Using Sound to Represent Uncertainty in Spatial Data

by

Nicholas Edward Bearman

Thesis submitted for the degree of Doctor of Philosophy at the

University of East Anglia, Norwich

School of Environmental Sciences

January 2013

Abstract

There is a limit to the amount of spatial data that can be shown visually in an effective manner, particularly when the data sets are extensive or complex. Using sound to represent some of these data (sonification) is a way of avoiding visual overload. This thesis creates a conceptual model showing how sonification can be used to represent spatial data and evaluates a number of elements within the conceptual model. These are examined in three different case studies to assess the effectiveness of the sonifications.

Current methods of using sonification to represent spatial data have been restricted by the technology available and have had very limited user testing. While existing research shows that sonification can be done, it does not show whether it is an effective and useful method of representing spatial data to the end user. A number of prototypes show how spatial data can be sonified, but only a small handful of these have performed any user testing beyond the authors' immediate colleagues (where n > 4). This thesis creates and evaluates sonification prototypes, which represent uncertainty using three different case studies of spatial data. Each case study is evaluated by a significant user group (between 45 and 71 individuals) who completed a task based evaluation with the sonification tool, as well as reporting qualitatively their views on the effectiveness and usefulness of the sonification method.

For all three case studies, using sound to reinforce information shown visually results in more effective performance from the majority of the participants than traditional visual methods. Participants who were familiar with the dataset were much more effective at using the sonification than those who were not and an interactive sonification which requires significant involvement from the user was much more effective than a static sonification, which did not provide significant user engagement. Using sounds with a clear and easily understood scale (such as piano notes) was important to achieve an effective sonification. These findings are used to improve the conceptual model developed earlier in this thesis and highlight areas for future research.

List of Contents

List of Figures

List of Tables

List of Acronyms & Glossary

AL2 (Address Layer 2) – *A dataset containing gecoded postal addresses available for purchase by companies and local government organisations. Forms a part of Ordnance Survey MasterMap data products. Used in CS1, see Chapter 4 for details.*

API (Application Programming Interface) - *An option that allows direct access to a program's code. Often used to allow add-ons to make use of existing tools, such as the Google Maps API used in this research.*

CS1 (Case Study 1) – *The first case study of the thesis, evaluating how sound could be used to represent positional location accuracy in the Address Layer 2 data set. See Chapter 4 for details.*

CS2 (Case Study 2) – *The second case study of the thesis, evaluating how sound could be used to represent uncertainty in future climate predictions, using the UKCP09 data set. See Chapter 5 for details.*

CS3 (Case Study 3) – *The third case study of the thesis, evaluating how sound could be used to represent distance in a landscape visualisation setting. See Chapter 6 for details.*

GI (Geographic Information) – *Any information that has a spatial location associated with it.*

GMAPI (Google Maps API) – *The API that allows the Google Maps interface to be customised, used in CS2.*

ICAD (International Community for Auditory Display / International Conference on Auditory Display) – *A research community that focuses on auditory display and the surrounding areas of perception, technology and applications. An annual conference is run to showcase these ideas.*

INSPIRE (Infrastructure for Spatial Information in the European Community) – *An EU led strategy to provide easier access to spatial information for the member countries.*

JJA (June, July & August) – *The period for the 'summer' data in UKCP09.*

MIDI (Musical Instrument Digital Interface) – *An interface that allows different computer based musical instruments to communicate to each other. Relevant for this research because it is used to generate the sounds used in the evaluations.*

OS (Ordnance Survey) – *National mapping agency of Great Britain.*

PQA (Positional Quality Accuracy) – *Status flag included as a part of Address Layer 2. Indicates how accurate the gecoded location is, with 5 different categories.*

Sonification – *The term used to refer to non-speech sounds used to represent data. See Chapter 2 for more details.*

Staccato (also known as accented) – *A term in music referring to non-held notes, just a press and release of the key on the piano. Used in CS1 and CS3, opposite to sustained.*

Sustained – *A term in music referring to a mechanism used to hold a note, such as the sustain pedal on a piano. This keeps the note playing until the pedal is released. Sustained notes are used in CS2, in comparison to staccato notes used in the other case studies.*

UKCIP (United Kingdom Climate Impacts Programme) – *Organisation run to create and promote use of climate change predictions in the UK. It also helps businesses create adaptation strategies for managing impacts of climate change. It was initially funded by DEFRA for these activities, but is now independently funded.*

UKCIP02, UKCIP98, CCIRG96 & CCIRG91 – *Previous iterations of climate projections for the UK. See Table 10 (page 133) for details.*

UKCP09 (United Kingdom Climate Predictions 2009) – *Data promoted by UKCIP, containing predicted future climate data for the UK for the 21st century. For more details see Chapter 5.*

UEA (University of East Anglia) – *University where the research was conducted, as well as providing a number of volunteers for the evaluation sessions.*

Acknowledgements

This PhD was funded by an interdisciplinary studentship from ESRC/NERC, with additional funding from Ordnance Survey. It has provided me with a wealth of experience, both professional and personal. I would like to thank Jenny Harding and Glen Hart at Ordnance Survey for helping me to refine my project ideas and provide a supply of willing volunteers for me to use in my evaluations.

I would also like to thank all of the volunteers who have taken part in my evaluations, including participants from Ordnance Survey, UKCIP and various local authority organisations within Norfolk. This is as well as those at UEA who have volunteered for the evaluations and those who helped me to develop the tools through the pilot stage, particularly when it was in a very rough and ready state.

I have also received much advice, time and assistance from many colleagues at UEA, both those who have finished their PhDs as well as those still working hard on them (most of the time!). Particularly Steve and Sue, whose lunchtime discussions always proved interesting and occasionally proved useful. Also to everyone at UEA using GIS, thank you for reminding me I am not alone when ArcGIS decides it doesn't like what I am attempting to do with it. I am grateful to all the staff I have come across in my non-academic work at UEA for helping me with my pestering and annoying questions. I would also like to thank my new colleagues at the European Centre for Environment and Human Health within the University of Exeter Medical School for their support in allowing me to finish my PhD.

My supervisor, Andrew, has been a real help with this research and it would be much less respectable without his input. His comments have always been useful even if I may not have wanted to hear them.

I also want to thank Louise, my wife, who has helped me balance the numerous aspects of my home and work life right up to the submission of this thesis, and provided me with some much needed rest and relaxation. Life would be very different without you.

List of Related Publications

The findings from this thesis have been published in a number of different forms. The journal papers and papers in conference proceedings or presentations are listed below:

Chapter 4 – Case Study 1: Using Sound to Represent Positional Accuracy of Address Locations

- Bearman, N. & Lovett, A. (2010) Using Sound to Represent Positional Accuracy of Address Locations, **The Cartographic Journal**, Volume 47, Number 4, pp. 308-314(7).
  This paper was written by Nick Bearman. Andrew Lovett contributed to the methodological design and reviewed drafts of the manuscript.

- Bearman, N. & Lovett, A. (2010) Using Sound to Represent Uncertainty in Address Locations, Presented 16/04/2010 at **GISRUK 2010** Conference, UCL, London, UK.
- Bearman, N. & Lovett, A. (2010) Using Sound to Represent Positional Accuracy of Address Locations, Presented 23/07/2010 at **Accuracy 2010** Conference, University of Leicester, Leicester, UK.
  Both conference presentations were written and presented by Nick Bearman.

Chapter 5 – Case Study 2: Using Sound to Represent Uncertainty in UKCP09

- Bearman, N. & Appleton, K. (2012) Using Google Maps to collect spatial responses in a survey environment, **Area**, Volume 44, Number 2, pp. 160-169(10).
- Bearman, N. & Appleton, K. (2012) Finding your way with Google Maps, **GeoConnexionUK**, Volume 10, Number 3, pp. 54-55(2).
  This paper and trade journal article focuses on the methodology of using the Google Maps API. They were written by Nick Bearman and Katy Appleton, with Nick as the lead author. Katy contributed her case study to the paper and advised on the drafts.

- Bearman, N., Jones, P. & Lovett, A. (submitted) Using Sound to Represent Uncertainty in UKCP09 Data with Google Maps API, **Atmospheric Science Letters**.
  Nick wrote this paper, with comments on the drafts and revisions from Phil Jones and Andrew Lovett.

- Bearman, N. & Appleton, K. (2011) Using Google Maps to collect spatial responses in a survey environment, Presented 28/04/2011 at **GISRUK 2011** Conference, University of Portsmouth, UK.
  Nick Bearman and Katy Appleton presented jointly, with Katy covering her case study and issues associated with it. Nick presented the rest of the material.

- Bearman, N. (2011) Using Sound to Represent Uncertainty in Future Climate Predictions for the UK, Presented 23/06/2011 at The 17th International Conference on Auditory Display (**ICAD-2011**), Budapest, Hungary.
- Bearman, N., Jones, P. & Lovett, A. (2011) Using Sound to Represent Uncertainty in UK Climate Predictions 2009 Data: Communication of Uncertainty, From Climate Research to Climate Services workshop, Presented 29/06/2011 at **Royal Meteorological Society Conference**, Exeter, UK.
  The ICAD and Royal Met Society conference presentations were written and presented by Nick Bearman.

