s11	goat	*lamb being slaughtered*
s12	sheep	*sheep*
s13	glass breaking	*glass breaking*
s14	church bell ringing	*a ship's fog bell*
s18	seagull	*seagulls*
s19	seal	*seal*
s20	horse	*horse*
s22	lion roaring	*motorcycle*
s23	power saw	*dentist's drill*
s26	coins counting	*money handling*
s28	heavy ball bouncing	*broom falling on wooden floor*

Table J-4*:* Descriptions by Participant 1 for the elements.

Examining the correlations among the elements for Participant 1, shown in Table J-5, we can say that element *s4* (the sound of an owl hooting) was definitely associated with *s6 (1.00)*, *s7 (0.97)*, and *s8 (0.85)* (two different types of bird song and the sound of a rooster crowing, respectively) while being definitely not associated with *s13 (-0.94)*, *s23 (-0.81)*, and *s28 (-0.85)* (the sound of glass breaking, the sound of a power saw in use, and the sound of a heavy ball bouncing, respectively). As previously stated, we will only consider the one element or sound and the elements with the most significant correlations to this element, in order to highlight the points of interest from the participant's data. The *s8* element (the sound of a rooster crowing) is the element most closely associated with the other elements as shown in Table J-6. The elements *s4* (the sound of an owl hooting) and *s6* (bird song version 2) are the next most closely associated elements.

	s1	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s18	s19	s20	s22	s23	s26	s28
s1	1.00	0.61	0.56	0.61	0.52	**0.91**	0.77	0.78	0.75	**0.92**	−0.51	−0.44	0.36	0.24	0.78	−0.43	−0.60	−0.29	−0.51
s4		1.00	0.73	**1.00**	**0.97**	**0.85**	0.59	0.43	0.44	0.58	**−0.94**	−0.77	−0.08	0.09	0.43	−0.76	**−0.81**	−0.66	**−0.85**
s5			1.00	0.73	0.57	0.66	0.22	0.18	0.08	0.38	−0.77	−0.18	0.07	0.28	0.18	−0.54	−0.79	−0.31	−0.54
s6				1.00	**0.97**	**0.85**	0.59	0.43	0.44	0.58	**−0.94**	−0.77	−0.08	0.09	0.43	−0.76	**−0.81**	−0.66	**−0.85**
s7					1.00	**0.80**	0.59	0.40	0.47	0.52	**−0.89**	**−0.87**	−0.15	−0.03	0.40	−0.77	−0.72	−0.73	**−0.88**
s8						1.00	0.74	0.65	0.70	**0.83**	−0.73	−0.69	0.10	0.05	0.65	−0.61	−0.69	−0.50	−0.75
s9							1.00	**0.97**	**0.94**	**0.94**	−0.43	−0.56	0.16	0.15	**0.97**	−0.39	−0.45	−0.35	−0.44
s10								1.00	**0.92**	**0.94**	−0.31	−0.36	0.33	0.28	1.00	−0.31	−0.42	−0.25	−0.31
s11									1.00	**0.87**	−0.25	−0.48	0.13	−0.07	**0.92**	−0.34	−0.28	−0.34	−0.42
s12										1.00	−0.44	−0.47	0.27	0.25	**0.94**	−0.33	−0.50	−0.22	−0.39
s13											1.00	0.65	−0.16	−0.34	−0.31	**0.88**	**0.94**	0.77	**0.87**
s14												1.00	0.14	0.16	−0.36	0.58	0.43	0.64	0.75
s18													1.00	**0.80**	0.33	−0.35	−0.45	−0.28	−0.14
s19														1.00	0.28	−0.27	−0.57	−0.13	−0.03
s20															1.00	−0.31	−0.42	−0.25	−0.31
s22																1.00	**0.87**	**0.96**	**0.95**
s23																	1.00	0.72	0.79
s26																		1.00	**0.93**
s28																			1.00

Table J-5: Correlation Analysis for Participant 1 Elements.

Element	Root-mean-square correlation
s1	0.61
s4	0.66
s5	0.46
s6	0.66
s7	0.64
s8	0.67
s9	0.59
s10	0.54
s11	0.52
s12	0.60
s13	0.64
s14	0.54
s18	0.28
s19	0.27
s20	0.54
s22	0.60
s23	0.65
s26	0.53
s28	0.62
Average of statistic	0.56
Standard deviation of statistic	0.12

Table J-6*:* Root-mean-square (average) correlation among elements analysis for Participant 1.

The interpretation of the first principal component in Table J-7 contrasts the animal/bird sounds with those where were object based. The second principal component in Table J-7 weakly supports the first principal component but given the lack of strong contrasts in this principal component it is difficult to determine the next main source of variation. The first two components are shown in Figure J-4, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure J-5.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
s1	−0.25	-	0.02	
s4	−0.25	-	−0.11	
s5	−0.22	-	0.02	
s6	−0.25	-	−0.11	
s7	−0.24	-	−0.15	
s8	−0.25	-	−0.09	
s9	−0.24	-	−0.06	
s10	−0.23	-	0.03	
s11	−0.22	-	−0.11	
s12	−0.24	-	−0.01	
s13	0.24	+	−0.00	
s14	0.24	+	0.22	(+)
s18	−0.11	(-)	0.66	+
s19	−0.12	(-)	0.65	+
s20	−0.23	-	0.03	
s22	0.24	+	−0.05	
s23	0.24	+	−0.12	
s26	0.24	+	−0.01	
s28	0.25	+	0.06	
Standard deviation	3.96		1.33	
Proportion of Variance	0.83		0.09	
Cumulative Proportion	0.83		0.92	

Table J-7*:* Principal–Components Analysis for Participant 1 Elements.

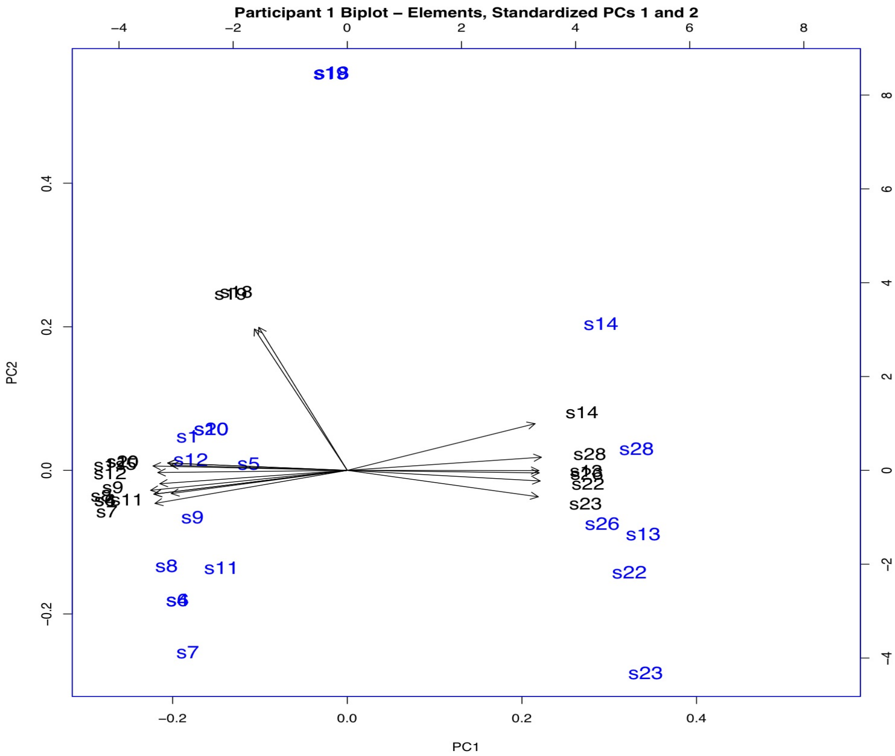


Fig. J-4: *The first two components of the principal–components analysis for Participant 1 Elements.*

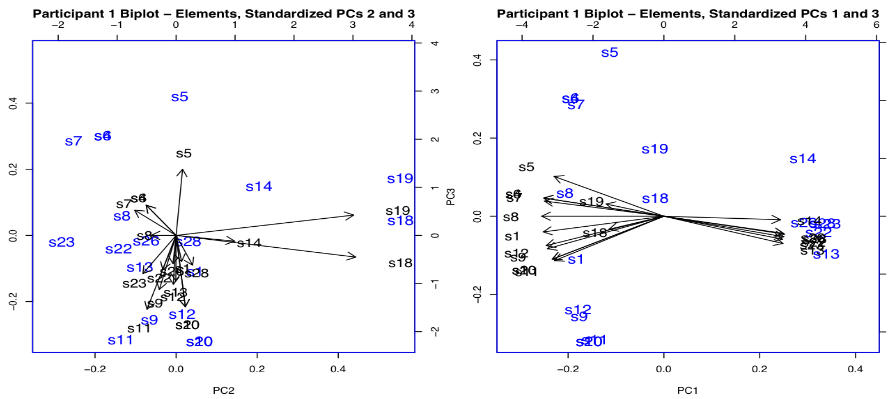


Fig. J-5: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 1 Elements.*

Participant 1 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 1 are shown in Figure J-6. There are two possible cutoff levels, at 7 and 8.15 and would result in 3 and 4 clusters respectively. The resulting clusters for cutoff level 8.15 were shown in Table J-8. Cluster 1 contained sounds that were seen as either animal sounds or mechanical / unpleasant. The shortest distance match in this cluster was at 5.65 between ”*critters – unpleasant*” and ”*farm life – everyday sounds*”. The next sub-cluster joined at a distance of 5.91 and was ”*animals/sea – faring boat sounds*”. The last sub-cluster joined at a distance of 6.55 and was ”*horse like animals – childrens pets*”. These constructs suggest that certain animal sounds are seen as more pleasant and natural sounds. These sounds and the derived constructs were construed in terms of their subject matter and the relevance to location. In some cases, while the constructs may appear to be similar, if they do not distinguish between elements in the same manner and from a similar or the same context then they are unlikely to be referring to the same concept. This is supported by the MST MDS analysis and by the principal component analysis for the participant’s constructs as shown previously in Table J-3.

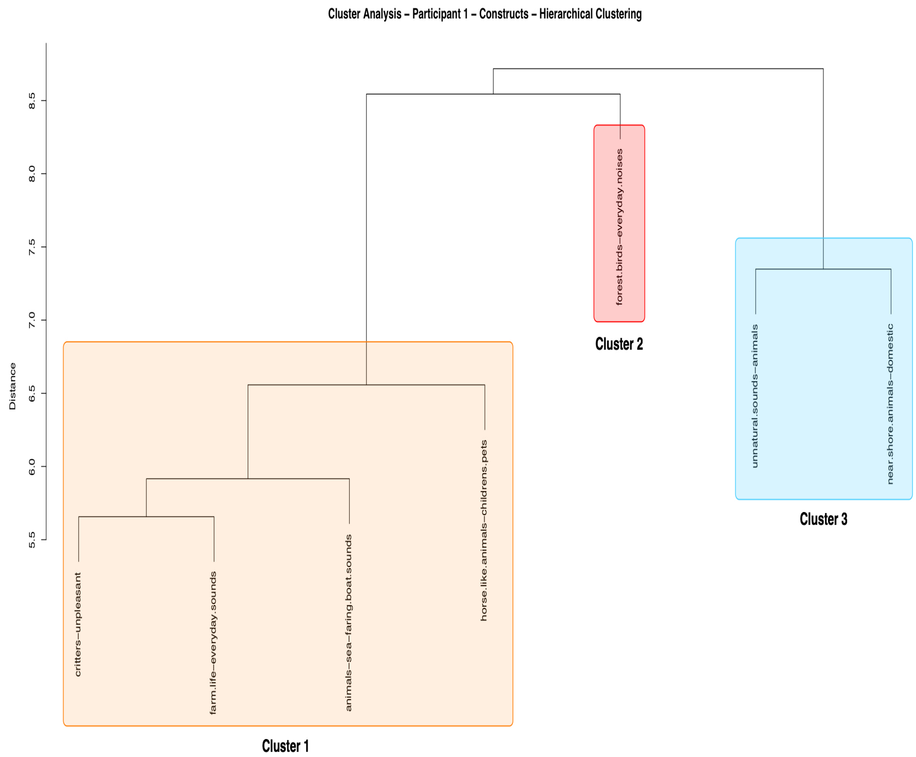


Fig. J-6: *The cluster analysis of the RGT constructs for Participant 1.*

A minimum spanning tree is a used to highlight possible distortions produced by the scaling solutions. These distortions are indicated by nearby points on the MDS plot not being linked by an edge of the tree. The MST multidimensional scaling analysis for Participant 1 is shown in Figure J-7, with its related Shepard diagram in Figure J-8. Examining Figure J-7 visually for patterns, we see that it is mostly likely a circumplex pattern as “*forest birds – everyday noises*” could not be interpreted to lie under the same line as the concepts of ”*critters – unpleasant*” or ”*farm life – everyday sounds*”. This indicates that there are two or

	Constructs *Emergent Pole* - *Implicit Pole*
1	critters — unpleasant farm life — everyday sounds animals/sea — faring boat sounds horse like animals — childrens pets
2	forest birds — everyday noises
3	unnatural sounds — animals near shore animals — domestic

Table J-8*:* Clusters obtained for Participant 1.

more dimensions effecting the participant's determination of constructs. A Shepard diagram is used to highlight discrepancies between the original dissimilarities (shown in diagram as X's) and the multidimensional scaling solution (the line in the diagram). The quality of the current multidimensional scaling solution shows that it is close but not quite an ideal solution as a number of the points do not fall on the bisecting line.

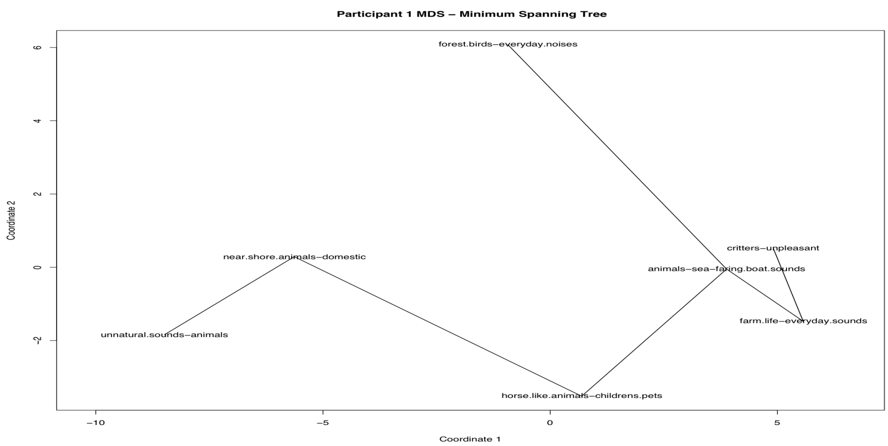


Fig. J-7*: MDS minimum spanning tree analysis of the RGT constructs for Participant 1.*

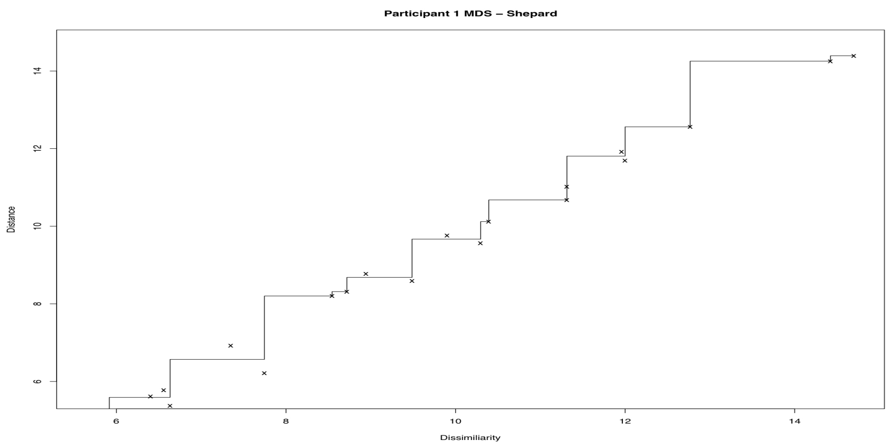


Fig. J-8: *Shepard diagram of the MDS analysis of the RGT constructs for Participant 1.*

Participant 1 - Cluster Analysis and MDS - Elements

The cluster analysis for Participant 1 was shown already in Figure J-9 with a cutoff level of approximately 0.8 which results in 4 clusters. These were objects, seaside birds & animals, farm animals, and birds. These were clustered into two larger clusters which were either object sounds or bird & animal sounds. Cluster 1 contained sounds that were seen as object like sounds. One noted exception to this cluster is s22, the sound of a lion roaring, which was found to have identification issues. This is discussed further in Section 5.1.1. Cluster 2 contains sounds that were seen as bird or animal like sounds. In some cases, while the elements or sounds may appear to be similar, if the participant has not distinguished between elements in the same manner and from a similar or the same context then it means that they are unlikely to be referring to the same concept.

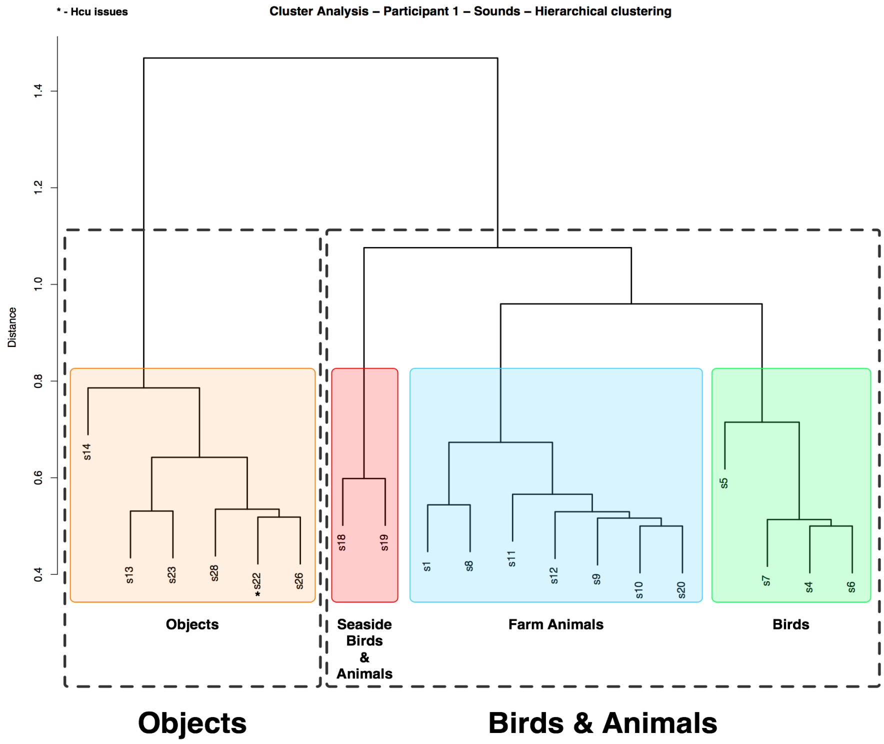


Fig. J-9: *The cluster analysis of the RGT elements (sounds) for Participant 1.*

The MST multidimensional scaling analysis for Participant 1 is shown in Figure J-10, with its related Shepard diagram in Figure J-11. Examining Figure J-10 visually for patterns, we see that it is mostly likely a circumplex pattern as *s7* could not be interpreted to lie under the same line as the elements *s5* or *s11*. This indicates that there are two or more dimensions effecting the participant's determination of elements. This somewhat supports the earlier findings of the principal component analysis for the participant's elements as shown in Table J-7.