Chapter 6 – Case Study 3: Landscape Visualisation

- Bearman, N. & Lovett, A. (2012) Can Sound Help Represent Distance in a Landscape Visualisation Setting?, Presented Poster 11-13/02/2012 at **GISRUK 2012**, Lancaster University, UK.
  Nick Bearman wrote and presented the research, with comments from Andrew Lovett on the draft of the paper.

Chapter 1. Introduction

1.1 How do Humans Understand the World?

As human beings the vast majority of our existence is directly reliant on interacting with the world around us and being able to navigate between different locations. Our senses enable us to experience the world through sight, sound, taste, touch and smell, and allow us to know where we are. Knowing where we are, where we want to go and how to get there is crucial at a variety of scales, whether it is from the bedroom to the kitchen, home to the local shop, home to work, home to another country or navigate through the structure of a book or report. Our senses provide a wide variety of input regarding our location and some of our senses are much more effective at providing navigation and location information than others.

Vision is often thought to be the dominant sense, with estimates of visual input around 90% of total cognitive input, compared to 10% for all of the other senses (Bujacz and Strumiłło, 2007). Sight is able to provide both a broad overview of a scene as well as focus in on one specific area. It can detect changes over time (for example in an animation), but it is difficult to do this over a large area, particularly if the information is densely packed. Sight is also a very good means of exploring data sets, whether the data is represented as tables, graphs or other visual formats.

Sound is the second most powerful sense after vision (Fortin et al., 2007). The auditory system is very good at hearing changes over time and even relatively subtle changes can be picked out very easily. The direction of a sound relative to the listener is easy to detect in the real world, for example whether sounds are coming from in front of us, behind us, to one side or the other, but there are limits on what can be easily replicated in an artificial environment. Left and right panning is easily achieved (with stereo headphones or stereo speakers) but forwards/backwards differentiation is difficult without a multiple speaker surround sound setup or specific information on the user. Sound is most frequently used to interpret speech or to understand alerts with a specific task (e.g. an alarm clock or a phone ringing). However, it can also be used to communicate data with specific sounds (e.g. piano notes) representing specific data (e.g. mean annual temperature). The use of sound to represent data is termed sonification and how sonification can be used to represent spatial data is the main focus of this thesis.

Touch (or haptic) is fundamental to our interaction with objects in the world and usually relates to experiences at the human scale. Touch is also often used to represent location as a substitute for vision, for example in tactile maps. These maps can be very effective, but are best used in combination with another sense, often audio (Jacobson and Kitchin, 1997). A tactile map is an intuitive representation of space, but is limited because only a small number of areas can be touched at once (unlike vision where you can get an overall view) so this method requires the user to build up a mental image in their head of the spatial representation. This self-built mental image a user creates is unlikely to be the most effective method of achieving any navigation or orientation aim as the information input phase is relatively lengthy.

Smell and taste are more difficult to comprehend on the same level as vision, sound and touch because they are much more personal and less easy to fit in to a numerical scientific framework. Sound and vision are easily experienced remotely from the object creating them (e.g. film or television). Haptic is less easy to implement, but has achieved wider exposure to users through gaming and force feedback devices. 'Smell-o-vision' has had brief periods of exploration since the 1950s, but a variety of technical and cost issues have limited its long term impact. Taste is still restricted to the immediate environment; there is no way of transmitting taste over long distances, without actually moving the food itself. Smell is very important for triggering responses because of its close links to memory in the brain (Yeshurun et al., 2009; Dowdey, 2012) and therefore if it could be technically handled, would create a very powerful immersive mechanism. However, there is no current way of easily exploiting smell in a sufficiently controllable manner (Brewster et al., 2006).

Whilst our senses are important for us to gain spatial information about the world around us; how we perceive, understand and recall this information (cognitive mapping) is also important. A wide range of studies have been conducted on how people perceive spatial information, demonstrating that it varies depending on how the information is communicated (vision, sonic, haptic, etc.), people's experiences and abilities as well as a range of other factors (Kitchin, 1994; Bonebright and Flowers, 2011). The understanding of spatial data and spatial concepts is important for humans, both directly and indirectly. We use spatial concepts directly, to be able to navigate to different places and identify where we are. We also often use spatial concepts indirectly in metaphors to help explain practical and abstract ideas (e.g. directions and atomic structures) (Mark, 1993).

The use of geospatial data is also important for a wide range of business and government uses and has increased over recent history as such data have become more easily available, primarily through the Internet. A widely quoted statement when explaining the importance of geospatial data is that around 80% of the information used for decision making is geospatial. The details of this quote have disappeared into the mists of GIS history (though see Haklay, 2010 for an interesting review), but it shows how important spatial data are for a wide variety of uses and users.

In the GIS field the terms spatial and geospatial (or sometimes geographical) knowledge are used frequently, sometimes interchangeably. For this thesis, spatial is used to mean any reference concerning an object in space, its dimensions, size, location and so on. Geospatial is a specific subset of spatial, referring to objects at the scale of ranging from a few meters to the size of the planet. Examples of geospatial objects include houses, people, roads, rivers and oceans, examples of spatial (but not geospatial) objects include molecules, atoms, Computer Aided Design applications (although they do have some overlap with geospatial), astronomy and the locations of stars in the galaxy and universe. Many of the cognitive processes discussed apply to both spatial and geospatial data as the difference is only scale. In general, the word spatial will be used unless there is a specific need to use geospatial or to differentiate between the two.

1.2 How do Humans Understand Spatial Data?

It is necessary for people to understand spatial data for it to be of any use. Many studies have investigated cognitive mapping, assessing how people receive, encode and subsequently access spatial data. While typically geographers may have been the primary users of maps and spatial data, they were not the only ones. Maps were often used by people in sciences, engineering and humanities as a valuable way to express ideas. Those who created maps must remember this diversity of uses and users (Robinson, 1995). In more recent years the variety of people using maps and spatial data has increased dramatically and therefore the 'typical map user' now has a much wider background than even just 10 years ago. There has also been a change in focus to understanding cartography from the point of view of the users of the map, each of whom may have a different interpretation of the map that they are using (Kitchin et al., 2012).

Understanding of spatial data varies between users and it is important for the user to gain the maximum benefit from a set of spatial data. The communication of spatial data will be improved by understanding the cognitive processes that underpin the understanding of spatial data. It may also provide new insights into more efficient storage of spatial data and may allow easier applications of new technology to communicating spatial data (Kitchin, 1994). Understanding the cognitive processes will also allow us to design geographic information (GI) interfaces in a more effective way, with default settings that provide the best experience for the majority of users but allow flexibility when required. It will also allow the interface to be customised for specific groups of users (Montello and Freundschuh, 2005), which will allow them to improve their performance with the interface.

If the cognitive processes are not understood well, or not reflected well in the tools used, it can result in the user spending more time dealing with the technical issues associated with the tools they are using than thinking about the real world phenomena they are dealing with. Additionally, if users misunderstand the data they are using (possibly because of a mismatch in the cognitive process and the software design) they can end up misusing the data (Kitchin, 1994), which can potentially have some very serious impacts (Joslyn et al., 2009).

When GIS data were only available to experts, a significant level of knowledge was required to make use of this information. Geographic information is now much more widely accessible than 10-20 years ago because of the developments in GIS technology, the Internet and the increasing availability of data through Open Data initiatives. This has resulted in experts in many other fields wanting to make use of GIS for their work. GIS has traditionally been technical and can be difficult to use correctly without a technical and/or geographical background. A new approach called IIGLU (Interactive, Integrated, Geographic Learning and Understanding) has been proposed (Weber et al., 2012), which focuses on the geographical concepts rather than the technical or theoretical aspects. It works very much on a case study basis, asking questions such as 'which of these data sets is the most appropriate for this work?', teaching issues such as uncertainty and scale through examples. It is hoped this approach will allow more users to make use of GI while still retaining the knowledge required for professional GIS use.

Understanding the cognitive processes that underpin spatial information is important, but it is also very complex to understand because we cannot see into someone's mind directly to see how they structure spatial information; unfortunately it is not like a mechanical tool that we can open the cover and see the mechanisms underpinning the process. Developments in EEG (electroencephalography) and MRI (magnetic resonance imaging) technologies allow insights into the activity levels of different sections of the brain when participants are shown different stimuli. There is potential to use this technique to gain new insights into the processes undertaken during interpretation of geospatial data; however no such research appears to have been conducted to date (Lachaux et al., 2012). Instead we have to ask participants to complete a series of evaluations with spatial data structured in different ways and make deductions from these results.

When considering the acquisition of spatial data, the first stage is to consider the different types of knowledge and the second is to consider how these types of knowledge may be stored in the brain. There are two different ways of categorising spatial knowledge, firstly by the nature of the knowledge itself and secondly by the source of that spatial knowledge (Mark, 1993). When categorising by the nature of spatial knowledge, three different levels are commonly used:

1. Declarative Geographic Knowledge
   Also known as geographic facts, this is declarative knowledge of facts about locations, size, capital cities and so on. Examples include 'Canberra is the capital of Australia', 'The

population of the City of Norwich is around 130,000', 'Snowdon is in Wales' and so on. This level of knowledge does not include the ability to locate these places on a map.

2. Procedural Geographic Knowledge

   Basic geographic knowledge that enables navigation through a familiar environment, with heavy use of cues to act as reminders from a previous journey, usually from an ego-centric point of view.

3. Configurational Geographic Knowledge

   'Map-like' knowledge of space, often has (or approximates) a Euclidean geometry. This includes purely topological knowledge, as well as building up to scale, relative and absolute distance and direction.

These are hierarchical in nature and can be built up from the bottom (1 -> 2 -> 3) or learnt at the top level (3) and forgotten in stages (3 -> 2 -> 1). Configurational knowledge is very dependent on the presence of maps and is thought not to be possible to learn without the concept of maps (Mark and McGranagham, 1986).

The second approach is classification by source (Mark, 1993). Again this is hierarchical in nature with the later stages building on the earlier ones. The first stage is haptic spaces, which are learnt through direct exposure and experience, for example walking around a room or local area. The second is pictorial spaces, originally interpreted by the user visually, either through direct experience or through a plan view map. The third is transperceptual spaces, which are composed or assembled in the mind through a series of experiences over time. This also allows transfiguration between haptic spaces and pictorial spaces, relating where you are and what you can see to features on a map.

While these are two different approaches to categorise the different types of spatial knowledge, they are not mutually exclusive. They can be used in conjunction with each other to achieve a greater understanding of spatial knowledge. Piaget considers spatial knowledge in terms of how it is gained and understood (Hart and Moore, 1973). He proposes that it comes from the exploration of the physical environment and therefore this type of interactive physical exploration of the real world is more powerful and direct than exploration of a static cartographic map (Hart and Moore, 1973; Parsons, 1995). The development of spatial knowledge develops in line with a child's physical growth, not reaching full development until adolescence (Hart and Moore, 1973).

A different approach is to consider which particular spatial elements we store in our brains and how they might be structured. Three examples are considered here, the concept of spatial links, hierarchical storage of spatial data and the impact of perception of distance.

The spatial concept of links between places (such as a road between two towns or a railway line stopping at several places) is thought to be stored in our mental representations of space and objects with the theory known as the anchor-point hypothesis (Couclelis et al., 1987). Specific points (e.g. stations on a railway line) act as objects that people can then link with other spatial objects. This is also shown in the conception of space from use of the London Underground, with users often knowing a lot about specific places in London, but not being able to link them together because they mainly travel via the tube. This is because spatial knowledge gained from travel through an environment (e.g. walking or getting a bus along the surface) is different to that gained from travel along the London Underground, a network with no external contextual information (Thorndyke and Hayes-Roth, 1982; Guo, 2011). The concept of anchor-points can happen at a variety of scales and may also work with distributions of data, with specific points in a data set being anchored at known points, and the locations in-between being scales between the two anchor points (Couclelis et al., 1987).

The second example of knowledge storage discussed here is a hierarchical method that stores spatial knowledge in different levels; for example information about towns and cities is stored in a different layer to information about states. Participants in the US were asked whether San Diego (California) was east or west of Reno (Nevada). The state of Nevada is generally east of California and the majority of the respondents said Reno was east of San Diego, even though the correct answer was west. It is because the spatial data is sorted hierarchically that it is easy to know Nevada is generally east of California (they share a border) and it is known that Reno is in Nevada and San Diego in California, therefore it is thought that Reno is east of San Diego because of the location of the states (Stevens and Coupe, 1978). Stevens and Coupe observe that spatial data is typically stored hierarchically (as shown using a different example in Figure 1) and while these relationships are stored at some levels (e.g. Ohio is east of Colorado) they are not stored at all levels (e.g. the relationship between Denver and Cincinnati). Therefore, the location of Denver relative to Cincinnati has to be deduced from the information stored at the state level, in this case that Cincinnati is east of Denver. In the vast majority of cases this is correct – it is only in a few cases where the user comes up with an incorrect answer based on this method, such as the California example above.

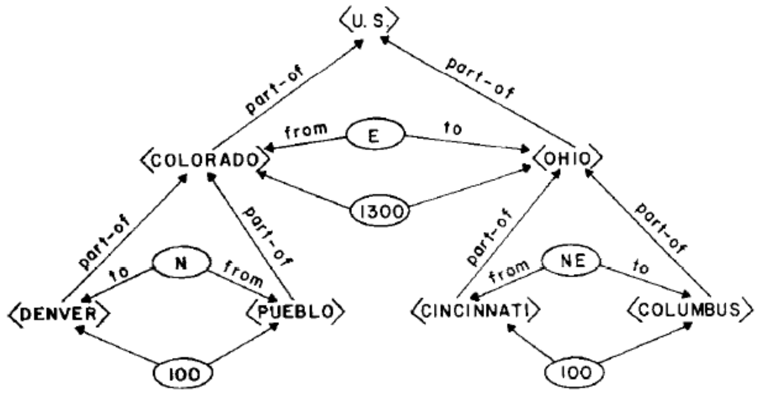


Figure 1. A hierarchical method of storing spatial information (Stevens and Coupe, 1978, fig. 6).

An additional observation related to the hierarchical approach is the impact of barriers on distance estimation. Distance estimation between two points separated by a barrier is generally higher than estimation of the same distance if the barrier is removed (Newcombe and Liben, 1982). In addition to barriers, map projection can effect perception of distance. For example, a study by Gilmartin and Lloyd (1991) showed that participants were less concerned about events happening in cities further away from their current location than events happening in cities closer to their current location. They were also shown a series of different map projections (with the same pair of cities highlighted on them) and they showed less concern about events in a city that appeared to be further away than events in a city that appeared closer to their current location, even when this apparent change in distance was just because of the change in map projection.

As well as considering how spatial knowledge is stored in the brain, work has been conducted looking at the process of learning spatial data and the representation of it as a whole in the brain. Piaget recognised the concept of an internal 'cognitive map' as a representation of the real world inside the mind. This is something built through an individual's exposure to different locations and therefore a direct consequence of their experiences (Hart and Moore, 1973; Parsons, 1995) and as a result, these vary between people. If the methods people use to store spatial information in this 'cognitive map' or 'cognitive representation' of the world are known, this can be used to inform representation of spatial data, which will reduce the cognitive load on the user. Kitchin (1994) reviews the development and use of the term 'cognitive map' and the following paragraphs provide an overview of key ideas. For more detailed information, the reader is referred to Kitchin (1994).

The term 'cognitive map' is used to conceptualise and discuss how people store spatial knowledge and it has had a number of different definitions, both varying over time and depending on who is defining or using the term. Unfortunately, this variation has muddied the water somewhat, with definitions of a cognitive map varying from being an actual cartographic map, to being likened to a cartographic map, to being used as if it were a cartographic map, to being a hypothetical construct used to conceptualise and discuss how people store spatial knowledge. While the multidisciplinary nature of the area initially led to differing views of what the term cognitive map meant, the different views are now becoming more integrated. In current work, a cognitive map is considered to be a hypothetical construct that is used to refer to and discuss spatial knowledge, rather than an object or process in itself (Kitchin, 1994). More generally, a similar convergence of cognitive science, behavioural geography and cartography has occurred to work collectively on the theoretical foundations for GI systems (Mark et al., 1999).

It is important to understand how such spatial information is processed and stored as it will help us design more effective teaching methods and interfaces to spatial data. There are a number of different conceptual schema, which have been developed and built on over the years, to increase the links between the different academic fields in the area and applicability of the schemas. This thesis will consider some examples to see how they have developed, but will not discuss them in detail. For more detailed information, the reader is referred to Kitchin (1996).

Downs' model (Figure 2) includes the cyclical nature of the process of searching for information in the real world and evaluating it with respect to a decision. If there is sufficient information then the decision can be made and the behaviour changed or action carried out, or if there is not sufficient information, the search will begin again.

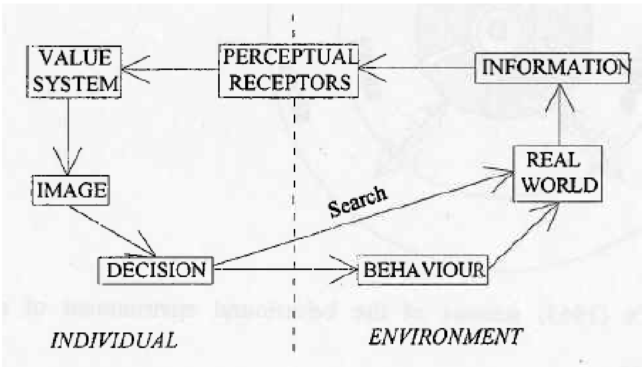


Figure 2. Downs' 1970 conceptual schema (Kitchin, 1996, fig. 2).

Gold's model (Figure 3) highlights the complexity of the processes happening in the model of the mind, often conceptualised as a black box. It highlights the cultural and personal influences on behaviour as well as how these decisions impact both the individuals' behaviour (behavioural environment) and the wider environment itself (objective environment).