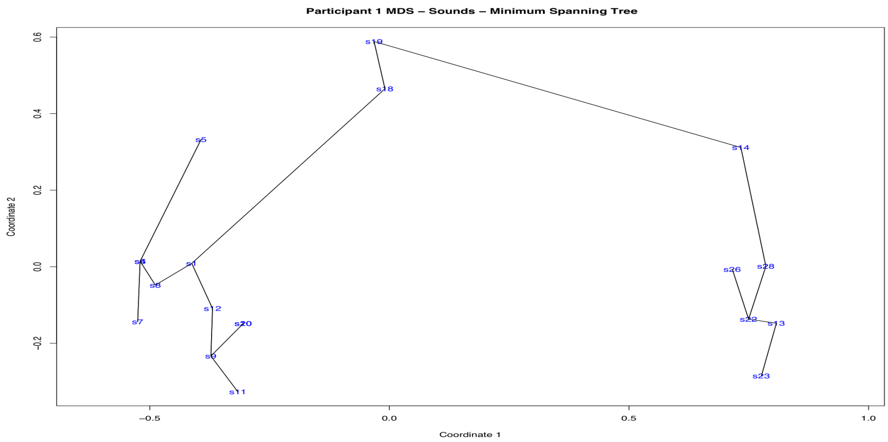


Fig. J-10: *MDS minimum spanning tree analysis of the RGT elements (sounds) for Participant 1.*

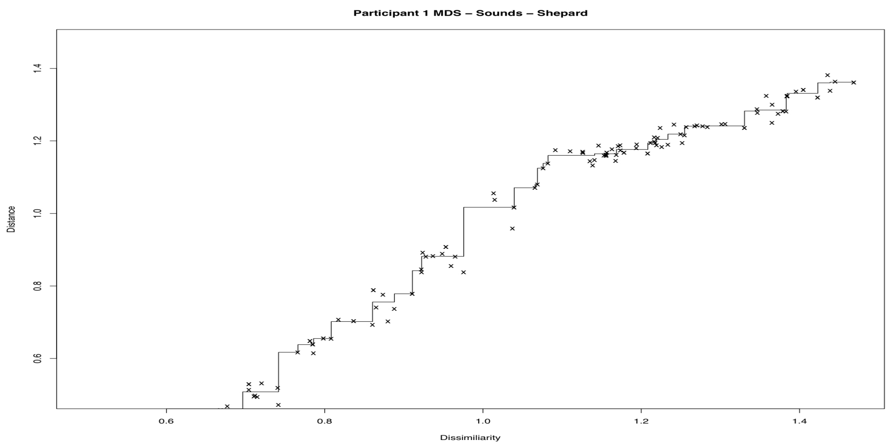


Fig. J-11: *Shepard diagram of the MDS analysis of the RGT elements (sounds) for Participant 1.*

J.1.1 Results and Observations for Participant 1

The results of the analysis of the constructs from Participant 1 from the PCA, CA, and MDS analysis seem to support that a certain animal and bird sounds were heard as more pleasant and natural than the other object sounds.The next factor was between the pleasantness of natural sounds and the unpleasantness of mechanical sounds and how this interaction fitted the subject matter with relevance to location. The results of the analysis of the elements (sounds) from Participant 1 from the PCA, CA, and MDS analysis suggested that the contrast of animal/bird sounds with those where were object based was a factor.

Participant 2 - PCA - Constructs

Examining the correlations among the constructs for Participant 2, shown in Table J-9, we can say that *animal sounds* (construct 1) is associated with *animal sounds* and *predominant* *animals sounds*, which were constructs 2 (0.98) and 6 (0.98) respectively. The participant viewed that the *animal sounds* (the first construct) were definitely not *object noise* (construct 4, -0.98). As previously stated, we will only consider the first construct and the constructs with the most significant correlations, in order to highlight the points of interest from the participant's data. The correlation between "*animal sounds - object sounds*" and "*predominant animals sounds - object sounds*" was highly significant (1.00). There was a significant correlation (0.98) between "*animal sounds - object sounds*" and "*object noise - animal noise*". There was also a significant correlation (0.98) between "*object noise - animal noise*" and "*predominant animals sounds - object sounds*". It is important to note that while the same labels may be used by participants, that their context may differ and as such they cannot be directly compared as being the same concept. The "*object noise – animal noise*" construct is the construct most closely associated with the other constructs as shown in Table J-10. "*animal sounds – sounds from objects*", "*animal sounds – object sounds*" and "*predominant animals sounds – object sounds*" were the next most closely associated constructs as they all correlated with the other constructs at 0.79.

	1	**2**	**3**	**4**	**5**	**6**	**7**
1 animal sounds — sounds from objects	1.00	**0.98**	0.57	**-0.98**	0.42	**0.98**	-0.20
2 animal sounds — object sounds		1.00	0.51	**-0.98**	0.43	**1.00**	-0.14
3 land animals — water animals			1.00	-0.56	0.24	0.51	-0.07
4 object noise — animal noise				1.00	-0.42	**-0.98**	0.23
5 farm animals — house pet sounds					1.00	0.43	-0.23
6 predominant animals sounds — object sounds						1.00	-0.14
7 bird sounds — animal sounds							1.00

Table J-9*:* Correlation Analysis for Participant 2.

The interpretation of the first principal component in Table J-11 measures the overall 'naturalness' of the sources of the sounds. The second principal component in Table J-11 contrasts the type of animal sounds between those of birds and of other kinds of animals, implying that after the 'naturalness' of the sources of the sounds is taken into account, the main source of variation is between sounds from the type of animal producing the sounds. The

Construct	Root-mean-square correlation
animal sounds — sounds from objects	0.79
animal sounds — object sounds	0.79
land animals — water animals	0.56
object noise — animal noise	0.80
farm animals — house pet sounds	0.51
predominant animals sounds — object sounds	0.79
bird sounds — animal sounds	0.41
Average of statistic	0.67
Standard deviation of statistic	0.16

Table J-10*:* Root-mean-square (average) correlation among constructs analysis for Participant 2.

first two components are shown in Figure J-12, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure J-13.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
animal sounds - sounds from objects	-0.46	-	-0.12	(-)
animal sounds - object sounds	-0.45	-	-0.15	(-)
land animals - water animals	-0.27	-	-0.15	(-)
object noise - animal noise	0.47	+	0.10	(+)
farm animals - house pet sounds	-0.23	-	0.62	+
predominant animals sounds - object sounds	-0.45	-	-0.15	(-)
bird sounds - animal sounds	0.18	+	-0.73	-
Standard deviation	2.41		0.84	
Proportion of Variance	0.83		0.10	
Cumulative Proportion	0.83		0.93	

Table J-11*:* Principal–Components Analysis for Participant 2.

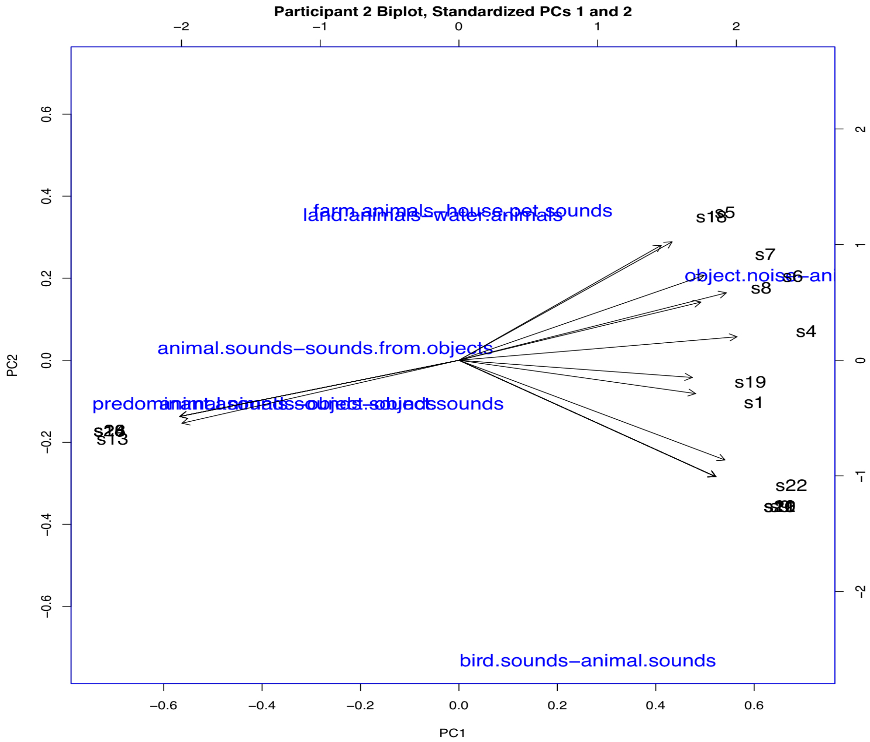


Fig. J-12: *The first two components of the principal–components analysis for Participant 2.*

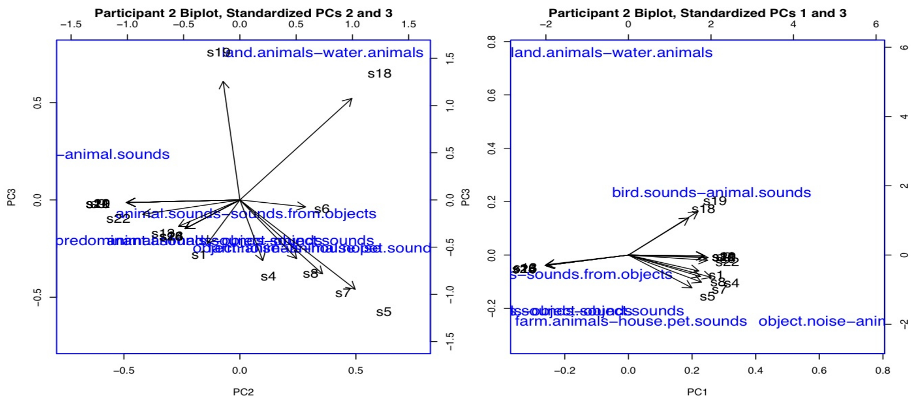


Fig. J-13: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 2.*

Participant 2 - PCA - Elements

The results of the elicitation task for Participant 2 are shown below in Table J-12. Participants responded in free-text format to what they thought each sound was as we can see from Table J-12, these text descriptions were often highly descriptive and described the events or actions.

ID	Sound Description	Participant's Description
s1	cat	*cat (short)*
s4	owl	*night sounds (owl)*
s5	bird song	*bird chirping*
s6	bird song	*jungle sounds*
s7	bird song	*bird chirping*
s8	rooster	*cockerel*
s9	donkey	*donkey*
s10	horse	*horse braying*
s11	goat	*sheep bleating (short)*
s12	sheep	*sheep bleating*
s13	glass breaking	*glass breaking*
s14	church bell ringing	*bell*
s18	seagull	*seagull at seaside*
s19	seal	*seal*
s20	horse	*horse braying*
s22	lion roaring	*lion roaring*
s23	power saw	*drill noise*
s26	coins counting	*coin being counted*
s28	heavy ball bouncing	*marble ball dropped onto table*

Table J-12*:* Descriptions by Participant 2 for the elements.

Examining the correlations among the constructs for Participant 2, shown in Table J-13, we can say that element *s13* (the sound of glass breaking) was definitely associated with *s14 (0.97)*, *s23 (0.97)*, *s26 (0.97)*, and *s28 (0.97)* (the sound of church bells ringing, the sound of a power saw in use, the sound of coins being counted, and the sound of a heavy ball bouncing, respectively) while being definitely not associated with *s4 (-0.88)*, *s6 (-0.88)*, *s7 (-0.80)*, *s18 (-0.84)*, and *s19 (-0.80)* (the sound of an owl hooting, two different versions of bird song, the sound of a seagull's cry, and the sound of a seal roaring, respectively). The elements *s4*, *s14*, *s23*, *s26*, and *s28* (the sound of an owl hooting, the sound of church bells ringing, the sound of a power saw in use, the sound of coins being counted, and the sound of a heavy ball bouncing, respectively) where the elements or sounds most closely associated with the other elements as shown in Table J-14.

	s1	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s18	s19	s20	s22	s23	s26	s28
s1	1.00	**0.80**	0.72	0.62	0.60	0.47	0.73	0.73	0.73	0.73	−0.73	−0.77	0.28	0.53	0.73	**0.85**	−0.77	−0.77	−0.77
s4		1.00	**0.86**	**0.90**	**0.94**	**0.91**	0.79	0.79	0.79	0.79	**−0.87**	**−0.88**	0.55	0.57	0.79	**0.84**	**−0.88**	**−0.88**	**−0.88**
s5			1.00	0.77	**0.90**	0.76	0.39	0.39	0.39	0.39	−0.76	−0.76	0.46	0.28	0.39	0.51	−0.76	−0.76	−0.76
s6				1.00	**0.89**	**0.88**	0.68	0.68	0.68	0.68	**−0.95**	**−0.88**	**0.80**	0.66	0.68	0.71	**−0.88**	**−0.88**	**−0.88**
s7					1.00	**0.97**	0.59	0.59	0.59	0.59	**−0.81**	**−0.80**	0.58	0.42	0.59	0.62	**−0.80**	**−0.80**	**−0.80**
s8						1.00	0.65	0.65	0.65	0.65	−0.77	−0.75	0.60	0.46	0.65	0.63	−0.75	−0.75	−0.75
s9							1.00	**1.00**	**1.00**	**1.00**	−0.70	−0.71	0.39	0.72	1.00	**0.98**	−0.71	−0.71	−0.71
s10								1.00	**1.00**	**1.00**	−0.70	−0.71	0.39	0.72	**1.00**	**0.98**	−0.71	−0.71	−0.71
s11									1.00	**1.00**	−0.70	−0.71	0.39	0.72	1.00	**0.98**	−0.71	−0.71	−0.71
s12										1.00	−0.70	−0.71	0.39	0.72	**1.00**	**0.98**	−0.71	−0.71	−0.71
s13											1.00	**0.97**	**−0.84**	**−0.80**	−0.70	−0.75	**0.97**	**0.97**	**0.97**
s14												1.00	**−0.80**	**−0.84**	−0.71	−0.77	**1.00**	**1.00**	**1.00**
s18													1.00	**0.80**	0.39	0.38	**−0.80**	**−0.80**	**−0.80**
s19														1.00	0.72	0.71	**−0.84**	**−0.84**	**−0.84**
s20															1.00	**0.98**	−0.71	−0.71	−0.71
s22																1.00	−0.77	−0.77	−0.77
s23																	1.00	**1.00**	**1.00**
s26																		1.00	**1.00**
s28																			1.00

Table J-13: Correlation Analysis for Participant 2 Elements.

Element	Root-mean-square correlation
s1	0.70
s4	0.83
s5	0.63
s6	0.79
s7	0.73
s8	0.72
s9	0.76
s10	0.76
s11	0.76
s12	0.76
s13	0.82
s14	0.83
s18	0.60
s19	0.69
s20	0.76
s22	0.79
s23	0.83
s26	0.83
s28	0.83
Average of statistic	0.76
Standard deviation of statistic	0.07

Table J-14*:* Root-mean-square (average) correlation among elements analysis for Participant 2.

The interpretation of the first principal component in Table J-15 contrasts the animal/bird sounds with those where were object based. The second principal component in Table J-15 contrasts the type of animal sounds between those of birds and of other kinds of animals, implying that after the 'naturalness' of the sources of the sounds is taken into account, the main source of variation is between sounds from the type of animal producing the sounds. The first two components are shown in Figure J-14, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure J-15.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
s1	−0.23	-	0.10	(+)
s4	−0.23	-	−0.08	
s5	−0.22	-	−0.38	-
s6	−0.23	-	−0.18	(-)
s7	−0.23	-	−0.27	-
s8	−0.23	-	−0.21	-
s9	−0.23	-	0.30	+
s10	−0.23	-	0.30	+
s11	−0.23	-	0.30	+
s12	−0.23	-	0.30	+
s13	0.23	+	0.13	(+)
s14	0.23	+	0.11	(+)
s18	−0.22	-	−0.29	-
s19	−0.23	-	0.11	(+)
s20	−0.23	-	0.30	+
s22	−0.23	-	0.25	+
s23	0.23	+	0.11	(+)
s26	0.23	+	0.11	(+)
s28	0.23	+	0.11	(+)
Standard deviation	4.27		0.73	
Proportion of Variance	0.96		0.03	
Cumulative Proportion	0.96		0.99	

Table J-15*:* Principal–Components Analysis for Participant 2 Elements.

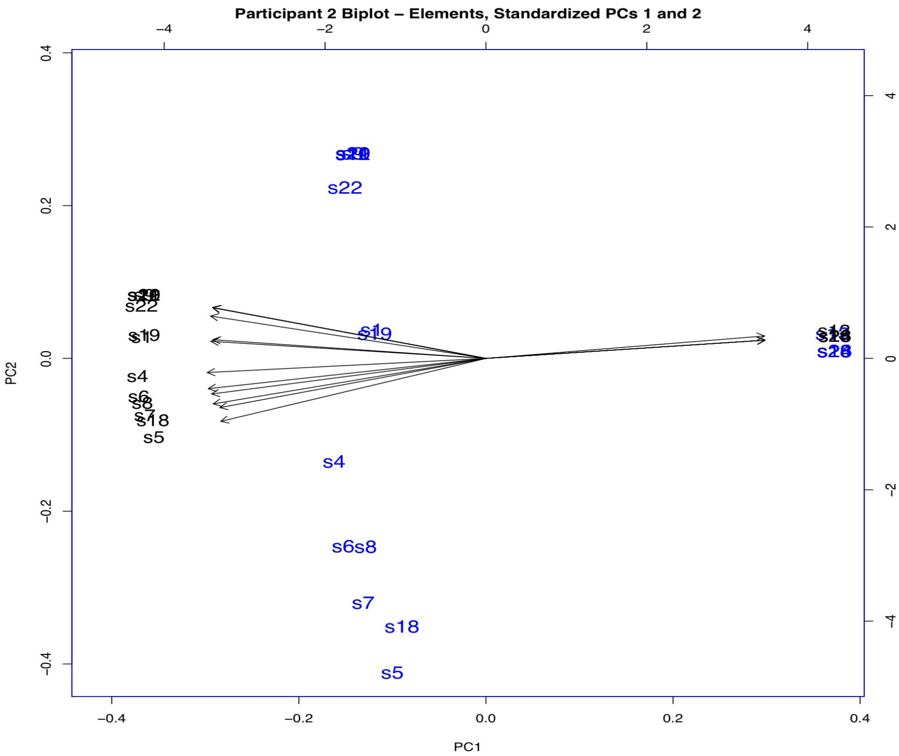


Fig. J-14: *The first two components of the principal–components analysis for Participant 2 Elements.*

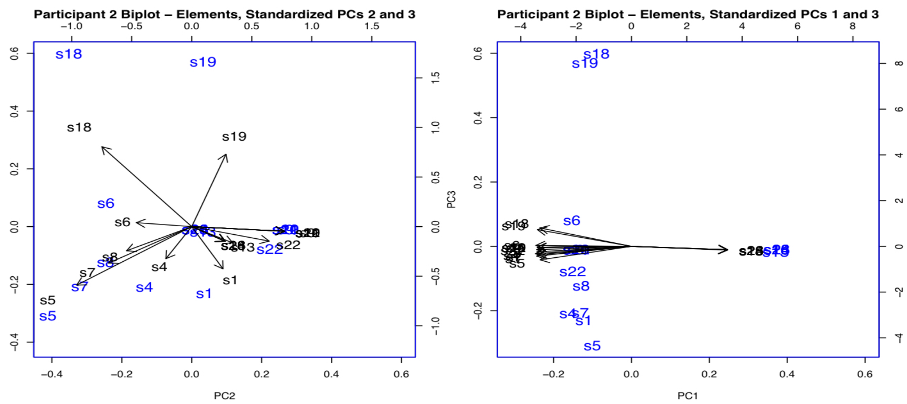


Fig. J-15: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 2 Elements.*

Participant 2 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 2 were shown in Figure J-16. The two possible cutoff levels, at 5.6 and 8.4 and would result in 3 and 5 clusters respectively. The resulting clusters for cutoff level 8.4 were shown in Table J-16. Cluster 1 contained sounds that were seen as either animal sounds or sounds from objects The shortest distance match in this cluster was 0 (zero) between "*animal sounds – object sounds*" and "*predominant animals sounds – object sounds*". The next sub-cluster joined at a distance of 1.73 and was "*animal sounds – sounds from objects*". The next sub-cluster joined at a distance of 6 and was "*land animals – water animals*". The last sub-cluster joined at a distance of 6.4 and was "*farm animals – house pet sounds*". These constructs suggest that the sounds were distinctly divided into animal or object sources.