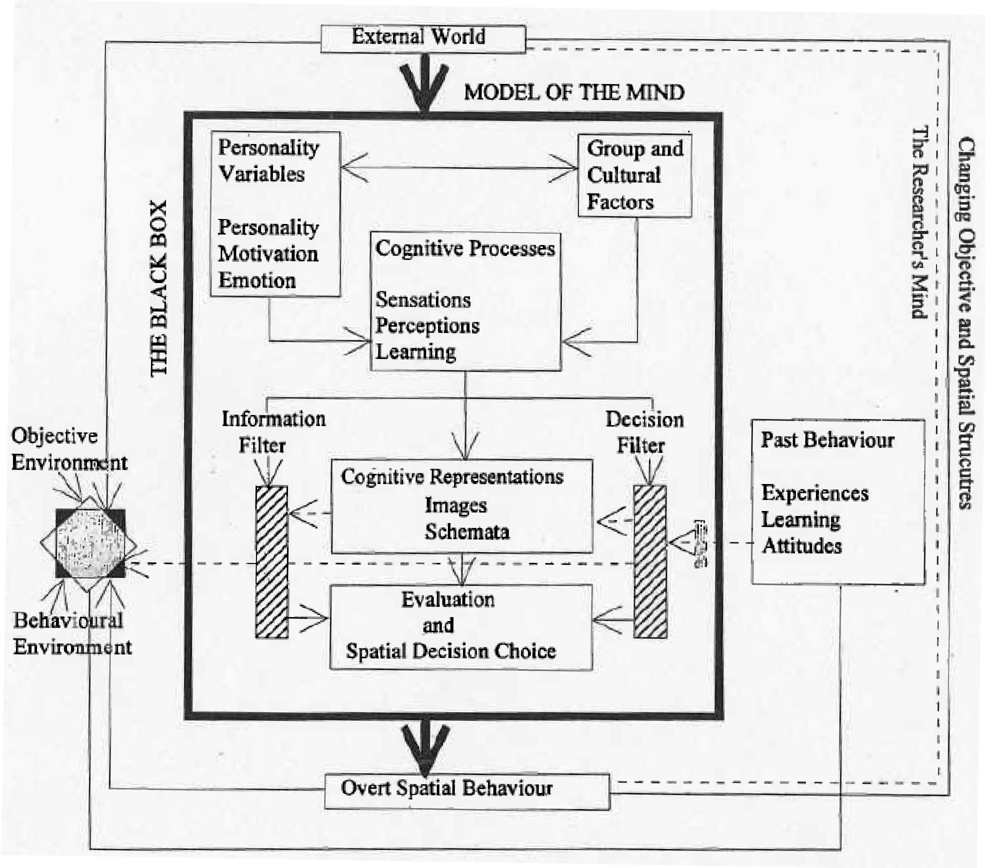


Figure 3. Gold's 1980 conceptual schema (Kitchin, 1996, fig. 6).

Many of these models ignore the environmental aspects that are important sources of information. Kitchin (1996) proposes a new approach, including both the environmental and social aspects of the "real" world as well as working and long term memory as distinct sections of a schema (Figure 4). The real world is split into primary and secondary sources, reflecting the importance of both primary (environmental) and secondary (social) sources of information.

The working memory is the main go-between for the short term memory and the long term memory (the event store, spatial store and characteristics store). It assesses the relevance of the current information to the current situation or decision required and accesses extra relevant information from the various stores. It brings together theories from a range of areas and provides a new theoretical framework for cognitive mapping as well as testable hypothesis.

However, there are still a wide variety of geographic concepts that lack formal definitions and Mark et al. (1999) identify a number of areas that require research, including the spatial dependency of cognitive processing and how perception of distance varies with the familiarity of the units used. This research on cognitive mapping sets the scene for this thesis and provides important building blocks on which the conceptual model created by this thesis sits.

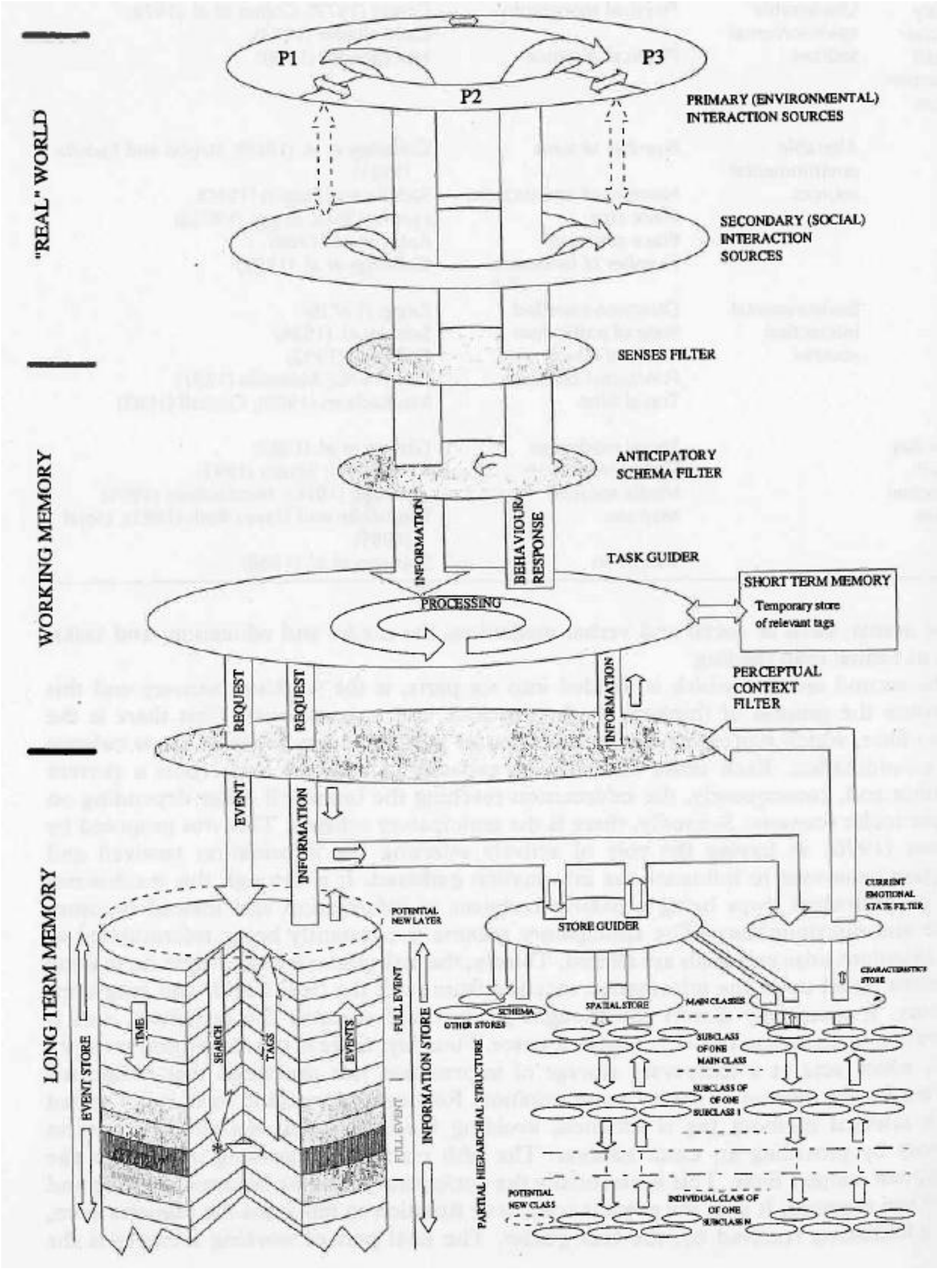


Figure 4. Kitchin's proposed conceptual schema (Kitchin, 1996, fig. 10).

1.2.1 How do humans understand maps and map symbols?

A map is ultimately a product that will be viewed by someone with a specific purpose in mind. There are a wide range of views on how maps are understood by users. A traditional view is that maps contain all the necessary information to solve a particular problem and any misunderstanding of the map is due to poor map design (Bertin, 1984; referenced by Unwin, 1994). The other end of the spectrum says that viewing a map is a subjective process and will be interpreted in a specific way by the user as intended by the map designer (Wood and Fels, 2009), but this may (and often will) vary between map users. How a map is perceived is also largely dependent on context (where it is presented) and framing (the surrounding elements on the map).

Being aware of how people understand maps is very important to extend the use of maps from the visual domain into the auditory domain. Gilmartin (1981) proposes that an approach including both cognitive and psychophysical relationships in perception is required to understand the processes occurring when people use maps. Often cartographers wish to understand how people use maps to improve the map itself, whereas psychologists are interested in learning about the process of communication and the map is just the tool being used. As these two aspects are linked, it is important to be aware of both, even if the research is primarily focused on one or the other.

There has been a vast amount of research on the symbols, colours and designs used on maps, with many standards and conventions regarding symbolisation. Monmonier (1992) discusses the different categories of symbols that can be used on maps as well as the design principles surrounding the symbols themselves. Conventions have been established amongst the producers of maps and have been learnt by map users. Some conventions have an origin in being the most appropriate symbol/colour for that object, but many have just been adopted through use.