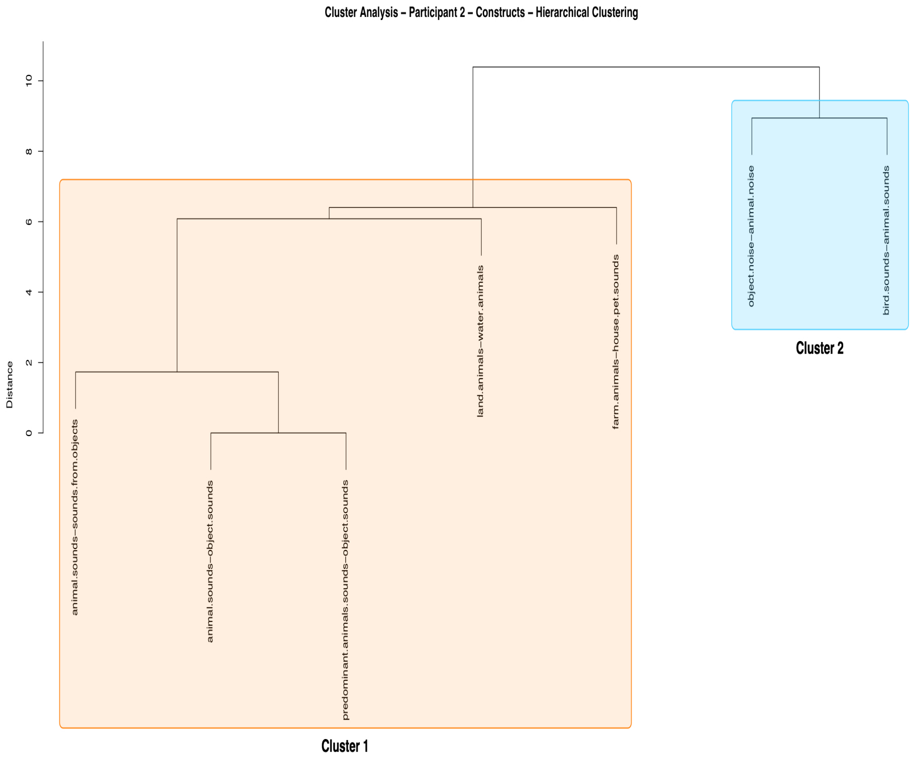


Fig. J-16*: The cluster analysis of the RGT constructs for Participant 2.*

	Constructs *Emergent Pole - Implicit Pole*
1	animal sounds — sounds from objects
	animal sounds — object sounds
	predominant animals sounds —object sounds
	land animals — water animals
	farm animals — house pet sounds
2	object noise — animal noise
3	bird sounds — animal sounds

Table J-16*:* Clusters obtained for Participant 2.

The MST multidimensional scaling analysis for Participant 2 is shown in Figure J-17, with its related Shepard diagram in Figure J-18. Examining Figure J-17 visually for patterns, we see that it is mostly likely a simplex pattern as all the constructs could be interpreted to lie under the same line. This indicates that there was a single dimension effecting the participant's determination of constructs. The similarity of constructs such as the construct "*animal sounds – sounds from objects*" to each other would seem to suggest that this dimension was as suggested in the previous principal component analysis of the participant's constructs that of the overall 'naturalness' of the sources of the sounds.

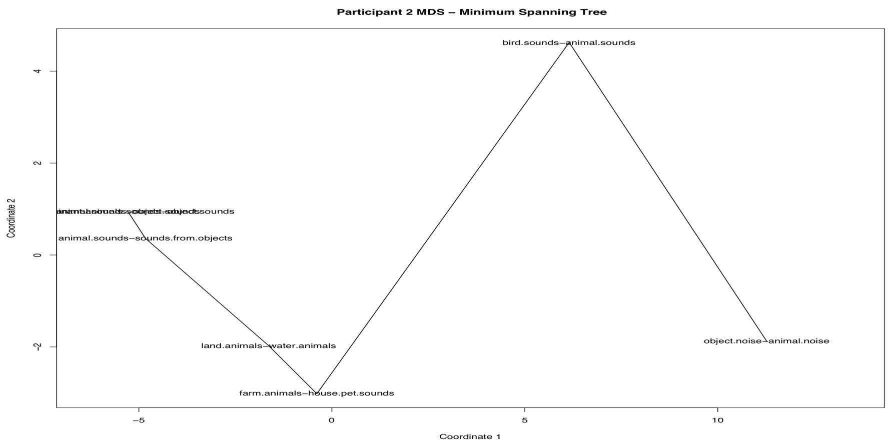


Fig. J-17: *MDS minimum spanning tree analysis of the RGT constructs for Participant 2.*

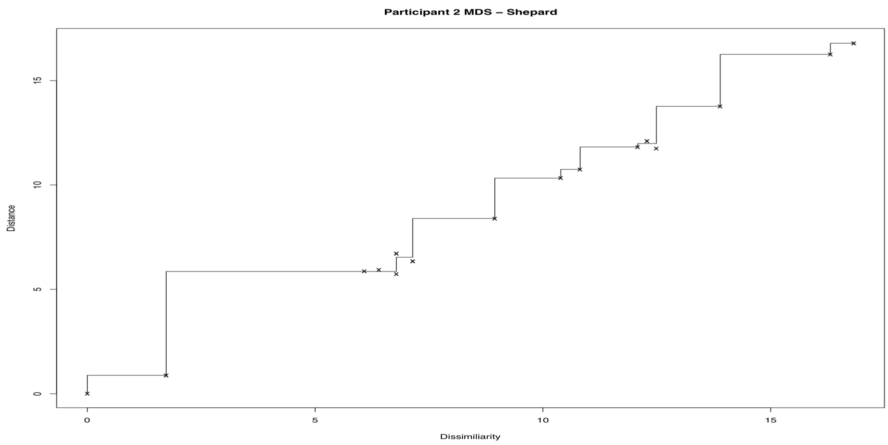


Fig. J-18: *Shepard diagram of the MDS analysis for Participant 2.*

Participant 2 - Cluster Analysis and MDS - Elements

The cluster analysis for Participant 2 was shown already in Figure J-19 with a cutoff level of approximately 0.65 which results in 4 clusters. These were objects, seaside birds & animals, animals, and birds. These were clustered into two larger clusters which were either object sounds or bird & animal sounds. Cluster 1 contained sounds that were seen as object like sounds. Cluster 2 contains sounds that were seen as bird or animal like sounds. Participant 2 is notable in not having any confusion over sound identification as discussed in Section 5.1.1.

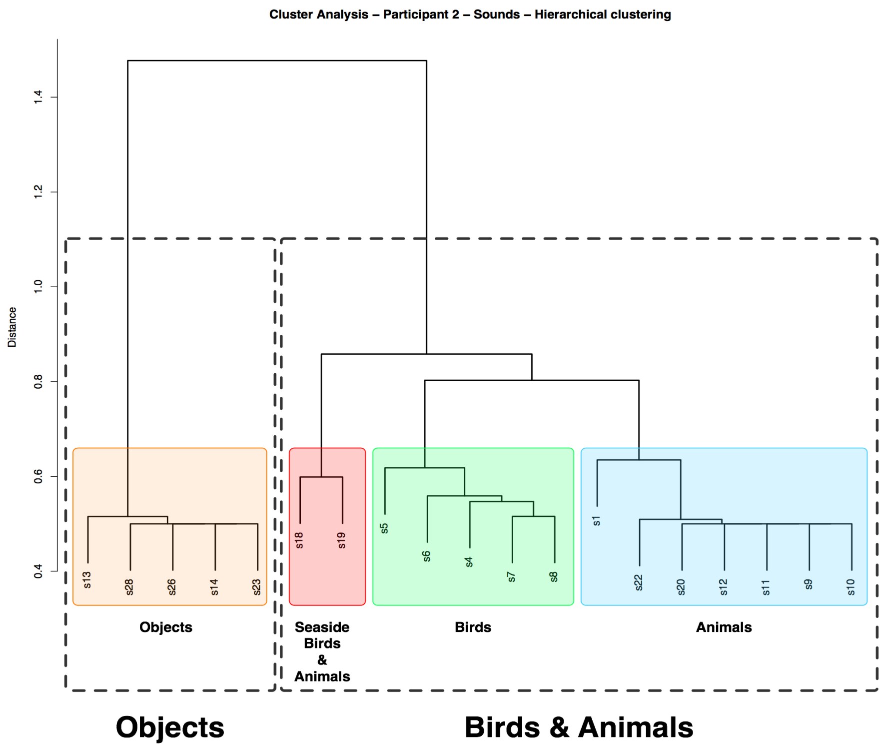


Fig. J-19: *The cluster analysis of the RGT elements (sounds) for Participant 2.*

The MST multidimensional scaling analysis for Participant 2 is shown in Figure J-20, with its related Shepard diagram in Figure J-21. The MST MDS analysis produces a circumplex pattern as *s19* could not be interpreted to lie under the same line as the elements *s28* or *s5*. This indicates that there are two or more dimensions effecting the participant's determination of elements. This somewhat supports the earlier findings of the principal component analysis for the participant's elements as shown in Table J-15.

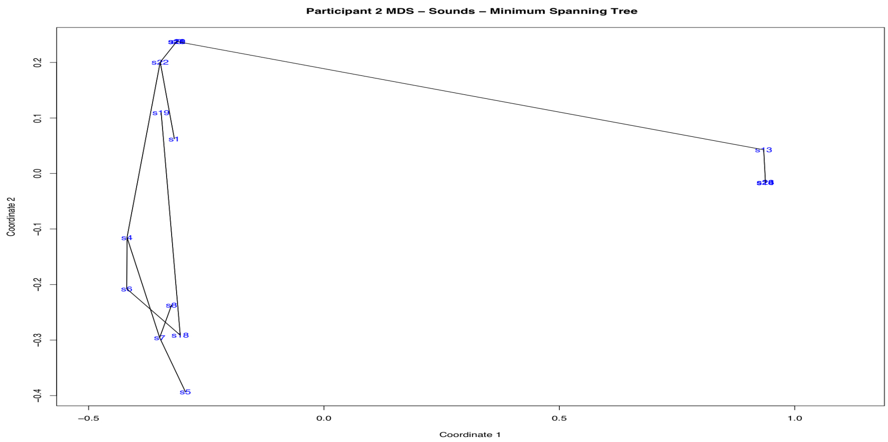


Fig. J-20: *MDS minimum spanning tree analysis of the RGT elements (sounds) for Participant 2.*

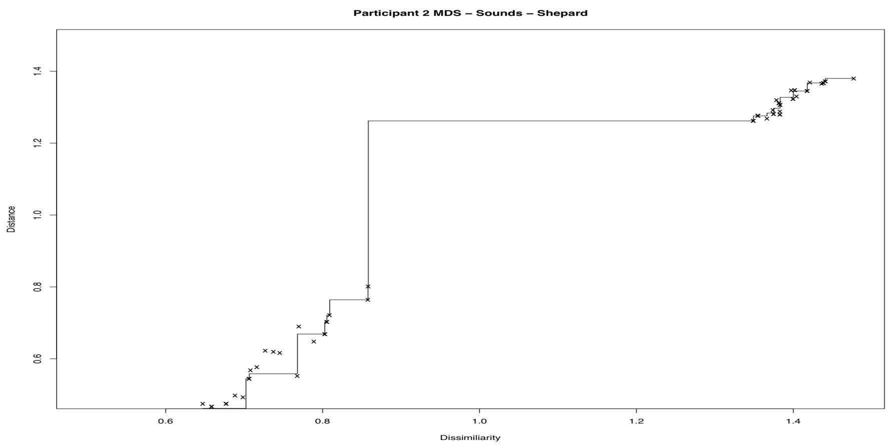


Fig. J-21: *Shepard diagram of the MDS analysis of the RGT elements (sounds) for Participant 2.*

J.1.2 Results and Observations for Participant 2

The results of the analysis of the constructs from Participant 2 from the PCA, CA, and MDS analysis suggest that the naturalness of a sound was a factor. This was further specified into a factor which split the type of animal or bird making the sound. The results of the analysis of the elements (sounds) from Participant 2 from the PCA, CA, and MDS analysis suggested that the contrast of animal/bird sounds with those where were object based was a factor.

Participant 3 - PCA - Constructs

Examining the correlations among the constructs for Participant 3, shown in Table J-17 in Appendix J.1.2, we can say that *time base night/early* (construct 1) is associated with *nature sounds* (construct 4, 0.70) and they are definitely not heard as *non-animal sounds* (construct 5, -0.60). There was a significant correlation (0.93) between "*nature sounds - man made*" and ”*non animal sounds - animal sounds*”. There was a significant correlation (0.89) between several of the constructs, ”*nature sounds/warning but not man made - man made/attention grabbing*” and the following: ”*nature sounds - man made*” and ”*non animal sounds - animal sounds*”. There was also a significant correlation (0.85) between ”*nature sounds - man made*” and ”*complete animal sounds - incomplete animal sounds*”. The ”*nature sounds – man made*” construct is the construct most closely associated with the other constructs as shown in Table J-18. ”*non animal sounds – animal sounds*” is the next most closely associated construct.

	1	**2**	**3**	**4**	**5**	**6**	**7**
1 time based night/ early — anytime/ attention grabber	1.00	0.60	-0.13	0.70	-0.60	-0.20	0.63
2 nature sounds/warning but not man made — man made/attention grabber		1.00	-0.05	**0.89**	**-0.89**	0.01	0.72
3 more familiar — less familiar			1.00	0.02	0.10	-0.24	0.25
4 nature sounds — man made				1.00	**-0.93**	-0.03	**0.85**
5 non animal sounds — animal sounds					1.00	-0.12	-0.74
6 sheep like sounds— non sheep like						1.00	-0.37
7 complete animal sound — incomplete animal sounds							1.00

Table J-17*:* Correlation analysis of constructs for Participant 3.

The interpretation of the first principal component in Table J-19 contrasts the animal like and alerting type of sounds with those where were man made and less familiar. The second principal component in Table J-19 contrasts the familiarity and the completeness of the sounds, implying that after the ‘alerting animal’ sounds are taken into account, the main source of variation is between sounds from the familiarity and sense of completeness of the sound. The first two components are shown in Figure J-22, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure J-23.

Construct	Root-mean-square correlation
time based night/early — anytime/attention grabber	0.62
nature sounds/warning but not man made — man made/attention grabber	0.70
more familiar — less familiar	0.40
nature sounds — man made	0.74
non animal sounds — animal sounds	0.72
sheep like sounds —non sheep like	0.42
complete animal sound — incomplete animal sounds	0.70
Average of statistic	0.61
Standard deviation of statistic	0.14

Table J-18*:* Root-mean-square (average) correlation among constructs analysis for Participant 3.

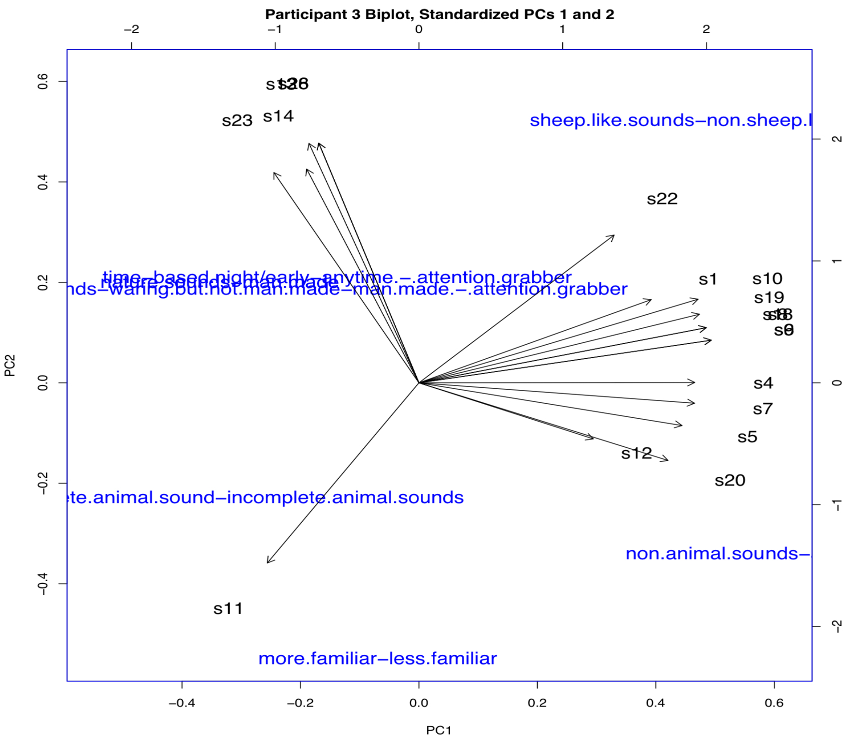


Fig. J-22*: The first two components of the principal–components analysis for Participant 3.*

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
time based night/early/anytime - attention grabber	-0.38	-	0.02	
nature sounds/waring but not manmade - manmade / attention grabber	-0.45	-	-0.12	
more familiar - less familiar	0.05		0.59	-
nature sounds - man made	-0.48	-	-0.05	
non animal sounds - animal sounds	0.47	+	0.20	
sheep like sounds - non sheep like	0.08	(+)	-0.71	+
complete animal sound - incomplete animal sounds	-0.43	-	0.30	(-)
Standard deviation Proportion of Variance Cumulative Proportion	2.21 0.70 0.70		1.26 0.23 0.93	

Table J-19*:* Principal–Components Analysis for Participant 3.

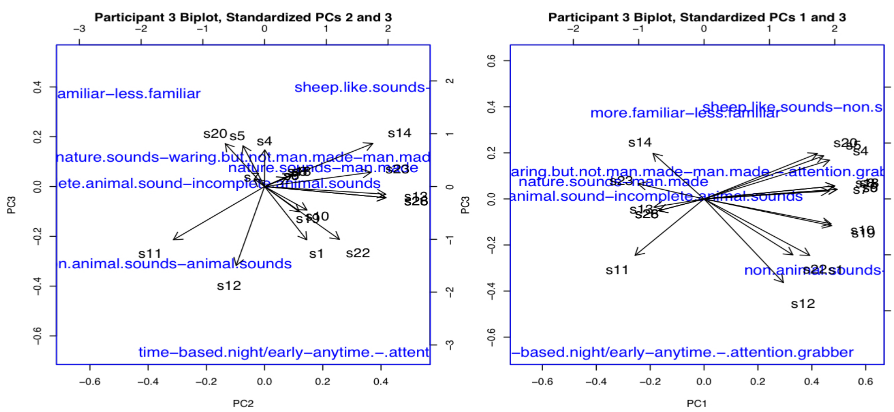


Fig. J-23*: The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 3.*

Participant 3 - PCA - Elements

The results of the elicitation task for Participant 3 are shown below in Table J-20. Participants responded in free-text format to what they thought each sound was as we can see from Table J-20, these text descriptions were often highly descriptive and described the events or actions.

ID	Sound Description	Participant's Description
s1	cat	*meow, singular (attentive)*
s4	owl	*hooting owls nightlife* *or cool iceberg 20% below*
s5	bird song	*piercing pitch birds (annoying)*
s6	bird song	*summer birds (warning)*
s7	bird song	*summer chirping, low pitched*
s8	rooster	*rooster, morning sound with* *speaker noise at mid to end*
s9	donkey	*donkey braying*
s10	horse	*neighing horse with pug completing nasal sounds*
s11	goat	*half a bah from sheep*
s12	sheep	*single long bleat from sheep*
s13	glass breaking	*glass breaking pitch increasing* *and something more*
s14	church bell ringing	*bells ringing from medium sized bell*
s18	seagull	*seagulls, seaside, fishermen, breeze,* *wave, and early in the day*
s19	seal	*donkey braying, water on sand from waves,* *and something else (unattractive)*
s20	horse	*horse neighing*
s22	lion roaring	*lion roaring, loud but base, not piercing*
s23	power saw	*plane, engine fans shutting off*
s26	coins counting	*coins, counting, and on a table*
s28	heavy ball bouncing	*marble dropping and bouncing on table*

Table J-20*:* Descriptions by Participant 3 for the elements.

Examining the correlations among the constructs for Participant 3, shown in Table J-21, we can say that element *s1* (the sound of cat meowing) was definitely associated with *s10 (0.91)*, *s19 (0.94)*, and *s22 (0.89)* (the sound of horse neighing, the sound of a seal roaring, and the sound of a lion roaring, respectively). The elements *s6*, and *s9* (the sound of an bird singing, and the sound of a donkey braying, respectively) where the elements or sounds most closely associated with the other elements as shown in Table J-22.

	s1	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s18	s19	s20	s22	s23	s26	s28
s1	1.00	0.50	0.47	0.73	0.72	0.68	0.73	**0.91**	−0.33	0.61	0.05	−0.25	0.68	**0.94**	0.42	**0.89**	−0.11	0.12	0.12
s4		1.00	**0.85**	**0.87**	**0.86**	**0.91**	**0.87**	0.77	−0.65	0.36	−0.33	−0.18	**0.91**	0.72	**0.84**	0.44	−0.37	−0.36	−0.36
s5			1.00	**0.83**	**0.87**	0.77	**0.83**	0.64	−0.44	0.23	−0.50	−0.33	0.77	0.69	**0.97**	0.32	−0.46	−0.44	−0.44
s6				1.00	0.79	**0.98**	**1.00**	**0.92**	−0.63	0.45	−0.21	−0.18	**0.98**	**0.92**	0.73	0.65	−0.37	−0.16	−0.16
s7					1.00	0.76	0.79	**0.80**	−0.39	0.44	−0.40	−0.42	0.76	**0.81**	**0.90**	0.53	−0.37	−0.37	−0.37
s8						1.00	**0.98**	**0.92**	−0.71	0.46	−0.15	−0.10	**1.00**	**0.89**	0.68	0.66	−0.32	−0.14	−0.14
s9							1.00	**0.92**	−0.63	0.45	−0.21	−0.18	**0.98**	**0.92**	0.73	0.65	−0.37	−0.16	−0.16
s10								1.00	−0.59	0.60	−0.01	−0.15	**0.92**	**0.98**	0.57	**0.86**	−0.20	0.01	0.01
s11									1.00	0.15	−0.44	−0.60	−0.71	−0.51	−0.33	−0.49	−0.37	−0.41	−0.41
s12										1.00	−0.34	−0.61	0.46	0.57	0.20	0.61	−0.55	−0.36	−0.36
s13											1.00	**0.88**	−0.15	−0.08	−0.61	0.33	**0.92**	**0.98**	**0.98**
s14												1.00	−0.10	−0.23	−0.44	0.07	**0.88**	**0.84**	**0.84**
s18													1.00	**0.89**	0.68	0.66	−0.32	−0.14	−0.14
s19														1.00	0.61	**0.83**	−0.25	−0.01	−0.01
s20															1.00	0.20	−0.51	−0.57	−0.57
s22																1.00	0.13	0.36	0.36
s23																	1.00	**0.90**	**0.90**
s26																		1.00	**1.00**
s28																			1.00

Table J-21: Correlation Analysis for Participant 3 Elements.

Element	Root-mean-square correlation
s1	0.54
s4	0.64
s5	0.62
s6	0.66
s7	0.65
s8	0.64
s9	0.66
s10	0.62
s11	0.52
s12	0.46
s13	0.45
s14	0.44
s18	0.64
s19	0.62
s20	0.61
s22	0.53
s23	0.49
s26	0.44
s28	0.44
Average of statistic	0.56
Standard deviation of statistic	0.09

Table J-22*:* Root-mean-square (average) correlation among elements analysis for Participant 3.

The interpretation of the first principal component in Table J-23 contrasts the animal/bird sounds with those where were object based. The second principal component in Table J-23 contrasts the location of the source of the sounds between those of could occur in a home / domestic environment and of an outdoor environment, implying that after the ‘naturalness’ of the sources of the sounds is taken into account, the main source of variation is between sounds from the location associated with where the sounds are typically heard. The first two components are shown in Figure J-24, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure J-25. These figures can help in illustrating the interpretation of the principal components.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
s1	−0.23	-	0.17	+
s4	−0.25	-	0.07	
s5	−0.25	-	−0.01	
s6	−0.25	-	0.12	(+)
s7	−0.26	-	0.00	
s8	−0.25	-	0.15	(+)
s9	−0.25	-	0.12	(+)
s10	−0.25	-	0.18	+
s11	0.09	(+)	−0.46	-
s12	−0.22	+	−0.11	
s13	0.21	+	0.32	+
s14	0.20	+	0.33	+
s18	−0.25	-	0.15	(+)
s19	−0.25	-	0.15	(+)
s20	−0.25	-	−0.06	
s22	−0.17	-	0.34	+
s23	0.23	+	0.27	+
s26	0.21	+	0.33	+
s28	0.21	+	0.33	+
Standard deviation	3.81		1.85	
Proportion of Variance	0.77		0.18	
Cumulative Proportion	0.77		0.95	

Table J-23*:* Principal–Components Analysis for Participant 3 Elements.

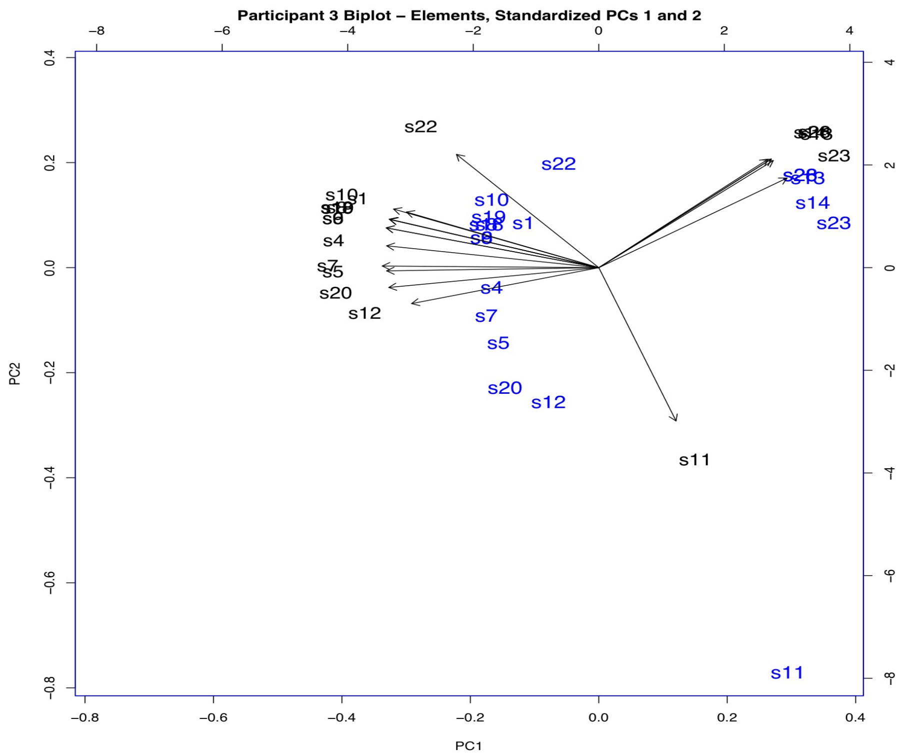


Fig. J-24: *The first two components of the principal–components analysis for Participant 3 Elements.*

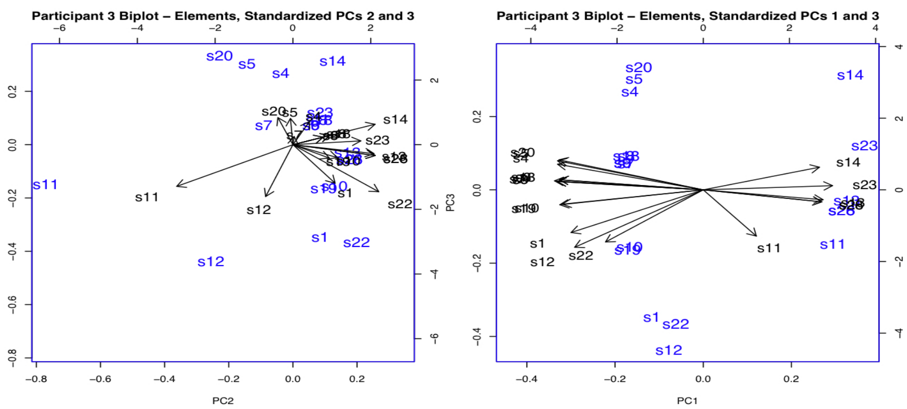


Fig. J-25: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 3 Elements.*

Participant 3 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 3 is shown in Figure J-26 with a cutoff level of 6.1 which resulted in 3 clusters. The resulting clusters are shown in Table J-24. Cluster 1 contained sounds that were seen as either natural sounds with reference to the time they normally occur or man made sounds which were attention grabbing. The shortest distance match in this cluster was 3.6 between ”*nature sounds/waring but not manmade – manmade/attention grabber*” and ”*nature sounds – man made*”. The next sub-cluster joined at a distance of 5.09 and was ”*complete animal sound – incomplete animal sounds*”. The next sub-cluster joined at a distance of 6.48 and was ”*time based night/early/anytime – attention grabber*”. The last sub-cluster joined at a distance of 7.14 and was ”*more familiar — less familiar*”. These constructs suggest that the sounds were distinctly divided into natural time referenced sounds or alarming / alerting man made sounds.

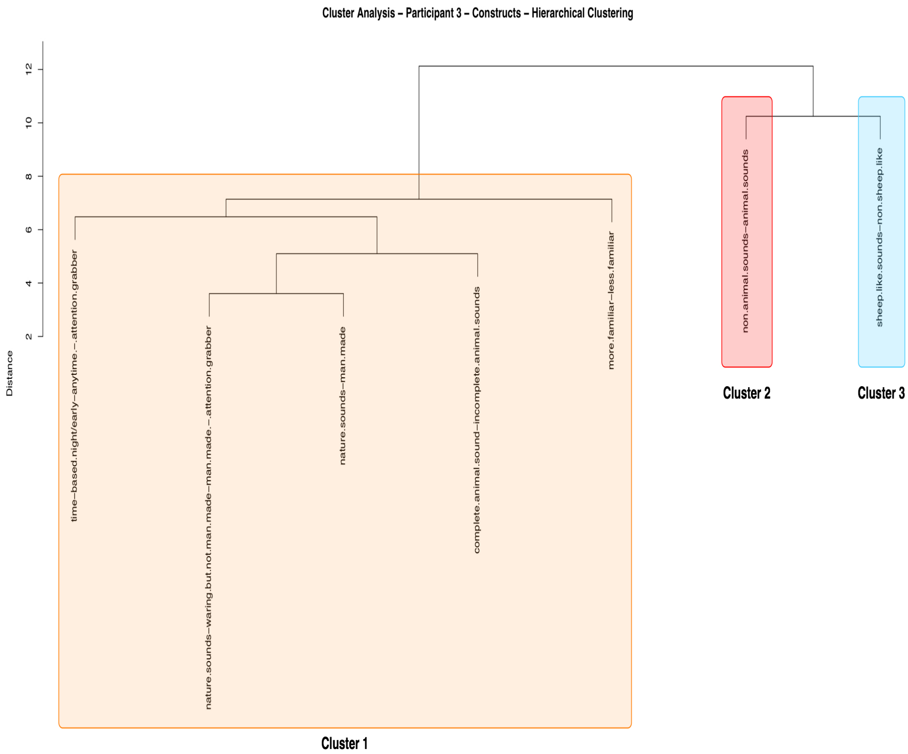


Fig. J-26*: The cluster analysis of the RGT constructs for Participant 3.*

	Constructs *Emergent Pole* - *Implicit Pole*
1	time based night/early/anytime — attention grabber nature sounds/waring but not manmade — manmade/attention grabber nature sounds — man made complete animal sound — incomplete animal sounds more familiar — less familiar
2	non animal sounds — animal sounds
3	sheep like sounds — non sheep like

Table J-24*:* Clusters obtained for Participant 3.

The MST multidimensional scaling analysis for Participant 3 is shown in Figure J-27, with its related Shepard diagram in Figure J-28. Examining Figure J-27 visually for patterns, we see that it is mostly likely a circumplex pattern as all the constructs could not be interpreted to lie under the same line. This indicates that there was two or more dimensions effecting the participant's determination of constructs. The dissimilarity of constructs as shown in Table J-24 to each other would seem to further support this interpretation.

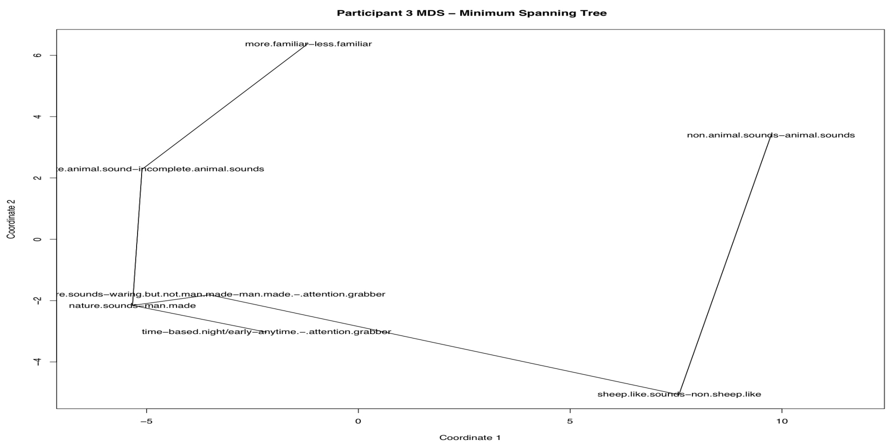


Fig. J-27: *MDS minimum spanning tree analysis of the RGT constructs for Participant 3.*

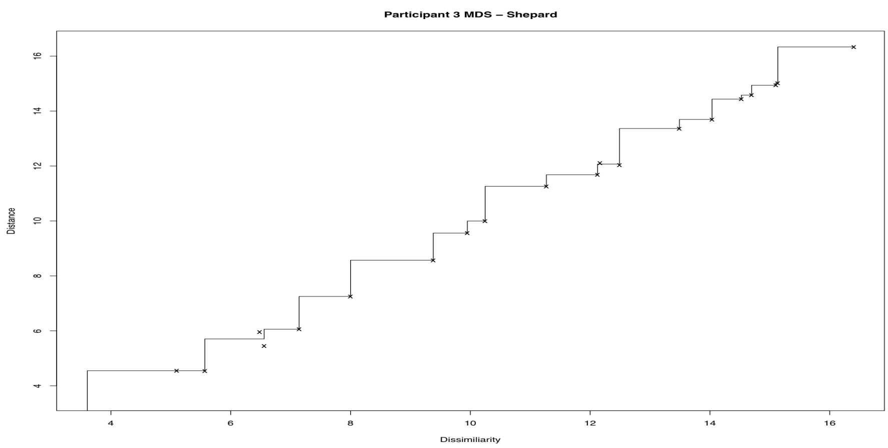


Fig. J-28: *Shepard diagram of the MDS analysis for Participant 3.*

Participant 3 - Cluster Analysis and MDS - Elements

The cluster analysis for Participant 3 is shown in Figure J-29 with a cutoff level of approximately 0.75 which results in 3 clusters with an outlier at a level of 1.2. This consisting of a single sound, s11, which is the sound of a goat bleating. These were birds & animals, animals, and objects. These were clustered into two larger clusters which were either bird & animal sounds or object sounds. Cluster 1 contains sounds that were seen as bird or animal like sounds. Cluster 2 contained sounds that were seen as object like sounds. One noted exception to this cluster is s23, the sound of a power saw in use, which was found to have identification issues. This is discussed further in Section 5.1.1. An outlier to the two clusters was the sound of a goat bleating, s11, as the later causal uncertainty studies (see Section 5.1.1) show, there was no difficulty in its identification so other factors must be involved and this is an area of further study.

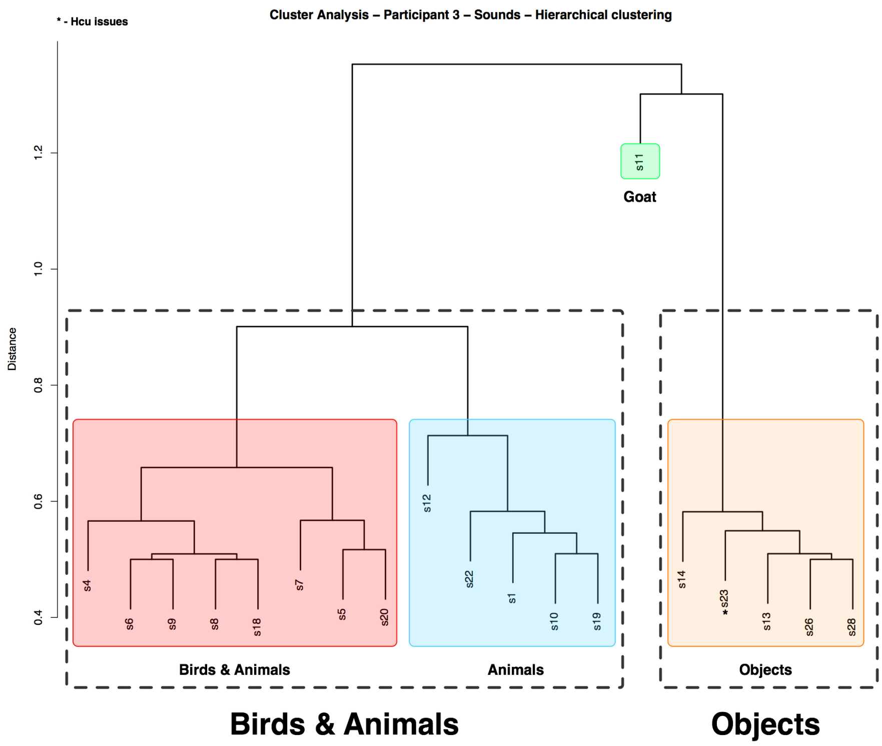


Fig. J-29: *The cluster analysis of the RGT elements (sounds) for Participant 3.*

The MST multidimensional scaling analysis for Participant 1 is shown in Figure J-30, with its related Shepard diagram in Figure J-31. Examining Figure J-30 visually for patterns, we see that it is mostly likely a radex pattern as *s11* could not be interpreted to lie under the same line as the elements *s23* or *s5*. This indicates that there are three or more dimensions effecting the participant's determination of elements. The number of dimensions and complexity for this interpretation requires further study to clarify the particular dimensions used by Participant 3 to classify the elements.