While the precise details of the knowledge storage procedure can be very complex (as discussed earlier), the need to understand them is simple – it is to be able to provide spatial data in suitable forms so it can be easily used and recalled when necessary. For gaining (and understanding) spatial knowledge, sight is often thought of as the sense ‘par excellence’ and from this two statements are often assumed to be true. Firstly, that people without sight will have significant problems learning and using spatial data and secondly, that vision is the only sense suitable for spatial data representation. There are a significant proportion of people who have low-vision or are completely blind. Several theories have been proposed about how people remember and encode spatial data, but very little work has been done on how blind people fit in to these categories (Kitchin et al., 1997). However, the view that the blind cannot understand spatial data in the same way as the sighted is not true and the methods they use to interpret spatial data have a wider application for sighted users understanding spatial data. In addition, while the majority of spatial data are represented visually, it is by no means the only way, nor the best way for all circumstances.

1.2.2 Using sound to represent spatial data

Sonification is a method of representing data where it is not possible or it may not be desirable to represent data visually, such as in multi-tasking environments where the user is looking elsewhere or on devices with very small or no screens. The field of sonification covers a wide range of applications, including both visual replacement technologies (such as in an ‘eyes busy’ environment or for blind or low vision users) and visual augmentation technologies where sound is used to add additional information to an existing visual interface, often to improve performance.

An example of a sonification for the blind and low vision users is auditory graphs (Flowers and Hauer, 1995). These are designed to be a replacement for traditional (visual) graphs for visually impaired users. The two dimensional nature of graphs has a very clear link to maps, but the methods used to sonify graphs are often very simple (limited to representing $y = x^2$ or similar). This representation method restricts the amount of information that can be gained from an auditory graph and means it is very difficult to use auditory graphs to represent much more complex spatial data.

Sonification is a new and relatively underdeveloped field. The creation of terms and definitions of key terms (including the term sonification itself) are still under development and discussion. The two academic areas of mathematics / computer science and music have very different approaches to sonification and while interdisciplinary advances are being made, there is still a major gap between the philosophical and practical approaches the two groups take to music (Volk and Honingh, 2012).

Hermann (2008) provides a good overview of how sonification fits into the framework of sounds in the real world. He relates sonification to organised sound, functional sounds and music & media arts (see Figure 5). Organised sound is the parent element in this diagram, where sounds are organised in a specific way for the listener(s).

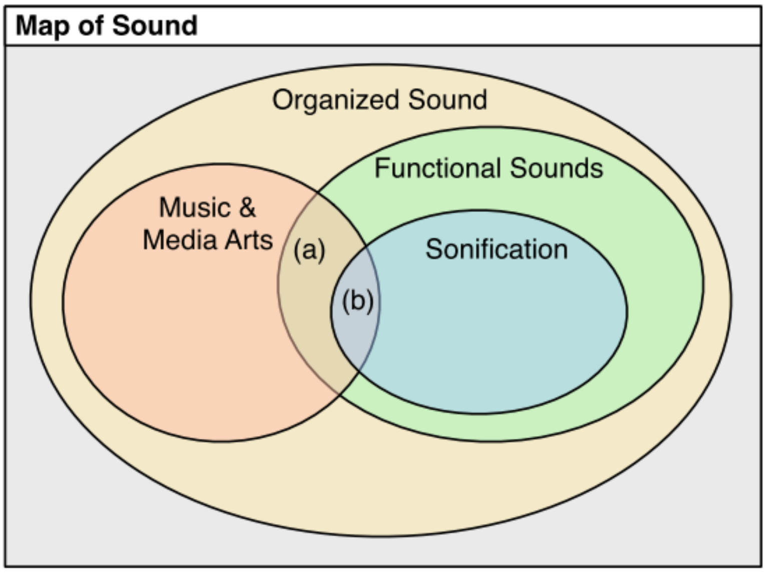


Figure 5. Overview of different types of organised sound, including sonification (Hermann, 2008, fig. 3).

Functional sounds can be any sound with a specific purpose, sonification is a subset of this which is a sound with a specific purpose, driven by data. The idea of sound being driven by data is not new and is something that is fundamental to sonification (Barrass and Kramer, 1999). However, not all functional sounds are necessarily sonifications. For example, the mosquito device (Compound Security, 2012) is a loudspeaker that produces a high frequency sound inaudible to older people, which drives away teenagers hanging around in front of shops. This sound is surely functional, yet it could not be described as a sonification as it is not driven by a data set. The music and media arts category also has some overlap with functional sounds and sonification, but it is definitely an organised sound because it is structured in a very particular way by the composer. The mosquito device mentioned above would fall in to the functional sound category but not the music category. The distinction between sonification and music is less clear, with sonifications often described as being 'musical'. However, this does vary with taste.

Hermann also states that it is important for sonifications to be interactive to be successful. While not a requirement (it is possible for sonifications to just be played back for the user to listen to) the more successful and useful sonifications have some level of interaction (Hermann, 2008; Walker and Nees, 2011). This can range from simple pause or repeat functions (Walker and Lowey, 2003) to allowing the user to customise which data are shown using sound and how the sound is mapped to the data (Harding and Souleyrette, 2010). This interaction allows the user to explore the data and it is crucial to providing new information and understanding (Fisher et al., 1993; Parsons, 1995).

1.2.3 Using sound and vision together to represent spatial information

The cognitive load and dual-coding theories together provide a framework which supports the idea that using sound and vision together will result in better performance than using either alone. The theory of working memory also supports the idea that using the two senses together is not twice as demanding as using each one individually. This section will describe the theories mentioned above and explain how they support the use of sound and vision in combination.

Cognitive load theory states that people have a limited working or short term memory and that the most effective information processing, both in terms of learning a new process and applying an already known process, occurs when users reduce all unnecessary cognitive loads (Sweller, 1994; Oviatt et al., 2004). Baddeley's theory of working memory says that short term or working memory consists of multiple different areas associated with different modes (Baddeley, 1992). The theory of working memory suggests that memory consists of three main parts - the central executive which controls the subjects attention and two sub-systems, the visualspatial sketch pad (which deals with visual and spatial information) and the phonological loop (which deals with auditory and speech based information) (see Figure 6). The two can work independently and in parallel, which allows multimodal interaction to result in better performance (Baddeley, 1992).

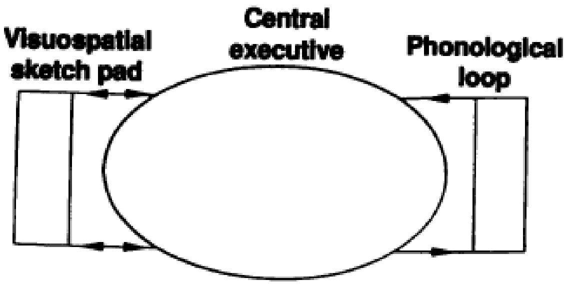


Figure 6. A simplified representation of the visualspatial and phonological areas of working memory (Baddeley, 1992, fig. 1).

The visual and auditory systems working independently allow a user to take in both types of information and process them simultaneously without a major loss in performance. Therefore effective working memory capacity can be enlarged by using multiple channels (Mousavi et al., 1995).

Dual-coding theory states that verbal and visual information is processed separately, but there is a level of interconnectedness between the two (Paivio, 1991), similar to Figure 6 as shown above. If one data set is encoded using both visual and audio variables, then it is likely that participant performance will increase because the data can be processed by both the visual and sonic areas of the brain (Nelson, 1996).

Shepherd and Bleasdale-Shepherd (2008) use Bertin's graphical sign system (Bertin, 1984) as a basis to provide a framework to use different senses for data presentation. The implementation of anything using senses to represent data must, at a minimum, have some model of how multi-sensory data can be represented for effective human understanding. Bertin also proposed a list of visual variables that could be altered (position, size, colour, etc.) as well as the type of data they could represent (categorical, nominal, etc.) and this needs to be applied to the other senses before they can be used in a sensible way. If multiple senses are used together (such as vision and sound) it is important to ensure that the audio and visual elements are synchronised correctly, otherwise the mismatch between the two could lead to motion sickness.

1.3 Human Computer Interaction

Human Computer Interaction (HCI) as a research field is concerned with how humans interact with computers and has expanded to cover a wide range of topics in conjunction with the expansion of technology in our everyday lives. The majority of HCI material covers human’s visual interactions with computers, because the vast majority of our interaction is visual. However, there has always been a haptic element (in terms of physical operation) and with the development of audio stimuli research on sonification has begun within the HCI community.

While there have been some studies using sonification published in the HCI literature (including Edwards et al., 1999; Oviatt et al., 2004; McGookin et al., 2010) there have been relatively few pieces of research considering how sonification fits into a HCI conceptual model (Frauenberger et al., 2007). Ibrahim and Hunt (2007) created an HCI model for evaluating the usability of sonification applications. They developed a traditional HCI model, based on Norman’s Model of HCI (see Figure 7).

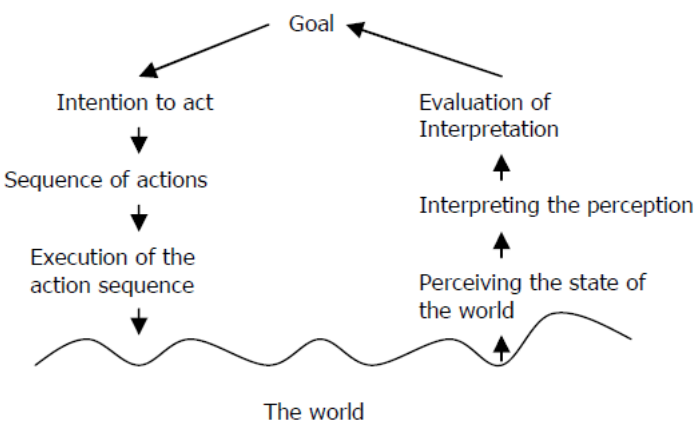
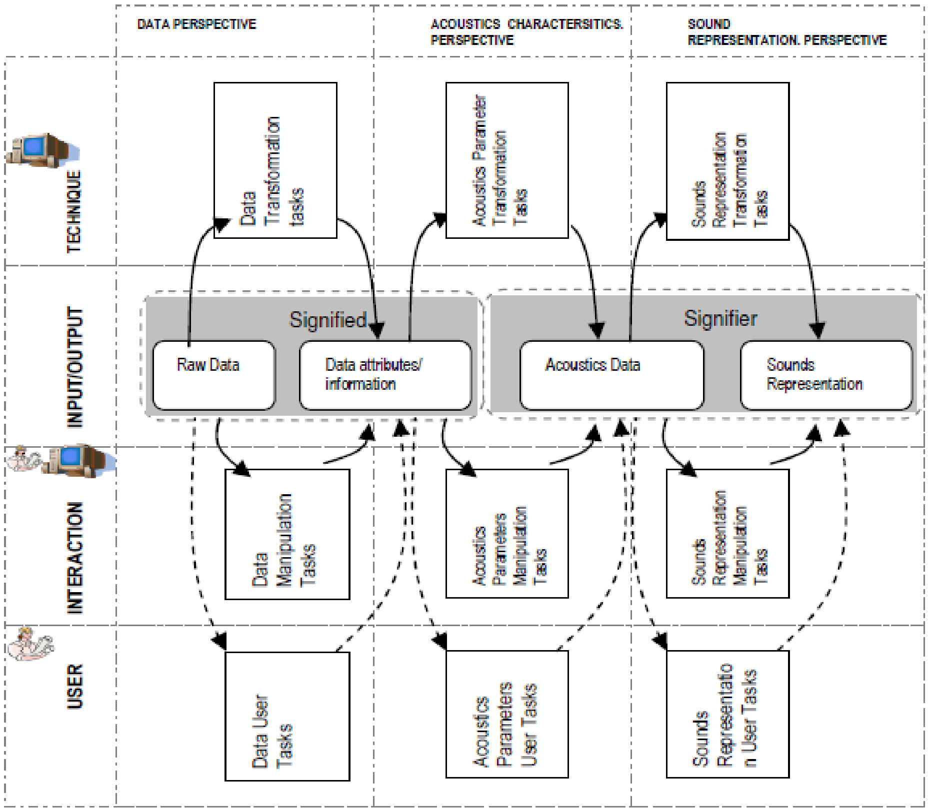


Figure 7. Norman’s model of Human Computer Interaction (Ibrahim and Hunt, 2007, fig. 1).



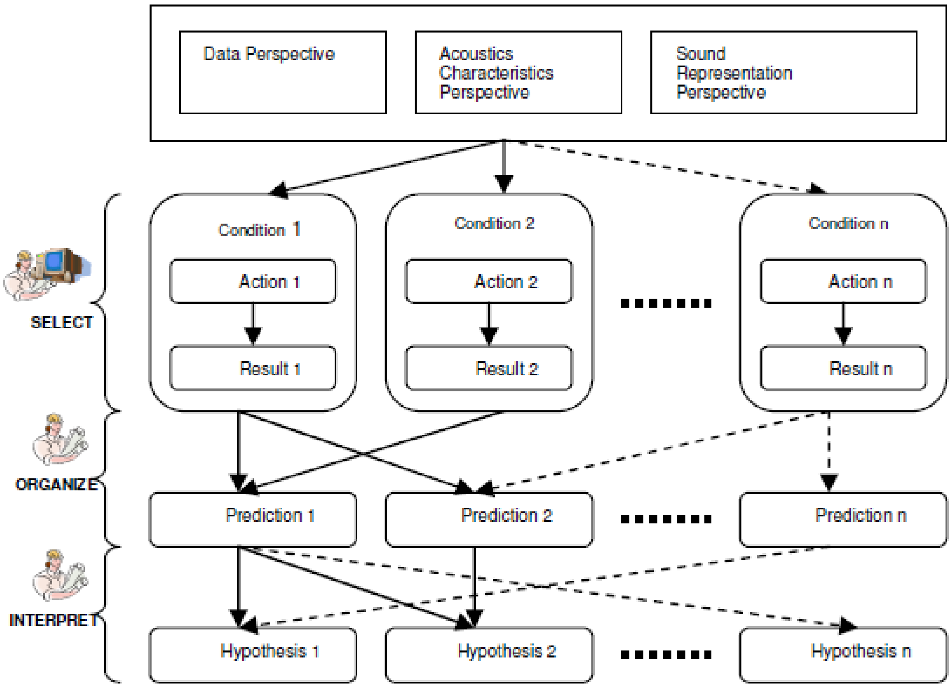


Figure 8. The SA (Sonification Application) Model (above) and the UIC (User's Interpretation Construction) Model (below) from Ibrahim and Hunt (2007, fig. 3 & 4).

Ibrahim and Hunt (2007) pulled out two main aspects from the model and referred to them as the Sonification Application Model and the User's Interpretation Construction Model (Figure 8). The Sonification Application (SA) Model is concerned with how the data (or the Input) is transformed into the sonic representation of the data. The User's Interpretation Construction (UIC) Model is concerned with how the application helps the user perceive the sound as a useful representation of the original data, with respect to the goals they are aiming to achieve. The separation of the issue into these two sub-issues is important, as it distinguishes the user interpretation of the sound from the transform of the data into the sound. However, there is a connection between the two, as for example a particular mapping of data onto sound (in the Sonification Application Model) may be more effective for one user and less effective for another user (e.g. they may prefer a different mapping). Once the mapping has been established, the UIC Model is concerned with how this sound is presented to the user and the context of the sonification. Figure 8 shows these relatively complex models. It appears that these models completed by Ibrahim and Hunt are not well known in the sonification community as no peer reviewed papers have been found that appear to use them, possibly because of the level of HCI knowledge required.

Frauenberger and Stockman (2009) outlined a framework called *paco* (pattern design in the context space) which is a process that helps designers with limited experience of the sonification environment to design sonification tools. It focuses on emerging scientific areas, by collecting a large number of ideas which are then evaluated by experts in the field and then extracting the common 'good' elements and creating a new round of ideas based on these. The iterative process is important in this context, with the first stage evaluation by experts focusing on the circle marked 'Rate' -> 'Extend context' in Figure 9, and the wider circle implemented for evaluation of extended prototypes. Feedback from these is gathered and then applied to the design process again, resulting in an improved prototype. Frauenberger and Stockman (2009) say that this principle could be applied to any design process, but their interest is in sonification so that is what they use as an example. The iterative nature of the process is important as it allows relatively quick and easy development of a prototype.

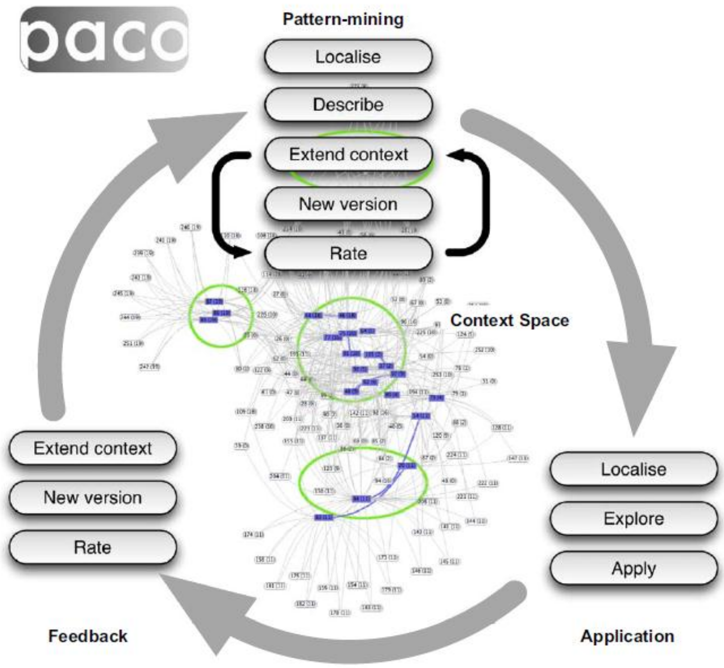


Figure 9. The *paco* framework for sonification design (Frauenberger and Stockman, 2009, fig. 1).