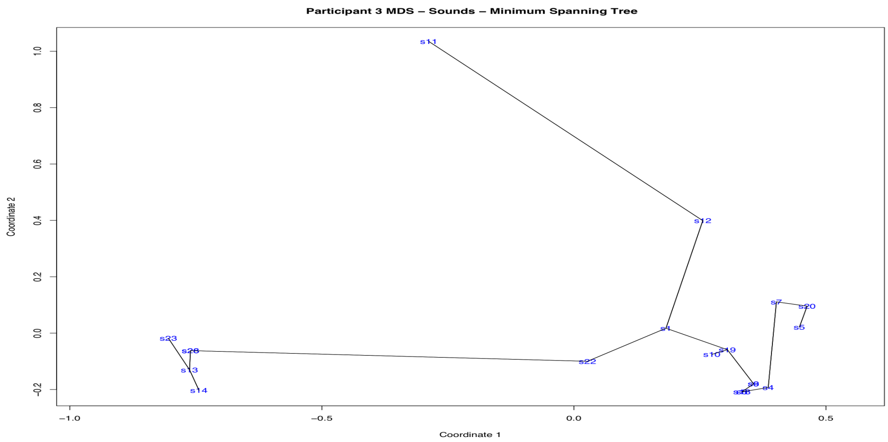


Fig. J-30: *MDS minimum spanning tree analysis of the RGT elements (sounds) for Participant 3.*

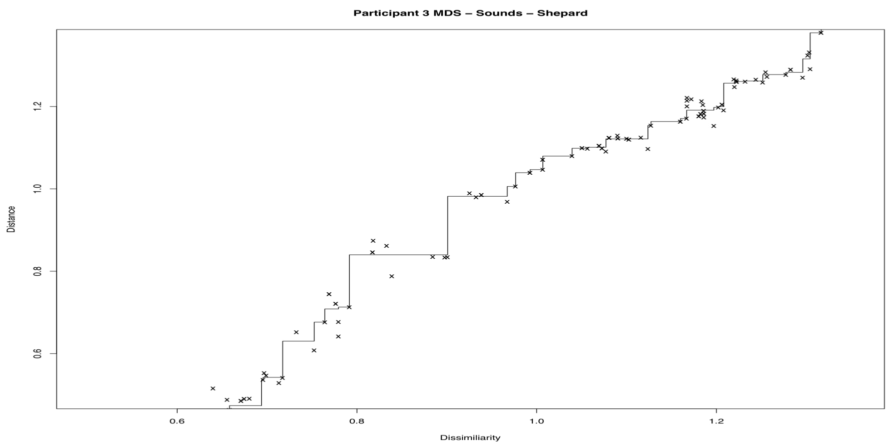


Fig. J-31: *Shepard diagram of the MDS analysis of the RGT elements (sounds) for Participant 3.*

J.1.3 Results and Observations for Participant 3

The results of the analysis of the constructs from Participant 3 from the PCA, CA, and MDS analysis suggest that natural sounds with reference to their familiarity versus unfamiliar man made sounds was a factor. A further factor was the completeness of the animal or bird sounds. The results of the analysis of the elements (sounds) from Participant 3 from the PCA, CA, and MDS analysis suggested that the contrast of animal/bird sounds with those where were object based was a factor. The typical location associated with the sounds was a further factor revealed by this analysis.

Participant 4 - PCA - Constructs

Examining the correlations among the constructs for Participant 4, shown in Table J-25, we can say that *hard to identify* (construct 1) was definitely not associated with *countryside* *sounds* (construct 7, -0.57). The rest of the correlations between construct one and the other constructs, apart from construct seven, do not have any significant correlation. There was only other significant correlation in the results for this participant was that *sheep* (construct 4) where not heard as *man made things* (construct 6, -0.58). The "*animals alone/single – animals in a farm*" construct and the "*relaxing open places – dark closed spaces*" construct were jointly (0.53) correlated with the other constructs as shown in Table J-26.

	1	**2**	**3**	**4**	**5**	**6**	**7**
1 hard to identify — easy to identify	1.00	0.39	0.37	-0.08	-0.27	0.12	-0.57
2 annoying — relaxing		1.00	0.42	0.07	-0.57	0.18	-0.31
3 animals alone/single — animals in a farm			1.00	0.33	-0.47	-0.42	-0.37
4 sheep — village				1.00	-0.18	**-0.58**	-0.14
5 relaxing open places — dark closed spaces					1.00	-0.19	0.54
6 man made things — animals						1.00	-0.06
7 countryside sounds — weird animal sounds							1.00

Table J-25*:* Correlation Analysis for Participant 4.

Construct	Root-mean-square correlation
hard to identify — easy to identify	0.49
annoying — relaxing	0.51
animals alone/single — animals in a farm	0.53
sheep — village	0.46
relaxing open places — dark closed spaces	0.53
man made things — animals	0.48
countryside sounds — weird animal sounds	0.52
Average of statistic	0.5
Standard deviation of statistic	0.03

Table J-26*:* Root-mean-square (average) correlation among constructs analysis for Participant 4.

The interpretation of the first principal component in Table J-27 contrasts the open, relaxing, identifiable sounds with those where were singular, closed and annoying. The second principal component in Table J-27 contrasts the man made nature sounds versus the natural sounds, implying that after the first principal component, the next main source of variation is between sounds from the naturalness of the sources of the sounds. The first two components are shown in Figure J-32 and this can help to illustrate the interpretation. The components 2 and 3 and the components 1 and 3 of the principal component analysis are shown in Figure J-33.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
hard to identify - easy to identify	-0.39	-	0.21	(+)
annoying - relaxing	-0.41	-	0.17	(+)
animals alone - single animals in a farm	-0.44	-	-0.31	(-)
sheep - village	-0.16	(-)	-0.59	-
relaxing open places - dark closed spaces	0.49	+	-0.11	
man made things - animals	0.02		0.69	+
countryside sounds - weird animal sounds	0.47	+	-0.09	
Standard deviation	2.04		1.43	
Proportion of Variance	0.60		0.29	
Cumulative Proportion	0.60		0.89	

Table J-27: Principal–Components Analysis for Participant 4.

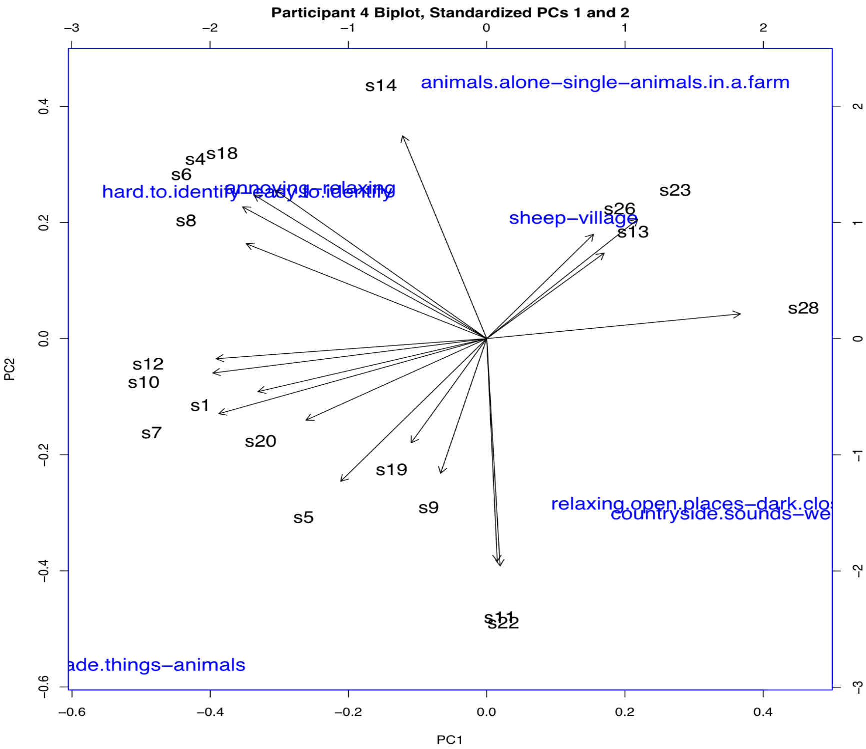


Fig. J-32: *The first two components of the principal–components analysis for Participant 4.*

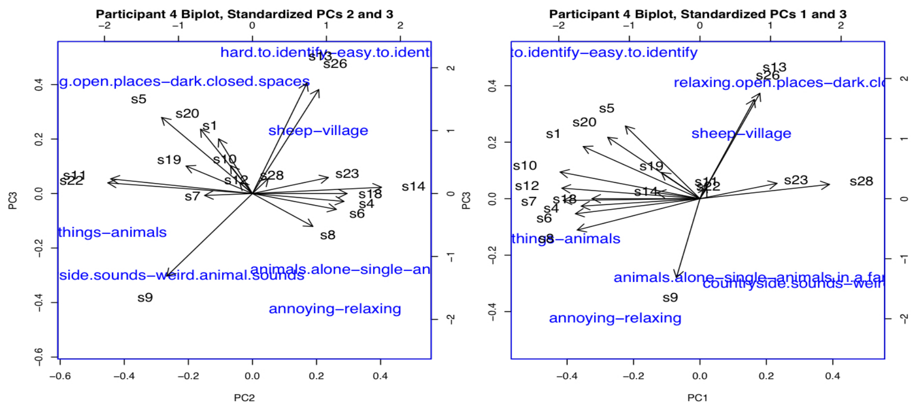


Fig. J-33: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 4.*

Participant 4 - PCA - Elements

The results of the elicitation task for Participant 4 are shown below in Table J-28. Participants responded in free-text format to what they thought each sound was as we can see from Table J-28, these text descriptions were often highly descriptive and described the events or actions.

ID	Sound Description	Participant's Description
s1	cat	*little kitten, cute and soft*
s4	owl	*owl in a forest*
s5	bird song	*high pitched bird*
s6	bird song	*rainforest*
s7	bird song	*little chicks, baby birds*
s8	rooster	*cock in the farm, waking, morning*
s9	donkey	*indistinguishable animal complaining*
s10	horse	*horse braying*
s11	goat	*cut off sheep noise,*
s12	sheep	*sheep*
s13	glass breaking	*breaking thin glass*
s14	church bell ringing	*church bells in a village by the river*
s18	seagull	*seagulls, relaxing seaside*
s19	seal	*donkey by the river*
s20	horse	*horse braying, dry, dark, even brown*
s22	lion roaring	*lion, low pitched*
s23	power saw	*hoover or hand dryer*
s26	coins counting	*counting coins*
s28	heavy ball bouncing	*marble falling on wooden floor*

Table J-28*:* Descriptions by Participant 4 for the elements.

Examining the correlations among the constructs for Participant 4, shown in Table J-29, we can say that element *s12* (the sound of a sheep bleating) was definitely associated with *s7 (0.85)*, and *s10 (0.95)* (the sound of a bird singing, and the sound of a horse neighing, respectively) while being definitely not associated with *s28 (-0.84)* (the sound of a heavy ball bouncing). The *s7* element (the sound of an bird singing) is the element or sounds most closely associated with the other elements as shown in Table J-30. The *s10* element (the sound of a horse neighing) is the next most closely associated element.

The interpretation of the first principal component in Table J-31 contrasts the animal/bird sounds with those where were object based. The second principal component in Table J-31 contrasts the location of the source of the sounds between those of could occur in a indoors environment (or on indoor medium such as radio or television) and of an outdoor environment,

	s1	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s18	s19	s20	s22	s23	s26	s28
s1	1.00	0.49	0.73	0.46	0.72	0.32	0.21	**0.90**	0.14	0.71	−0.09	0.26	0.50	0.47	0.77	0.31	−0.33	−0.04	−0.58
s4		1.00	0.06	**0.98**	0.56	**0.86**	−0.13	0.63	−0.60	0.63	−0.15	0.75	**0.98**	0.07	0.21	−0.57	−0.19	−0.10	−0.59
s5			1.00	0.04	0.63	0.05	0.02	0.61	0.58	0.48	0.12	−0.32	0.07	0.63	0.62	0.49	−0.67	0.02	−0.43
s6				1.00	0.60	**0.94**	−0.12	0.68	−0.51	0.73	−0.24	0.64	**0.91**	−0.07	0.28	−0.56	−0.20	−0.17	−0.70
s7					1.00	0.68	0.24	**0.85**	0.18	**0.85**	−0.46	−0.03	0.49	0.46	0.68	0.25	−0.65	−0.54	−0.75
s8						1.00	−0.11	0.64	−0.34	0.78	−0.37	0.38	0.75	−0.12	0.27	−0.50	−0.34	−0.34	−0.77
s9							1.00	0.16	0.34	0.06	**−0.82**	−0.18	−0.13	0.25	−0.01	0.58	−0.27	−0.74	−0.18
s10								1.00	0.14	**0.95**	−0.29	0.19	0.55	0.20	0.81	0.14	−0.38	−0.20	**−0.84**
s11									1.00	0.14	−0.21	**−0.87**	−0.66	0.13	0.30	0.73	−0.45	−0.19	−0.24
s12										1.00	−0.36	0.09	0.50	0.00	0.76	0.02	−0.38	−0.27	**−0.92**
s13											1.00	0.18	−0.06	0.00	−0.03	−0.28	0.36	**0.92**	0.49
s14												1.00	**0.82**	−0.06	−0.08	−0.64	0.36	0.28	−0.03
s18													1.00	0.20	0.14	−0.56	−0.18	−0.04	−0.45
s19														1.00	0.19	0.47	−0.59	−0.26	0.09
s20															1.00	0.44	−0.09	−0.02	−0.50
s22																1.00	−0.22	−0.35	0.08
s23																	1.00	0.50	0.52
s26																		1.00	0.33
s28																			1.00

Table J-29: Correlation Analysis for Participant 4 Elements.

Element	Root-mean-square correlation
s1	0.48
s4	0.50
s5	0.40
s6	0.52
s7	0.56
s8	0.50
s9	0.29
s10	0.53
s11	0.41
s12	0.51
s13	0.34
s14	0.38
s18	0.47
s19	0.28
s20	0.38
s22	0.43
s23	0.40
s26	0.33
s28	0.50
Average of statistic	0.43
Standard deviation of statistic	0.08

Table J-30: Root-mean-square (average) correlation among elements analysis for Participant 4.

in particular between animal sounds and outdoor sounds with birds, implying that after the 'naturalness' of the sources of the sounds is taken into account, the main source of variation is between sounds from the location associated with where the sounds are typically heard. The first two components are shown in Figure J-34, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure J-35.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
s1	−0.30	-	0.03	
s4	−0.21	-	−0.29	-
s5	−0.21	-	0.23	+
s6	−0.23	-	−0.27	-
s7	−0.31	-	0.03	
s8	−0.26	-	−0.23	-
s9	−0.14	(-)	0.25	+
s10	−0.31	-	−0.03	
s11	−0.06		0.38	+
s12	−0.30	-	−0.06	
s13	0.24	+	−0.10	(-)
s14	−0.02		−0.38	-
s18	−0.19	-	−0.31	-
s19	−0.14	(-)	0.24	+
s20	−0.25	-	0.10	(+)
s22	−0.03		0.39	+
s23	0.26	+	−0.16	(-)
s26	0.23	+	−0.15	(-)
s28	0.31	+	0.05	
Standard deviation	3.16		2.50	
Proportion of Variance	0.52		0.33	
Cumulative Proportion	0.52		0.85	

Table J-31*:* Principal–Components Analysis for Participant 4 Elements.

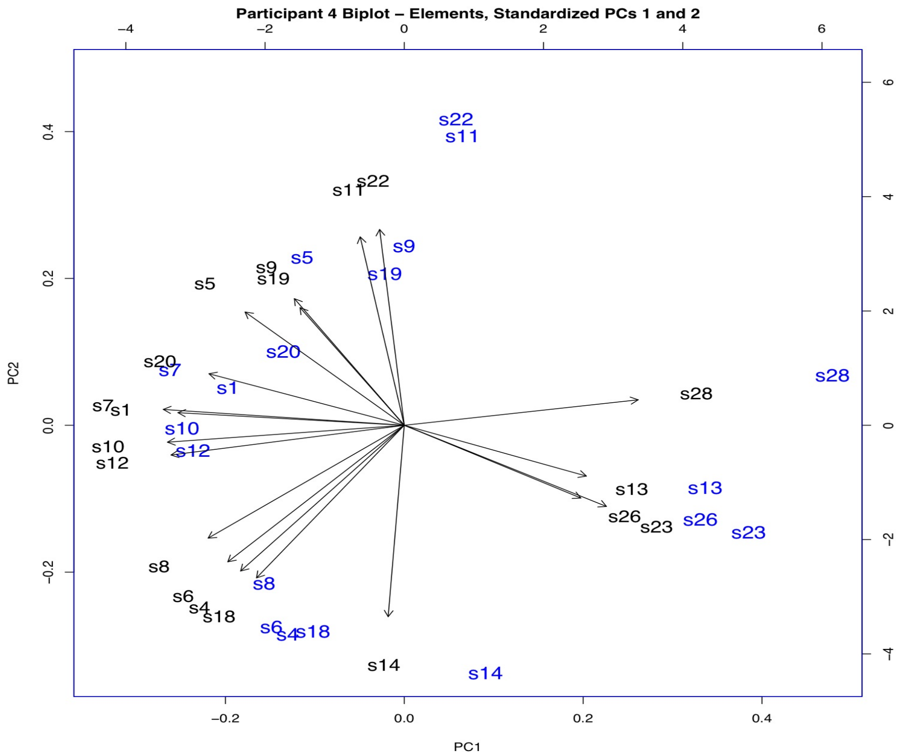


Fig. J-34: *The first two components of the principal–components analysis for Participant 4 Elements.*

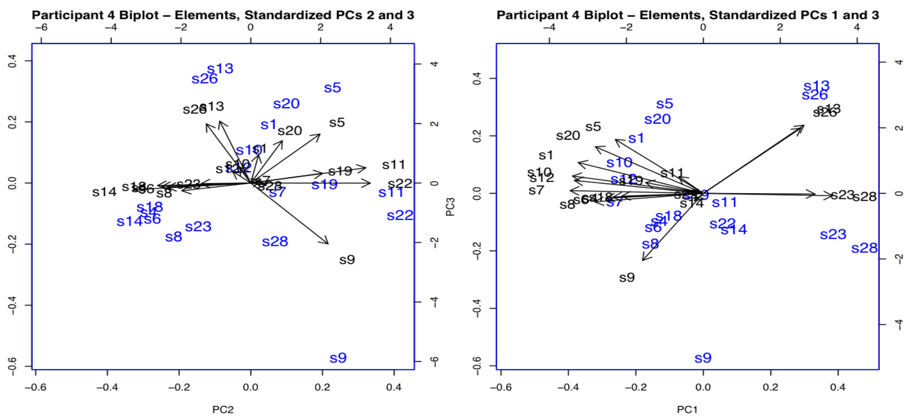


Fig. J-35: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 4 Elements.*

Participant 4 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 4 is shown in Figure J-36 with two possible cutoff levels. These are at 7.3 and 8.4 and would result in 3 and 7 clusters respectively. The resulting clusters for cutoff level 8.4 are shown in Table J-32. Cluster 2 had the shortest significant distances from the set of clusters and contained sounds that were seen as either sounds from nature or dark / weird sounds. The shortest distance match in this cluster was 6.32 between ”*relaxing open places – dark closed spaces*” and ”*countryside sounds – weird animal sounds*”. This cluster only contained two constructs. These constructs suggest that the sounds were distinctly divided into sounds which occurred in outdoors and darker / weird sounds.

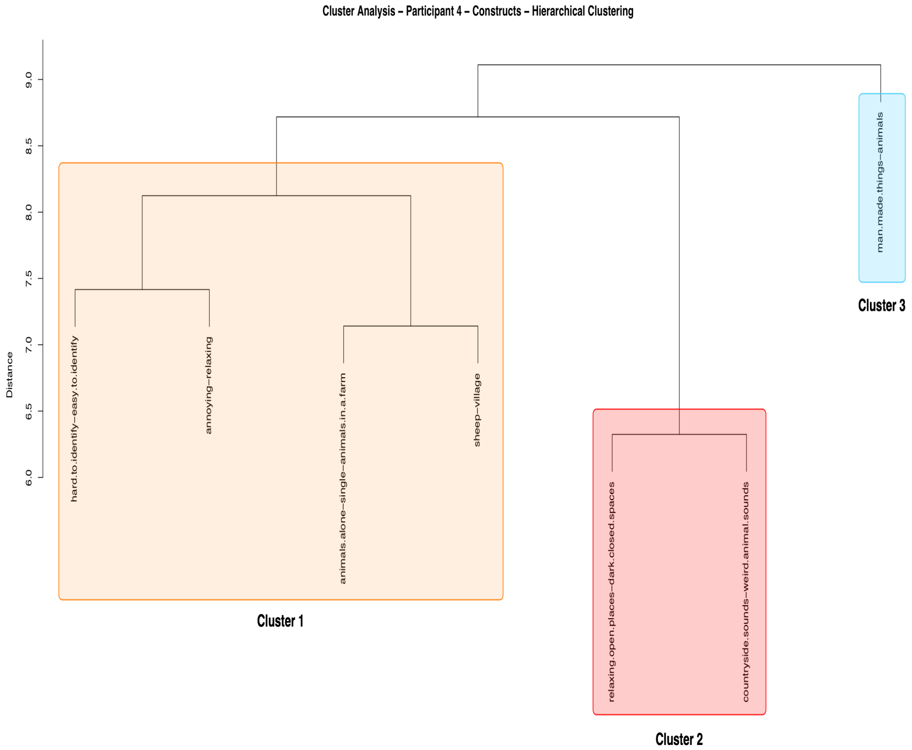


Fig. J-36: *The cluster analysis of the RGT constructs for Participant 4.*

	Constructs *Emergent Pole* - *Implicit Pole*
1	hard to identify — easy to identify annoying — relaxing animals alone — single animals in a farm sheep — village
2	relaxing open places — dark closed spaces countryside sounds — weird animal sounds
3	man made things — animals

Table J-32: Clusters obtained for Participant 4.

The MST multidimensional scaling analysis for Participant 4 is shown in Figure J-37, with its related Shepard diagram in Figure J-38. Examining Figure J-37 visually for patterns, we see that it is mostly likely a simplex pattern as all the constructs could be interpreted to lie under the same line. This indicates that there was a single dimension effecting the participant's determination of constructs. The similarity of several of constructs to the construct "*relaxing open places – dark closed spaces*" would seem to suggest that this as a dimension. In the previous principal component analysis of the participant's constructs the first principal component was related to a dimensional ranging from open, relaxing, identifiable sounds to those which were were singular, closed and annoying.

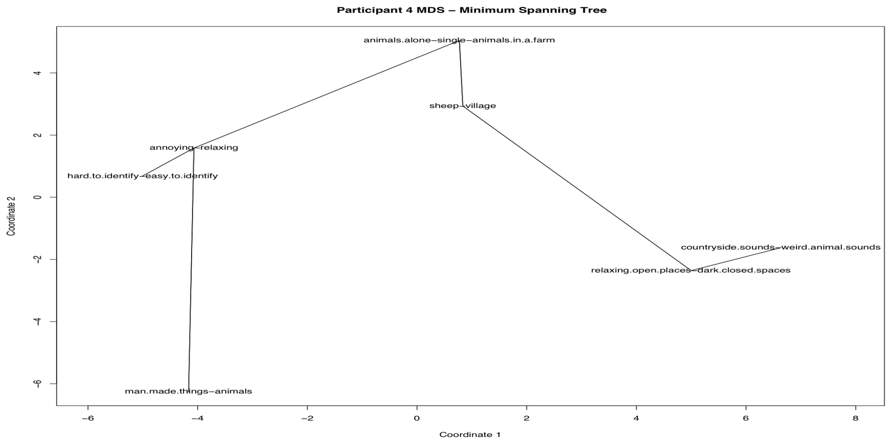


Fig. J-37: *MDS minimum spanning tree analysis of the RGT constructs for Participant 4.*

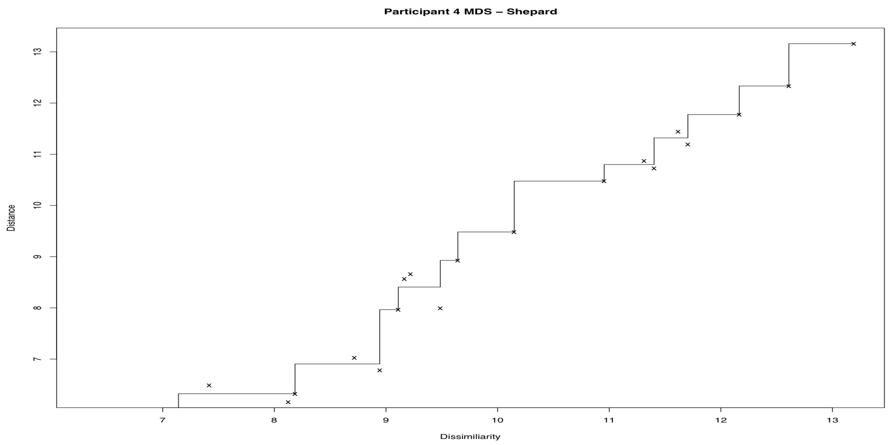


Fig. J-38: *Shepard diagram of the MDS analysis for Participant 4.*

Participant 4 - Cluster Analysis and MDS - Elements

The cluster analysis for Participant 4 is shown in Figure J-39 with a cutoff level of approximately 1 which results in 3 clusters with a single outlier at a level of 0.8. This consisting of a single sound, s14, which is the sound of a sea buoy bell ringing. These were birds & animals, bird calls, and objects. The outlier of s14 is particularly interesting as when shown in

Figure J-39 as it is somewhat related to the sound, s18, which is the sound of a seagull's cry. Cluster 1 contains sounds that were seen as bird or animal like sounds. Cluster 2 contained sounds that were seen as bird or bird like calls. Cluster 3 contained sounds that were seen as object like sounds. One noted exception to this cluster is s23, the sound of a power saw in use, which was found to have identification issues. This is discussed further in Section 5.1.1. An outlier to the two clusters was the sound of a sea buoy bell ringing, s14, as the later causal uncertainty studies (see Section 5.1.1) show, there was no difficulty in its identification so other factors must be involved and this is an area of further study.

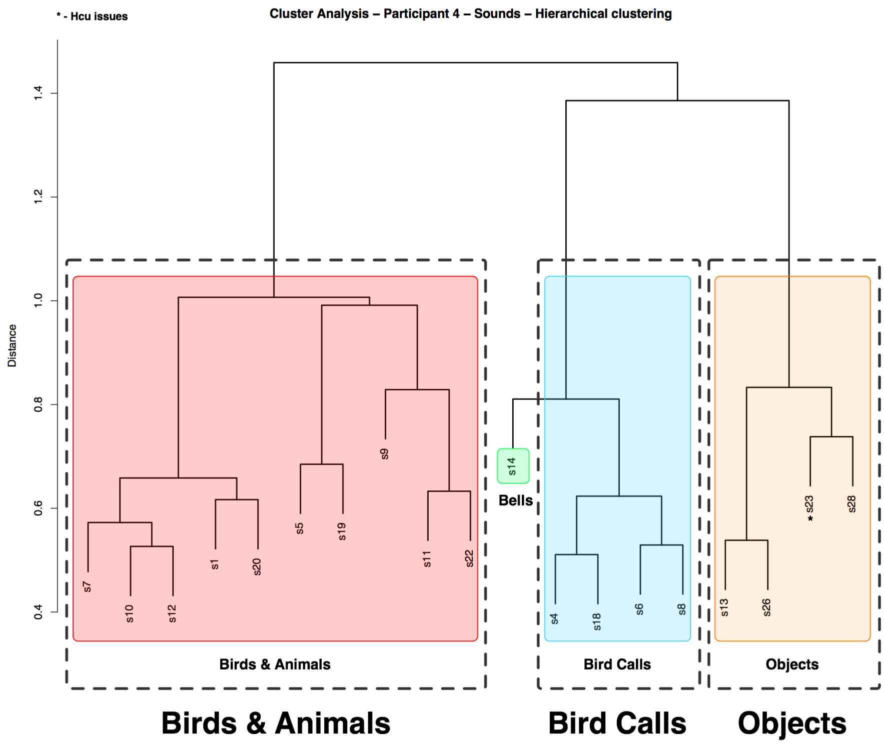


Fig. J-39: *The cluster analysis of the RGT elements (sounds) for Participant 4.*

The MST multidimensional scaling analysis for Participant 1 is shown in Figure J-40, with its related Shepard diagram in Figure J-41. Examining Figure J-30 visually for patterns, we see that it is mostly likely a circumplex pattern as *s19* could not be interpreted to lie under the same line as the elements *s28* or *s7*. This indicates that there are two or more dimensions effecting the participant's determination of elements. This hypothesis is somewhat supported by the earlier findings of the principal component analysis for the participant's elements as shown in Table J-23.

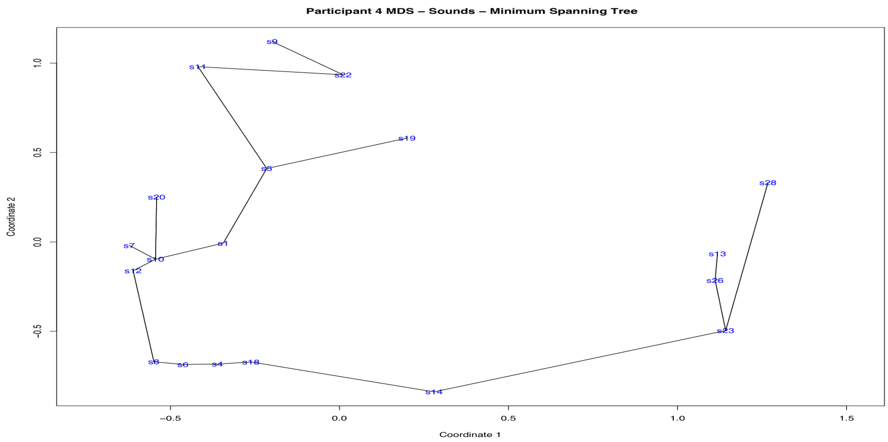


Fig. J-40: *MDS minimum spanning tree analysis of the RGT elements (sounds) for Participant 4.*

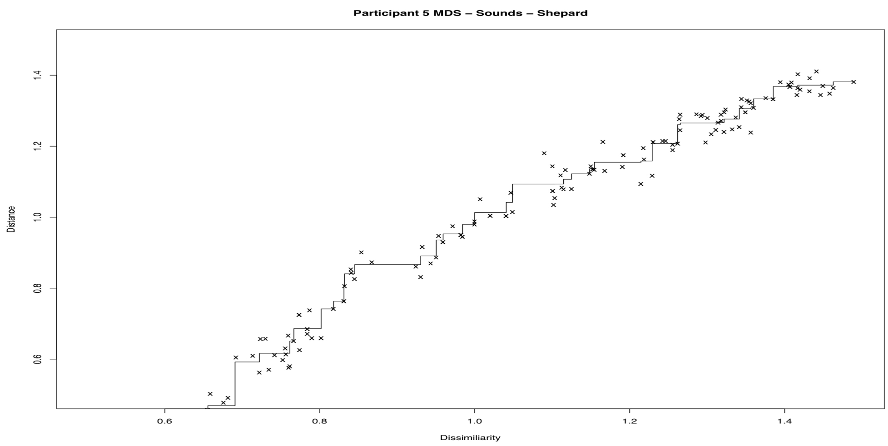


Fig. J-41: *Shepard diagram of the MDS analysis of the RGT elements (sounds) for Participant 4.*

J.1.4 Results and Observations for Participant 4

The results of the analysis of the constructs from Participant 4 from the PCA, CA, and MDS analysis suggest that outdoor sounds versus dark / weird sounds was a factor. The constructs further suggested that after open, relaxing, identifiable sounds and singular, closed sounds was the naturalness of the sound. The results of the analysis of the elements (sounds) from Participant 4 from the PCA, CA, and MDS analysis suggested that the contrast of animal/bird sounds with that of animal/bird calls was a factor. The typical location associated with the sounds was a further factor revealed by this analysis.

Participant 5 - PCA - Constructs

Examining the correlations among the constructs for Participant 5, shown in Table J-33, we can say that *(welcome/greeting/recognition* (construct 1) was definitely not associated or heard as *concentrating* (construct 5, -0.89) or as *dark/dead/lonely* (construct 3, -0.85). As previously stated, we will only consider the first construct and the constructs with the most significant correlations, in order to highlight the points of interest from the participant's data. There was a significant correlation (0.84) between "*dark dead lonely - outdoors life*" and "*concentrating - attention seeking*", which means that the *dark/dead/lonely* sounds are associated with the *concentrating* sounds. The *concentrating – attention seeking* construct is the construct most closely associated with the other constructs as shown in Table J-34. *welcome/greeting/recognition – night/mechanical* is the next most closely associated construct.

	1	**2**	**3**	**4**	**5**	**6**	**7**
1 welcome/greeting/ recognition — night/mechanical	1.00	0.47	**-0.85**	-0.18	**-0.89**	-0.15	0.61
2 comfortable — sheep		1.00	-0.47	0.47	-0.30	0.53	-0.08
3 dark/dead/lonely — outdoors life			1.00	0.15	**0.84**	0.13	-0.49
4 leaving/fading — cut off				1.00	0.37	0.70	-0.41
5 concentrating — attention seeking					1.00	0.34	-0.67
6 relaxing/becalmed — pleading						1.00	-0.53
7 insistent — unfeeling/detached							1.00

Table J-33*:* Correlation Analysis for Participant 5.

The interpretation of the first principal component in Table J-35 contrasts the welcoming, relaxing, open sounds with those where were short, mechanical-like and annoying. The second principal component in Table J-35 contrasts familiar and complete sounds versus sounds which were pleading and incomplete, implying that after the first principal component, the next main source of variation is between the completeness and urgency conveyed by the sounds. The first two components are shown in Figure J-42 and this can help to illustrate the interpretation. The components 2 and 3 and the components 1 and 3 of the principal component analysis are shown in Figure J-43.

Construct	Root-mean-square correlation
welcome/greeting/recognition — night/mechanical	0.67
comfortable — sheep	0.54
dark/dead/lonely — outdoors life	0.65
leaving/fading — cut off	0.54
concentrating — attention seeking	0.69
relaxing/becalmed — pleading	0.56
insistent — unfeeling/detached	0.60
Average of statistic	0.61
Standard deviation of statistic	0.06

Table J-34*:* Root-mean-square (average) correlation among constructs analysis for Participant 5.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
welcome greeting recognition - night/mechanical	-0.50	-	-0.17	(-)
comfortable - sheep	-0.17	(-)	-0.58	-
dark dead lonely - outdoors life	0.47	+	0.24	(+)
leaving fading - cut off	0.20	+	-0.47	-
concentrating - attention seeking	0.51	+	0.05	
relaxing/becalmed - pleading	0.20	+	-0.53	-
insistent - unfeeling detached	-0.41	(-)	0.28	
Standard deviation	2.16		1.48	
Proportion of Variance	0.67		0.31	
Cumulative Proportion	0.67		0.98	

Table J-35*:* Principal–Components Analysis for Participant 5.

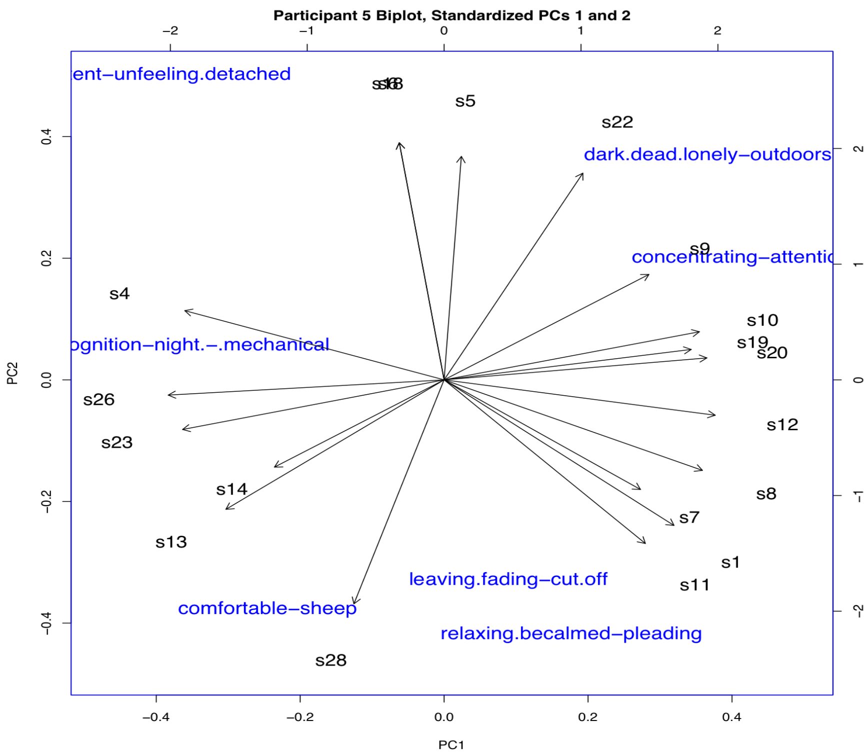


Fig. J-42: *The first two components of the principal–components analysis for Participant 5.*

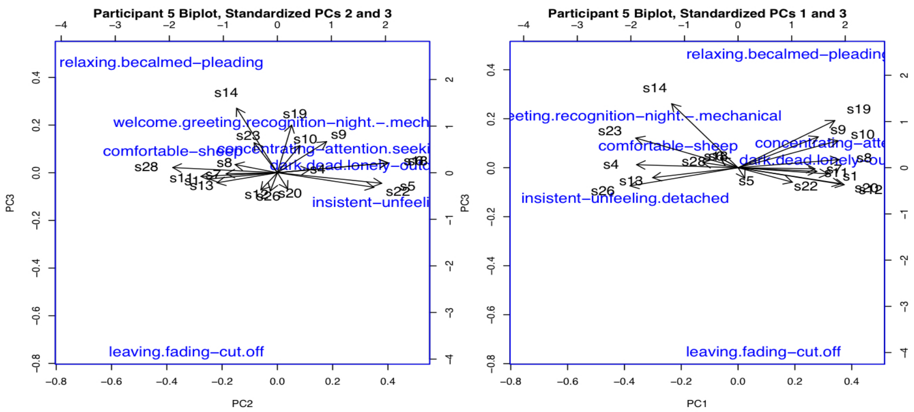


Fig. J-43: *The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 5.*

Participant 5 - PCA - Elements

The results of the elicitation task for Participant 5 are shown below in Table J-36. Participants responded in free-text format to what they thought each sound was as we can see from Table J-36, these text descriptions were often highly descriptive and described the events or actions.