1.4 The Context of this Thesis

The ability to understand spatial data is important as it enables us to navigate to different locations in our day-to-day lives. Understanding space over larger areas (e.g. continents and the world) and representations of this space (e.g. world maps or globes) allows us to comprehend locations that we have not directly experienced. This chapter has discussed how spatial data are understood and remembered by different people. The cognitive methods discussed (such as the hierarchical method of storing spatial knowledge) are relevant to how people understand spatial data and may suggest appropriate methods for representation of spatial data (see Section 1.2).

The vast majority of spatial data are represented visually and many pieces of research have been completed on different visual methods of spatial data representation (e.g. Bertin, 1984; Robinson, 1995) and why these different visual methods are more or less effective (Tufte, 2001). The first question the literature review in the next chapter will consider is 'How does vision influence the interpretation of spatial data?'. This is important to consider because visual representation forms such a large part of the way people understand spatial data. However, there are limits on the amount of information that can be displayed visually without overloading the user. Cognitive load theory can address these issues and suggest ways that could be used to reduce the users' cognitive load (see Section 1.2.3).

The use of sound to represent spatial data has had some research completed using a representation in a desktop GIS environment as well as using sound as an aid to navigation for both the visually able and visually impaired. The second question that will be addressed in the literature review is ‘How does sound influence the interpretation of spatial data?’, because it is important to review the research that has been completed in this area and draw out any relevant information for this thesis. It is also important to consider issues of sonification design, particularly the merits of iterative sonification design, as outlined in Frauenberger and Stockman’s (2009) *paco* framework (see Section 1.3).

While there has been research on how spatial data can be represented using vision and how spatial data can be represented using sound, it is also important to explore how they could be used together to represent spatial data. Cognitive load theory and dual coding theory have both been discussed as different approaches to representing data using multiple stimuli (see Section 1.2.3). The third question that will be addressed in the literature review is ‘How do vision and sound influence the interpretation of spatial data?’ because it is important to review any research that uses sound and vision together to represent spatial data.

Uncertainty is inherent to some degree in all spatial information because of the complex nature of the world and the limited amount of data that can be stored in a spatial database (Longley et al., 2001). However, details regarding uncertainty are rarely collected or made available to the user in a form that is useful. In some data sets recently published, uncertainty is included explicitly within the data set and forms a key part of the data itself. Therefore there is a clear need for this information to be represented in a way that is effective and is understood by the end user. The presence or absence of this data will also impact the understanding of the uncertainty data, so the final question that will be addressed in the literature review is ‘How does uncertainty influence the interpretation of spatial data?’. The user’s previous experience will have an impact on their current understanding, as shown by two of the conceptual schemas discussed in this chapter. Specifically, the past experiences in Gold’s 1980 model (see Figure 3) and the importance of the long term memory in Kitchin’s 1996 model (see Figure 4).

The four questions outlined above form the main structure for the literature review in Chapter 2. Justification for such a focus stems from the issues considered in this chapter. The literature review chapter will build on the research previously covered, addressing each of the four questions in turn:

1. How does vision influence the interpretation of spatial data?
2. How does sound influence the interpretation of data?
3. How do vision and sound influence the interpretation of spatial data?
4. How does uncertainty influence the interpretation of spatial data?

The literature review in Chapter 2 will help define more specific research aims and hypotheses as well as forming the basis of a conceptual model for this research. This will then be followed by a methodology chapter, which explains and justifies the common methodological features used to address the research aims and hypothesis. Subsequently there are three case study chapters evaluating how sound can be used to represent uncertainty in three different examples of spatial data. The results from the case study evaluations will be used to develop and refine the conceptual model to form one of the main outputs for this thesis in Chapter 7. Due to the audio nature of the topic of this thesis a number of supporting materials are available in the included DVD (see Appendix A for details and structure of the DVD).

Chapter 2. Literature Review

2.1 How does Vision Influence the Interpretation of Spatial Data?

2.1.1 Using the visual sense – geovisualization and cartography

For users to make use of spatial data, it is necessary for them to be able to capture, store, manipulate, edit and access the data. In the times before widely available paper and printing, maps had to be individually hand drawn and were very precious resources. With the development of printing presses capable of handling graphics, maps could be reproduced much more easily but the initial design and production of maps was still complex (Robinson, 1995). Even with subsequent advances in technology and data availability, many would say it is still difficult to produce a good map, but now it is remarkably easy to produce a map and many people do (Weber et al., 2012). Many of these maps are of very low cartographic quality, with some being misleading, either intentionally or through ignorance (Monmonier, 1996).

Handling information and presenting it in an effective and easy to understand way is a key concept of information science. Tufte (1983, 1990, 1997) talks in detail about the issues associated with making good and bad graphics, including maps. There are many principles from graphic design including clarity, minimalism and the concept of chart junk. Chart junk is the extra additions to a chart or graphic designed to make them look more aesthetic and Tufte presents many examples of maps that have extra 'graphics' intended to make the map more aesthetically pleasing. However, these extra graphics just distract from the map, and in some places obscure important information. He emphasises the importance of simplicity in design and this is equally important with maps and graphics. These are summarised very well by these four points of graphical excellence:

1. *Well-designed presentation of interesting data*
   – a matter of substance, of statistics and of design
2. *Complex ideas communicated with clarity, precision and efficiency*
3. *Giving the viewer the greatest number of ideas in the shortest time*
   with the least ink in the smallest space
4. *Telling the truth about the data*

Tufte (1990, p.51)

While Tufte does mention many of the negative issues with map design, there are some positives. Maps are crucial as a tool to represent data that varies spatially, as they create a unique way of representing this information so patterns in the spatial element of the data can be seen quickly and easily. It is a much more efficient way of representing information; indeed Tufte estimates that a map of age-adjusted rates of cancer for the 3,056 counties in the US is equivalent to around 21,000 numbers (Tufte, 1983). A graphic (any graphic, not just maps) can show a lot of information, much more than representing it in words or a table of numbers. For example, Longley et al. (2010) describes a map as being worth 'a million bytes'. Design of the map is very important and Longley et al. (2010, p.306) outlines some important factors: "a good design is one that looks good, is simple and elegant, and most importantly, leads to a map that is fit for the intended purpose". Also, thinking about the design of the visual product (be it a map, graph or diagram) is important to make it possible to represent a large amount of information in an effective, interesting and useful manner.

The field of cartography has changed substantially over the last 60 years or so, some might say beyond recognition. Unwin (1994) provided an interesting review on what has happened to the field of cartography, how it has changed and how it relates to the fields of scientific visualisation and GIS. However, this development and use of technology has created two different schools of thought about mapping and the presentation of spatial data. Geovisualization makes extensive use of technology to create new and different ways of representing traditional types and new types of spatial data. Cartography focuses on the tradition of 'making a good map' and uses technology in a different way to achieve its aims. The two are not necessarily opposed, but frequently do not overlap.

Wood and Fels (2009) approached the discussion of maps from a cartographic point of view and discussed extensively 'The Natures of Maps' considering how the content of maps has changed and how these changes reflect the motivations and decisions of the map makers. For example, there appeared to be many unstated assumptions and 'western' views implicit in maps published in National Geographic as well as embedded in the content from decisions taken on colours and design aimed at creating a certain impression for the reader. Wood and Fels also state that maps should not (and cannot) be considered 'the truth'. For example, political boundaries shown on a map often have no physical basis (apart from those that follow geographical features such as rivers), but often maps fail to explicitly show this and represent the boundary as a 'hard line' on the map. A similar occurrence happens when attempting to communicate the path of hurricanes in the US. The forecast paths were often shown as a 'cone of uncertainty' widening over time, but if it is represented with a hard line the viewers would believe that the hurricane will not cross that line, as well as making other assumptions about the graphics which are not true (Broad et al., 2007).

In contrast, the geovisualization field has developed many new techniques of displaying traditional and new spatial data, driven forward by technological developments and availability of data. This has resulted in a shift in visualisation from a presentation tool (i.e. we will create this nice visualisation to show our results) to an analysis tool (i.e. the visualisations help us understand and analyse our data better). Additionally, the advent of large, complex data sets (such as atmospheric general circulation models) means that data has to be shown visually as it cannot be shown any other way (e.g. in an equation) (Unwin, 1994). This use of visualisation as an analysis tool is now fundamental to how any modern GIS functions and techniques such as brushing (linked displays) are a significant part of this, as shown in software such as OpenGeoDa (GeoDa Center, 2012). The visual element has always attracted public interest and the launch of Google Earth (as well as other virtual globes) initiated a significant public interest in and demand for free spatial data.

At the point Unwin's article was written (1994) visualisation techniques had begun to mature, particularly in terms of printing detailed maps in colour, but animation techniques were still very new with some rules of use being developed (Monmonier, 1992) and sonification, while being mentioned, had very little development. Additionally, there is relatively little knowledge on the perception of patterns in motion, which is of particular relevance to animation (Ware, 2004).