ID	Sound Description	Participant's Description
s1	cat	*annoying, yelpy*
s4	owl	*dark, lonely, echoy*
s5	bird song	*tweet waiting*
s6	bird song	*passing by, languid*
s7	bird song	*insistent compelling*
s8	rooster	*greeting, welcoming, outdoors*
s9	donkey	*creak, fading, departing*
s10	horse	*greeting, indoors*
s11	goat	*cut short, snatched*
s12	sheep	*outdoors, shearing*
s13	glass breaking	*crisp, sharp*
s14	church bell ringing	*water, tidal, night*
s18	seagull	*sea, leaving*
s19	seal	*needy, attentive, seeking*
s20	horse	*natural sound*
s22	lion roaring	*rumble, relaxing*
s23	power saw	*sucky / mechanical*
s26	coins counting	*calculating and mechanical*
s28	heavy ball bouncing	*fall, echoey, dead*

Table J-36*:* Descriptions by Participant 5 for the elements.

Examining the correlations among the constructs for Participant 5, shown in Table J-37, we can say that element *s28* (the sound of a heavy ball bouncing) was definitely associated with *s13 (0.87)*, *s22 (0.95)*, and *s26 (0.93)* (the sound of glass breaking, the sound of a lion roaring, and the sound of coins being counted, respectively) while being definitely not associated with *s4 (-0.85)*, *s6 (-0.85)*, and *s7 (-0.88)* (the sound of an owl hooting, and two versions of bird song, respectively). The *s12* element (the sound of a sheep bleeting) is the element most closely associated with the other elements as shown in Table J-38. The elements *s1*, *s8*, *s23*, and *s26* (the sounds of a cat meowing, the sound of rooster crowing, the sound of a power saw in use, and the sound of coins being counted, respectively) are the next most closely associated elements.

The interpretation of the first principal component in Table J-39 contrasts the animal/bird

	s1	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s18	s19	s20	s22	s23	s26	s28
s1	1.00	0.61	0.56	0.61	0.52	**0.91**	0.77	0.78	0.75	**0.92**	−0.51	−0.44	0.36	0.24	0.78	−0.43	−0.60	−0.29	−0.51
s4		1.00	0.73	**1.00**	**0.97**	**0.85**	0.59	0.43	0.44	0.58	**−0.94**	−0.77	−0.08	0.09	0.43	−0.76	**−0.81**	−0.66	**−0.85**
s5			1.00	0.73	0.57	0.66	0.22	0.18	0.08	0.38	−0.77	−0.18	0.07	0.28	0.18	−0.54	−0.79	−0.31	−0.54
s6				1.00	**0.97**	**0.85**	0.59	0.43	0.44	0.58	**−0.94**	−0.77	−0.08	0.09	0.43	−0.76	**−0.81**	−0.66	**−0.85**
s7					1.00	**0.80**	0.59	0.40	0.47	0.52	**−0.89**	**−0.87**	−0.15	−0.03	0.40	−0.77	−0.72	−0.73	**−0.88**
s8						1.00	0.74	0.65	0.70	**0.83**	−0.73	−0.69	0.10	0.05	0.65	−0.61	−0.69	−0.50	−0.75
s9							1.00	**0.97**	**0.94**	**0.94**	−0.43	−0.56	0.16	0.15	**0.97**	−0.39	−0.45	−0.35	−0.44
s10								1.00	**0.92**	**0.94**	−0.31	−0.36	0.33	0.28	**1.00**	−0.31	−0.42	−0.25	−0.31
s11									1.00	**0.87**	−0.25	−0.48	0.13	−0.07	**0.92**	−0.34	−0.28	−0.34	−0.42
s12										1.00	−0.44	−0.47	0.27	0.25	**0.94**	−0.33	−0.50	−0.22	−0.39
s13											1.00	0.65	−0.16	−0.34	−0.31	**0.88**	**0.94**	0.77	**0.87**
s14												1.00	0.14	0.16	−0.36	0.58	0.43	0.64	0.75
s18													1.00	**0.80**	0.33	−0.35	−0.45	−0.28	−0.14
s19														1.00	0.28	−0.27	−0.57	−0.13	−0.03
s20															1.00	−0.31	−0.42	−0.25	−0.31
s22																1.00	**0.87**	**0.96**	**0.95**
s23																	1.00	0.72	0.79
s26																		1.00	**0.93**
s28																			1.00

Table J-37: Correlation Analysis for Participant 5 Elements.

Element	Root-mean-square correlation
s1	0.62
s4	0.61
s5	0.43
s6	0.46
s7	0.51
s8	0.62
s9	0.54
s10	0.59
s11	0.58
s12	0.63
s13	0.59
s14	0.45
s18	0.46
s19	0.56
s20	0.60
s22	0.52
s23	0.62
s26	0.62
s28	0.50
Average of statistic	0.55
Standard deviation of statistic	0.07

Table J-38*:* Root-mean-square (average) correlation among elements analysis for Participant 5.

sounds with those where were object based. The second principal component in Table J-39 contrasts the distance of the source of the sounds between those near to the listener and those in the distance, implying that after the 'naturalness' of the sources of the sounds is taken into account, the main source of variation is between the perceived distance of the sound by the listener. The first two components are shown in in Figure J-44, the components 2 and 3, and the components 1 and 3 of the principal component analysis are shown in Figure J-45. These figures can help in illustrating the interpretation of the principal components.

	PC1	*PC1 Simplified*	*PC2*	*PC2 Simplified*
s1	−0.24	-	−0.22	-
s4	0.27	+	0.11	(-)
s5	−0.03		0.40	+
s6	0.05		0.40	+
s7	−0.24	-	−0.20	-
s8	−0.26	-	−0.14	(-)
s9	−0.25	-	0.16	(-)
s10	−0.28	-	0.06	
s11	−0.22	-	−0.26	-
s12	−0.28	-	−0.07	
s13	0.25	+	−0.19	(-)
s14	0.25	+	−0.16	(-)
s18	0.05		0.40	+
s19	−0.28	-	0.03	
s20	−0.28	-	0.02	
s22	−0.17	-	0.32	+
s23	0.28	+	−0.06	
s26	0.28	+	−0.00	
s28	0.12	(+)	−0.37	-
Standard deviation	3.54		2.47	
Proportion of Variance	0.66		0.32	
Cumulative Proportion	0.66		0.98	

Table J-39*:* Principal–Components Analysis for Participant 5 Elements.

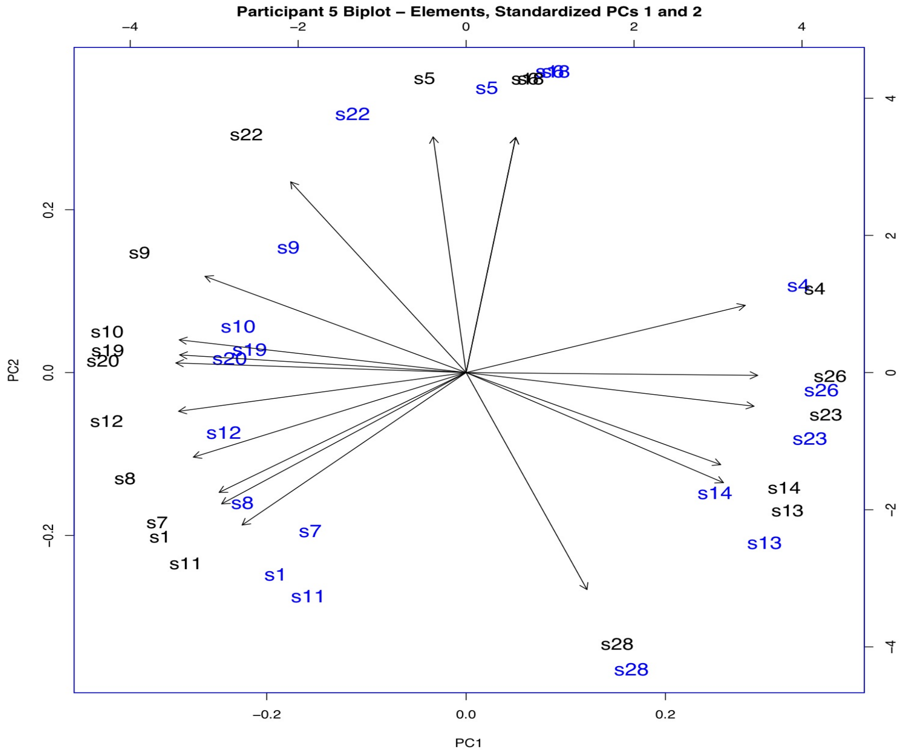


Fig. J-44*: The first two components of the principal–components analysis for Participant 5 Elements.*

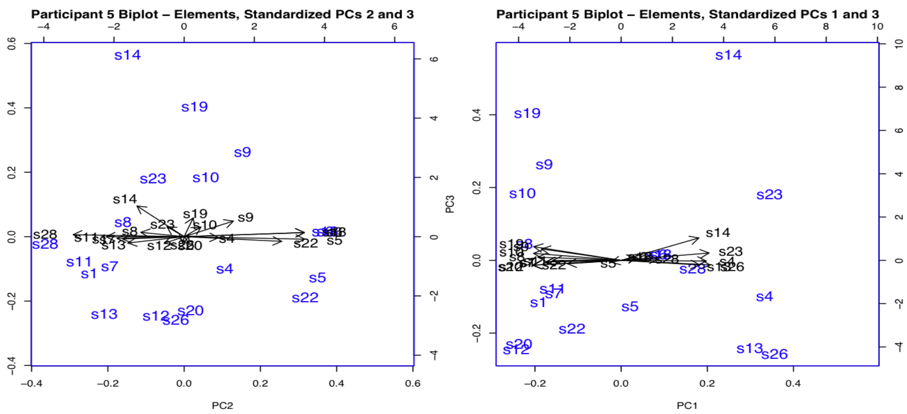


Fig. J-45*: The components PC 2–3 and PC 1–3 of the Principal–Components Analysis for Participant 5 Elements.*

Participant 5 - Cluster Analysis and MDS - Constructs

The cluster analysis for Participant 5 was shown already in Figure J-46 with two possible cutoff levels. These are at 5.35 and at 6.5 and would result in 4 and 7 clusters respectively. The resulting clusters for cutoff level 6.5 are shown in Table J-40. Cluster 4 had the shortest significant distances from the set of clusters and contained sounds that were seen as either dark / dead sounds or alarming / attention grabbing sounds. The shortest distance match in this cluster was 3.6 between ”*dark / dead / lonely – outdoors life*” and ”*concentrating – attention seeking*”. This cluster contained only these two constructs. These constructs suggest that the sounds were distinctly divided into sounds which occurred in outdoors and attention seeking / alarming type sounds.

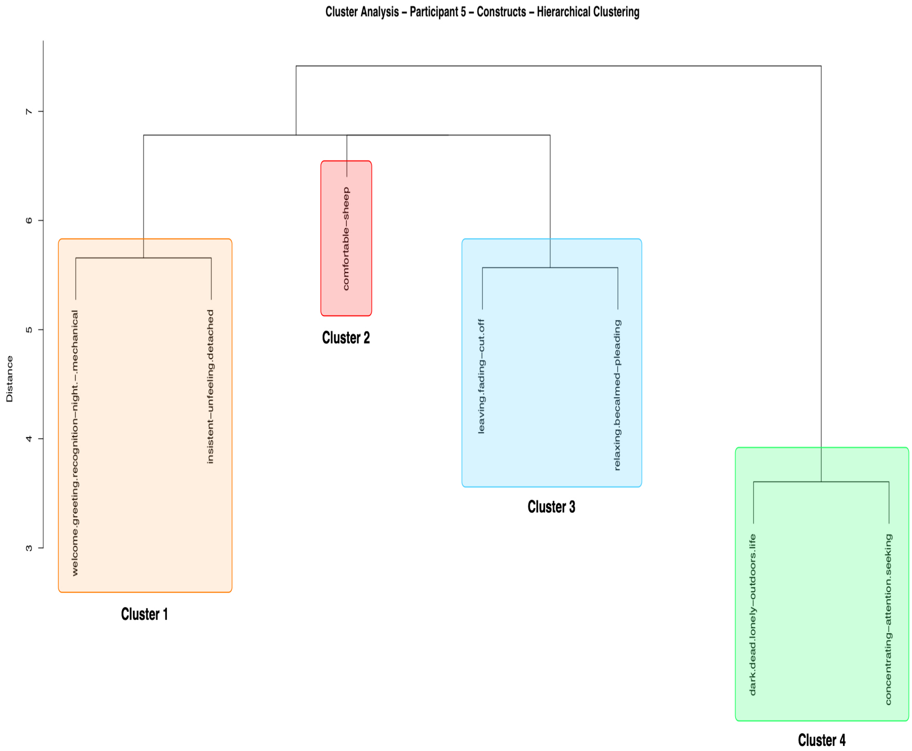


Fig. J-46*: The cluster analysis of the RGT constructs for Participant 5.*

	Constructs *Emergent Pole* - *Implicit Pole*
1	welcome greeting recognition — night / mechanical insistent — unfeeling / detached
2	comfortable — sheep
3	leaving / fading — cut off relaxing / becalmed — pleading
4	dark / dead / lonely — outdoors life concentrating — attention seeking

Table J-40*:* Clusters obtained for Participant 5.

The MST multidimensional scaling analysis for Participant 5 is shown in Figure J-47, with its related Shepard diagram in Figure J-48. Examining Figure J-47 visually for patterns, we see that it is mostly likely a circumplex pattern as all the constructs could be not be interpreted to lie under a single line. This indicates that there was no single dimension which influenced the participant's determination of constructs. The dissimilarity of several of constructs would seem to suggest that there are are two or more dimensions which influenced the participant's constructs.

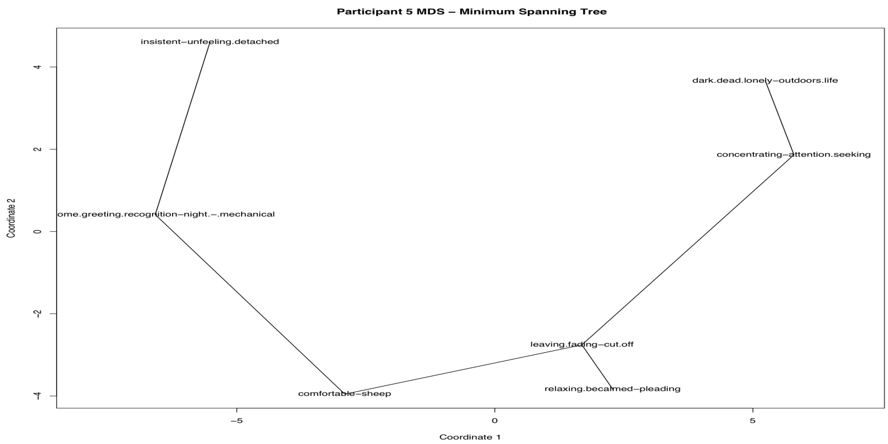


Fig. J-47: *MDS minimum spanning tree analysis of the RGT constructs for Participant 5.*

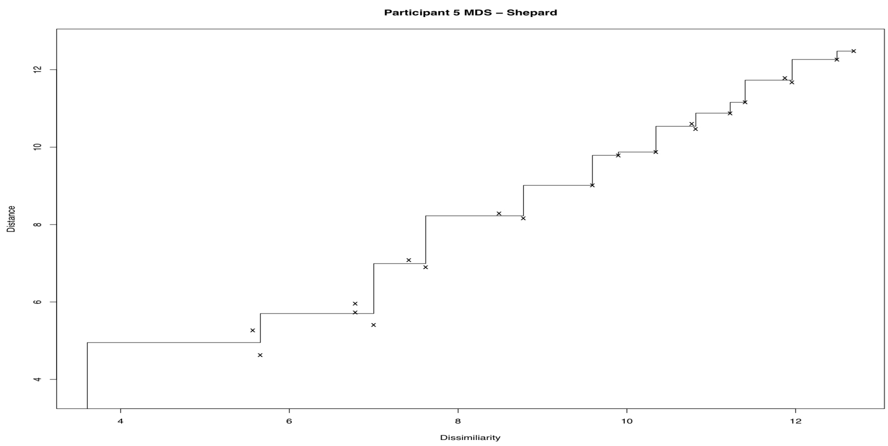


Fig. J-48: *Shepard diagram of the MDS analysis for Participant 5.*

Participant 5 - Cluster Analysis and MDS - Elements

The cluster analysis for Participant 5 is shown in Figure J-49 with a cutoff level of approximately 0.8 which results in 5 clusters. These were continuous impact like sounds, objects impacts, birds, farm animals & birds, and animals. These were clustered into two larger clusters which were either object sounds or bird & animal sounds. Cluster 1 contains sounds that were heard as continuous impact like sounds. One noted exception to this cluster is s23, the sound of a power saw in use, which was found to have identification issues. This is discussed further in Section 5.1.1. Cluster 2 contained sounds that were heard as object like sounds. Cluster 3 contained sounds that were heard as bird like sounds. One noted exception to this cluster is s22, the sound of a lion roaring, which was found to have identification issues and this is examined further in Section 5.1.1. Cluster 4 contained sounds that were heard as farm animals & birds sounds. One noted exception to this cluster is s8, the sound of a rooster crowing, which was found to have identification issues and this is examined further in Section 5.1.1. Cluster 5 contained sounds that were heard as animal sounds.

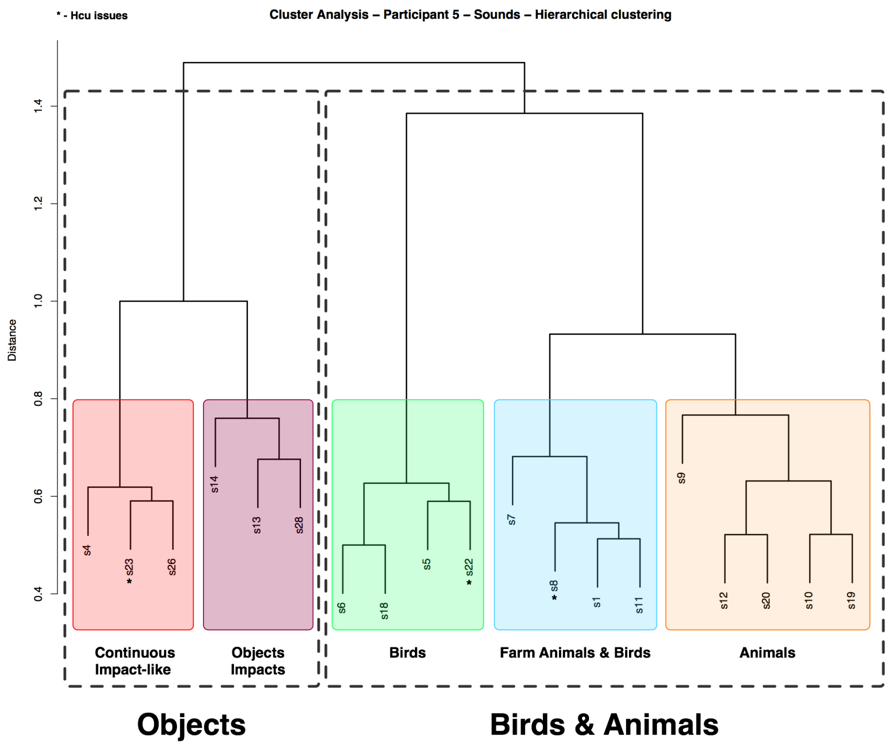


Fig. J-49*: The cluster analysis of the RGT elements (sounds) for Participant 5.*

The MST multidimensional scaling analysis for Participant 5 is shown in Figure J-50, with its related Shepard diagram in Figure J-51. Examining the Shepard plot in Figure J-51, there are several discrepancies between the original dissimilarities (shown in diagram as Xs) and the multidimensional scaling solution (the line in the diagram). This would indicate that the quality of the multidimensional scaling solution for this participant is not particularly good. In the case of this exploratory study the given inexactness in fit of the solution was regarded as acceptable. Examining Figure J-30 visually for patterns, we see that it is mostly likely a circumplex pattern as *s5* could not be interpreted to lie under the same line as the elements *s7* or *s14*. This indicates that there are two or more dimensions effecting the participant's determination of elements. This hypothesis is somewhat supported by the earlier findings of the principal component analysis for the participant's elements as shown in Table J-23.

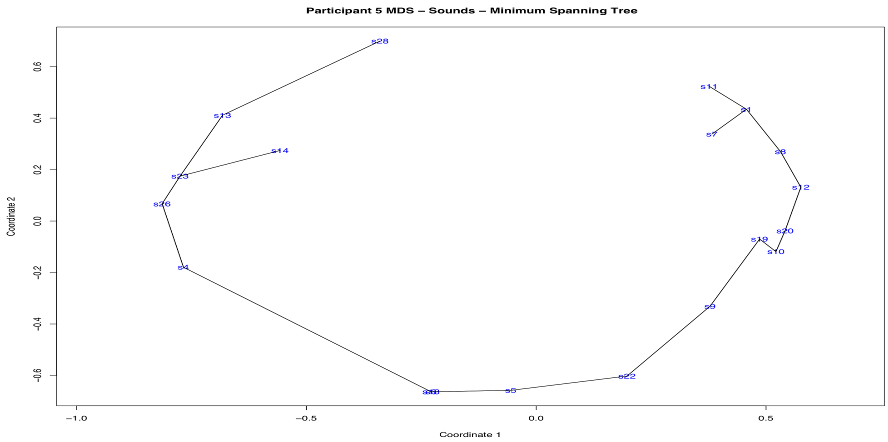


Fig. J-50: *MDS minimum spanning tree analysis of the RGT elements (sounds) for Participant 5.*

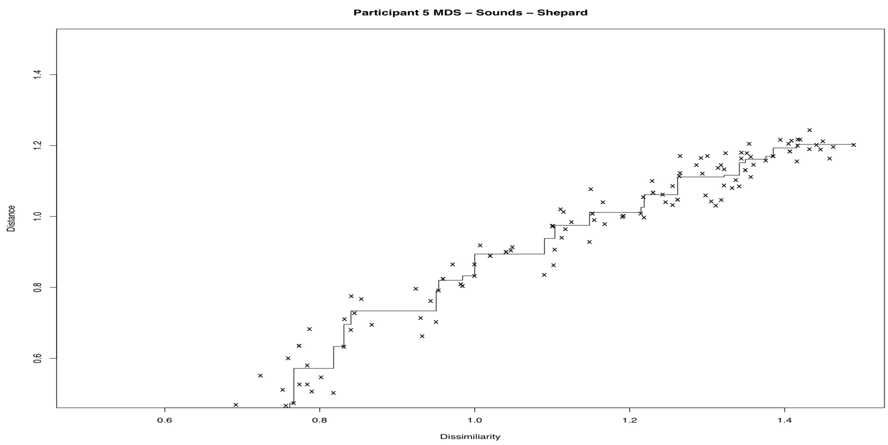


Fig. J-51: *Shepard diagram of the MDS analysis of the RGT elements (sounds) for Participant 5.*

J.1.5 Results and Observations for Participant 5

The results of the analysis of the constructs from Participant 5 from the PCA, CA, and MDS analysis suggest that alarming / alerting sounds versus dead / dark sounds was a factor. The constructs further suggested that the urgency and completeness of the sounds when contrasted to incomplete or pleading sounds was a factor. These incomplete sounds were typically short, mechanical, and found to be annoying by Participant 5. The results of the analysis of the elements (sounds) from Participant 5 from the PCA, CA, and MDS analysis suggested that sounds were split into birds, animals, objects, or impact sounds. The 'naturalness' of the sound's source was a factor. The perceived distance from the sound by the listener was another factor.

K.1 Appendix K - Chapter 6 - Second study - Questionnaire

This appendix holds the questions asked in the survey carried out for the second study presented in Chapter 6. Survey response data can be useful in asking if lifestyle, education, age or other factors may influence how the Auditory Icons are perceived. They can also be used as shown in Chapter 4, to question participants views on aesthetics, ease of use, frustration or other interface specific qualitative questions. These types of questions can be used separately or jointly depending on what is being examined by the study and the survey. This section looks at the questionnaire shown below, which explored participants childhood living environments, current living environments, and their musical training.

To help us clarify if perception is affected by age, gender, your local environment, the environment you grew up in or musical ability, please answer the questions below:

Q1: Do you have any known hearing problems: (Please tick appropriate box)

Yes	☐	**No**	☐

Q2: What is your age: (Please tick appropriate box)

Under 13	☐	**25 – 34**	☐
13 – 18	☐	**35 – 44**	☐
19 – 20	☐	**45 – 54**	☐
21 – 24	☐	**55 and over**	☐

Q3: What is your gender: (Please tick appropriate box)

Male	☐	**Female**	☐

Q4: What type of environment did you grow up in: (Please tick any boxes that apply)

Urban – City or larger	☐	**Near or by Sea**	☐
Urban – Town or smaller	☐	**Near or by River / Lake**	☐
Suburban – City / Town but some rural elements	☐	**Near of by Mountains / Hills**	☐
Rural – Village or smaller	☐	**Near or by Forest / Woods**	☐
Rural - Countryside	☐	**Other:** (Please Specify)	

Q5: What type of environment do you live in now: (Please tick any boxes that apply)

Urban – City or larger	☐	**Near or by Sea**	☐
Urban – Town or smaller	☐	**Near or by River / Lake**	☐
Suburban – City / Town but some rural elements	☐	**Near of by Mountains / Hills**	☐
Rural – Village or smaller	☐	**Near or by Forest / Woods**	☐
Rural - Countryside	☐	**Other:** (Please Specify)	

Continued on the next page

Q6: Have you ever had formal music training: (Please tick appropriate box)

Yes ☐ **No** ☐

Q6b: If the answer was Yes to Q6, how many years of formal music training did you receive: (Please tick appropriate box)

<1 month ☐ **<3 years** ☐
<6 months ☐ **<4 years** ☐
<1 year ☐ **<5 years** ☐
<2 years ☐ **>5 years or more** ☐

Q6c: If the answer was Yes to Q6, what kind of formal music training did you receive: (Please tick any boxes that apply)

Instrumental - Classical ☐ **Instrumental - Traditional** ☐
Voice / Singing ☐ **Suzuki Method** ☐
Music Composition ☐ **Music Theory** ☐
Music by Book (e.g. tablatures) ☐ **Other:** (Please Specify)

K.1.1 Questionnaire Results and Analysis

The tables below show the results of the questionnaire for the second study are shown. Due to the small number of participants, these results did not show any significant results but it is still useful to analyse them as the methods that are applied to them are useful to showcase.

Childhood Area Type

	Hearing Issues	Age	Gender	Urban City	Urban Town	Sub-Urban	Rural Village	Rural	Near Sea	Near River or Lake	Near Hills or Mountains
p1	N	25-34	M			Y					
p2	N	35-44	M			Y					
p3	N	25-34	M				Y	Y		Y	Y
p4	N	25-34	M				Y	Y		Y	Y
p5	N	35-44	F				Y	Y			

Table K-1*:* Questionnaire Results - Part 1 - Childhood Area Type

Current Abode/Home Area Type

	Hearing Issues	Age	Gender	Urban City	Urban Town	Sub-Urban	Rural Village	Rural	Near Sea	Near River or Lake	Near Hills or Mountains
p1	N	25-34	M			Y					
p2	N	35-44	M	Y							
p3	N	25-34	M			Y					Y
p4	N	25-34	M		Y					Y	
p5	N	35-44	F	Y							

Table K-2*:* Questionnaire Results - Part 2 - Current Living Area Type

Types of Musical Training

	Musical Training	Length of Training	Classical Instrument	Voice Training	Composition	Books or Tablatures	Traditional Instrument	Suzuki Method	Musical Theory
p1	Y	$\leq$2 years	Y						
p2	N								
p3	Y	$\leq$2 years					Y		Y
p4	Y	$\leq$2 years					Y		Y
p5	N								

Table K-3*:* Questionnaire Results - Part 3 - Musical Training

The results of the questionnaire are shown in the tables below to provide an overview. Analysing the associations of categorical variables can provide additional information, one method for this type of analysis is through the use of a contingency table (Agresti, 1996) or table of counts. The results of the questionnaire are summarised in two tables below, the first table was used to perform a cross-sectional study of the participants questionnaire data where the link between musical training and the type of environment that the participant grew up in as a child was examined. The second table performs a similar cross-sectional study where the link between musical training and the type of environment that the participant currently lives was examined. In Figure K-1, a graphical representation called a mosaic plot is used to visualise the decompositions of the variables associations. It is interesting to note that for people with a suburban childhood environment learnt classical instruments and music theory while participants from a rural childhood environment learnt traditional instruments. This type of response data from surveys can be useful in asking if variables such as musical training or length of the training and sound likes or dislikes are related and whether the length of musical training may influence these likes or dislikes. The work in this experiment concentrated on exploring the technique and its possibility for use in Auditory Display design and the questionnaire is shown here as another complementary technique. The results of the questionnaire did not return any significant data and further investigation would be required to explore these questions. It is well established that childhood environment has an effect on learning and development (Matheny Jr. et al., 1995, Roberts et al., 1999), how this environment as well as a person's current environment or musical skills could influence their perceptions or descriptions of everyday sounds are topics that deserve more detailed explorations in larger studies.

	Classical Instrument	Voice	Composition	Books or Tablatures	Traditional Instrument	Suzuki Method	Music Theory
Urban City	0	0	0	0	0	0	0
Urban Town	0	0	0	0	0	0	0
Suburban	1	0	0	0	0	0	1
Rural Village	0	1	1	0	2	0	1
Rural Countryside	0	1	1	0	2	0	1
Near Sea	0	0	0	0	0	0	1
Near River or Lake	0	1	1	0	2	0	1
Near Mountains or Hills	0	1	1	0	1	0	0

Table K-4*:* A contingency table exploring the type of musical training and the type of neighbourhood or area where participant grew up in

	Classical Instrument	Voice	Composition	Books or Tablatures	Traditional Instrument	Suzuki Method	Music Theory
Urban City	0	0	0	0	0	0	0
Urban Town	0	1	1	0	0	0	0
Suburban	1	0	0	0	0	0	1
Rural Village	0	0	0	0	2	0	0
Rural Countryside	0	0	0	0	2	0	0
Near Sea	0	0	0	0	0	0	0
Near River or Lake	0	0	0	0	0	0	0
Near Mountains or Hills	0	0	0	0	0	0	0

Table K-5*:* A contingency table exploring the type of musical training and the type of neighbourhood or area where participant is currently living

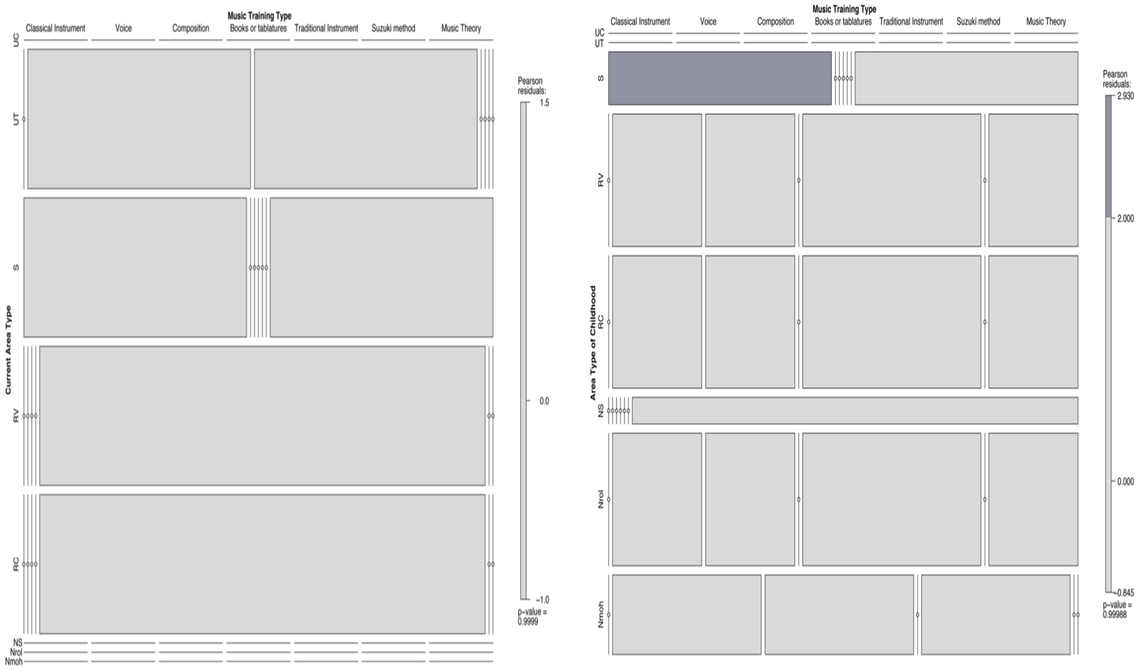


Fig. K-1: *Mosaic plots illustrating the results of the contingency tables*

L.1 Appendix L - Companion DVD

This appendix contains a DVD containing the applications, scripts, Auditory Icons, and data analysis results presented in the experiments from Chapters 4 to 6.

M.1 Appendix M - Publication List

This appendix contains a full listing of publications produced during the period of this dissertation and while a number of these were not directly referenced they contributed to the work within the thesis.

Book Chapters

Brazil, E., Fernström, M. and Ottaviani, L. (2003*a*), *The Sounding Object*, Mondo Estremo, Firenze, Italy, chapter Psychoacoustic validation and cataloguing of sonic objects: 2D browsing, pp. 257–294.

Fernström, M. and Brazil, E. (in press MIT Press, due 2010), *Principles of Sonification and Auditory Display*, T. Hermann, A. Hunt, J. Neuhoff, eds., chapter Auditory Icons, pp. in press.

Journal Papers

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

Master Thesis

Brazil, E. (2003), Investigation of multiple visualisation techniques and dynamic queries in conjunction with direct sonification to support the browsing of audio resources, *Supervisor:* M. Fernström., University of Limerick, 2003.

Refereed Conference or Workshop Publications

Brazil, E., Fernström, J. and Ottaviani, L. (2003*b*), A new experimental technique for gathering similarity ratings for sounds, *in* E. Brazil and B. Shinn-Cunningham, eds, 'International Conference on Auditory Display (ICAD-03)', pp. 238–242

Ottaviani, L., Brazil, E. and Fernström, M. (2003), Psychoacoustic experiments for validating sound objects in a 2-d space using the sonic browser, *in* 'Proceedings of the XIV Colloquium on Musical Informatics (XIV CIM 2003)', Firenze, Italy, pp. 90–94.

Brazil, E. and Fernström, M. (2004), Interactive radio: Exploring visitor stories using a radio interface, *in* S. Barrass and P. Vickers, eds.,'Proceedings of ICAD 2004 - The 10th International Conference on Auditory Display', Sydney, Australia.

Fernström, M. and Brazil, E. and Bannon, L. (2004), An investigation of soft-button widgets using sound, *in* 'Proceedings of 2004 Le Journees de Design Sonore', Paris, France.

Fernström, M. and Brazil, E. (2004), Human-computer interaction design based on interactive sonification – hearing actions or instruments/agents, *in* T. Hermann and A. Hunt, eds., 'Proceedings of the 2004 International Workshop on Interactive Sonification', Bielefeld, Germany

Brazil, E. and Fernström, M. (2006), Investigating concurrent Auditory Icon recognition, *in* 'Proceedings of ICAD 2006 - The 12th International Conference on Auditory Display', Queen Mary, London., pp. 51–58.

Brazil, E. and Fernström, M. (2007), Investigating ambient auditory information systems, *in* G. P. Scavone, ed., 'International Conference on Auditory Display (ICAD-07)', Montreal, Canada, pp. 326–333.

Brazil, E. and O'Callaghan, T. and Fernström, M. and McLoughlin, M. (2008), Where does usability fit in an industrial academic research programme, *in* 'Proceedings of IHCI 2008 - The Second Irish HCI Conference', Cork, Ireland, pp. 11–15.

Hermann, T. and Willamson, J. and Murray-Smith, R. and Visell, Y. and Brazil, E. (2008), Sonification for sonic interaction design, *in* 'CHI-08 Workshop on Sonic Interaction Design: Sound, Information, and Experience', Florence, Italy, pp. 35–40.

Brazil, E. and Fernström, M. (2009), Subjective experience methods for early conceptual design of Auditory Displays, *in* 'Proceedings of ICAD 2009 - The 15th International Conference on Auditory Display', Copenhagen, Denmark, pp. 11–18.

Fernström, M. and Brazil, E. (2009), The Shannon Portal: Designing an Auditory Display for casual users in a public environment, *in* 'Proceedings of ICAD 2009 - The 15th International Conference on Auditory Display', Copenhagen, Denmark, pp. 27–30.

Brazil, E. and Fernström, M. and Bowers, J. (2009), Exploring concurrent Auditory Icon recognition, *in* 'Proceedings of ICAD 2009 - The 15th International Conference on Auditory Display', Copenhagen, Denmark, pp. 56–59.

Brazil, E. and Fernström, M. (2009), Empirically based Auditory Display design, *in* 'Proceedings of SMC 2009 - The 6th Sound and Music Computing Conference', Porto, Portugal, pp. 7–12.

Non-Refereed Conference, Seminar or Workshop Abstracts

Brazil, E. (2005), Auditory Displays & Public Spaces: Informing the design of Auditory Icons based on listening test analysis, *in* 'CONVIVIO Workshop Understanding Public Spaces: Towards a Methodology', Killaloe, Ireland.

Brazil, E. (2008), Echoes, Whispers, and Footsteps from the Conflux of Sonic Interaction Design and of Public Spaces, *in* 'Aural City Symposium', Berlin, Germany.