Cartography (as a discipline) has broadly been replaced by technology (GIS and ViSC (scientific visualisation)) but there is still an element that remains, part design, part understanding of maps that Unwin does not appear to be able to incorporate into GIS and ViSC. Some of this is now in Human Factors / Human Machine Interface / Human Computer Interaction, some in Design, some in Cognitive Mapping and some in GIS. How much one experiences of this depends on ones colleagues, supervisors and their experiences, but at times the area can appear fragmented. While the categorisation of skills may vary depending on whom you ask, map making is a combination of a wide range of skills, as shown by this quote:

> *"In additional to being familiar with computer graphics technology,*
> *the current generation of cartographers must be part artist and*
> *part cognitive scientist as they try to construct better maps."*
> *Mark et al. (1999) p.752*

2.1.2 Visual variables

Mapping has a long history and it is widely recognised that there are six main visual variables that can be used to represent information; size, shape, greytone value, texture, orientation and hue (Bertin, 1984; Robinson, 1995). These can be applied, as shown by Monmonier (1996) in Figure 10, to points, lines and areas and they are commonly found on all maps. Sometimes two (or more) visual variables are combined to reinforce information; for example motorways are often shown in blue (hue) and as a thick line (size). Texture and orientation are often categorised as secondary rather than primary visual variables because they create patterns in relation to other objects around them, or they are very dependent on their context rather than being distinct in themselves (Robinson, 1995).

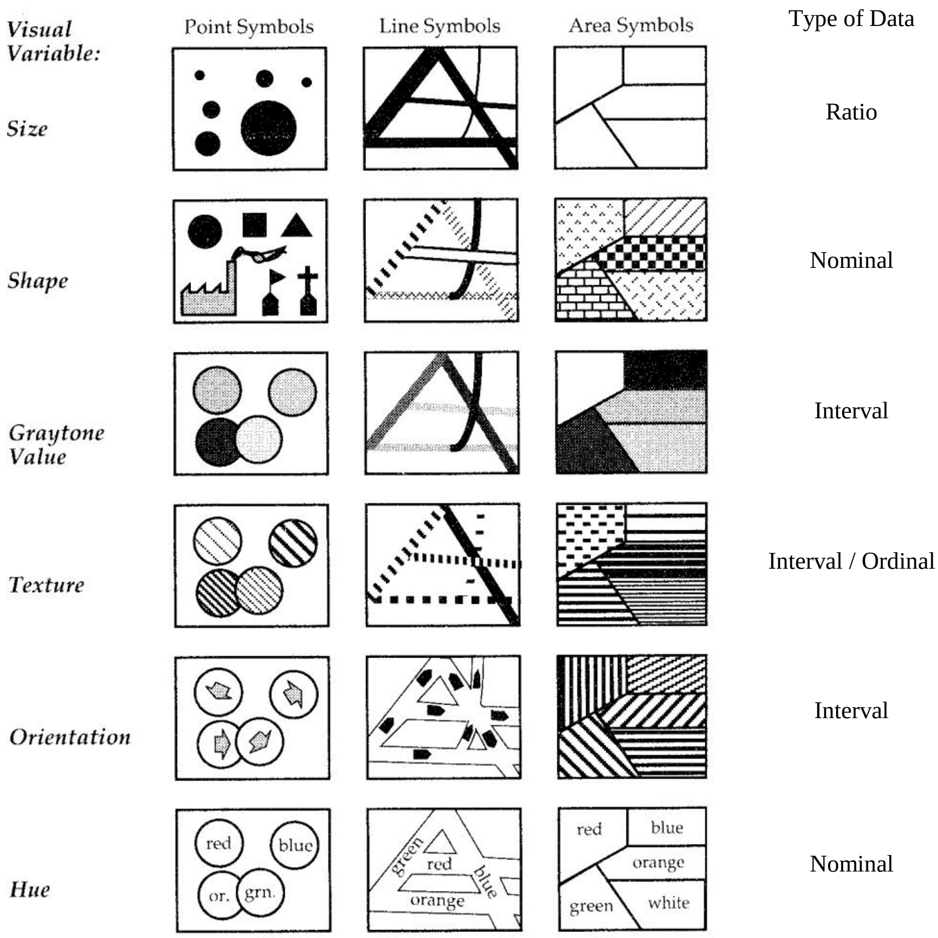


Figure 10. The six principal visual variables with the type of data they can represent (after Monmonier, 1996, fig. 2.11).

Different visual variables are appropriate for different data types. For example, shape and hue have no particular inherent order, as it is meaningless to say blue is bigger than yellow. So shape and hue can only be used with nominal data. Greytone, texture and orientation work well with interval data, as the distance between values is meaningful. For example, two different orientations can be a number of degrees apart. If the proportion between values is not meaningful, then only ordinal data can be used. For example, this may be the case with texture if it is not possible to easily identify the differences between the different textures. Size is one of the most useful variables as it can represent ratio data, where proportion is meaningful, as well as the value of 0.

Maps are combinations of colours and symbols that have some meaning for the user. The symbols and colours can either have an intuitive link to a particular meaning or a learned meaning. One linkage that could be considered intuitive is that red relates to hotter conditions / higher temperatures and blue relates to colder conditions / lower temperatures. Even though people say they think that red is inherently a hotter colour and blue is inherently a colder colour, research has shown that it is a purely intellectual effect. For instance, in one study participants were located in either a red or blue room and asked to judge their thermal comfort. It was found that the colour of the room made no different to the participants thermal judgment (Bennett and Rey, 1972; Candas and Dufour, 2005).

While colour is often used for symbols, it is also used in thematic maps in a very different way. Choice of colour for thematic maps is very important from a usability point of view, both in terms of the colour blind and those with normal vision and this choice is something many map creators find challenging. Cynthia Brewer addressed this problem by creating a tool that allows users to specify the type of data they have, number of classes and so on and provides a range of appropriate colour palettes that the user can select from, based on the nature of the data, the number of classes and the end-use environment for the map (Harrower and Brewer, 2003).

Colour blindness effects about 1% of the female population and 10% of the male population (Ware, 2004). Often it is not noticed by the sufferer unless they are in a specific profession where colour is important (Ware, 2004). Colour blindness is a particular problem for users of some Ordnance Survey products and in response Ordnance Survey designed a number of customisable colour palettes that are more effective for people with a variety of types of colour blindness (Mainwaring, 2011).

The development of map symbol terminology has gone through a similar process of repetitive exposure and improvement (Wright, 1944) to achieve a set of terms everyone uses and understands. If the majority of associations used in map colours and symbols are learnt through exposure, then it should be possible to learn associations between particular sounds and variables. While some of these may be fairly obvious (e.g. bird calls representing the presence of birds) other associations may not be (e.g. bird calls representing a rural environment). The sounds chosen need to be chosen carefully and tested to make the interface as accessible as possible.

Associations of colours with particular meanings have a variety of origins. Blue is often associated with water on a map representation, but it is very rare that the particular shade of blue used on a map is found in water, except in “the minds of wishful environmentalists, self-serving tourist operators and gullible map readers” (Monmonier, 1996, p.171). Blue is also used to represent motorways, following the convention of the use of blue signs on motorways. There was no specific reason that blue was chosen, other than it made a clear and obvious sign and was sufficiently different from existing signs (Design Museum, 2006). While some symbols have a scientific justification for being chosen to represent a specific object, many are arbitrarily chosen and some (such as the use of greens and browns for height in some topographic maps) bear no relationship to what they are trying to represent. Often they also have no relationship to each other for the eye itself. For example, brown often represents higher elevations, but in itself it is no 'higher' or 'lower' than green (Robinson, 1952).

While associations such as blue representing rivers or the ocean may be universally accepted, these associations (and conceptualisation of wider geographic concepts) can vary across cultures (Mark et al., 1999). For example, red is often used to mean 'stop' or as a warning, such as on octagonal stop signs or red traffic lights. However, in parts of Asia red is often associated with heroism or good luck and it can be traditional for brides to wear red wedding dresses. So while some colours have been universally learnt, others may have different connotations depending on the cultural background.

Differences in the way people perceive information (be it spatial or non spatial) can occur for different reasons and be categorised accordingly. Cultural and linguistic differences include examples such as the perception of the colour red, as mentioned above, as well as definitions of specific words, which may have different meanings in different languages or cultures. Disciplinary or experimental differences result from specialised training in a specific field with different experts considering the same objects in very different ways. For example, a forester may see a forest as a series of stands of trees, to be logged in a specific order, while an ecologist might see the same forest as plant and ecological communities with no sharp boundaries (Mark, 1993). The final category of differences is that of individual differences, where people's perceptions vary because of a range of factors, some of which are highly individual. These are discussed in the next section.

2.1.3 How does understanding of visual variables vary between users?

As well as variations between cultures, there are also significant variations between different groups of users. Gender is the most explored aspect and is discussed below with a range of other relevant factors.

Many studies have shown that males perform more effectively on spatial tasks than females, particularly those involving complex 3D rotation of shapes (Montello et al., 1999). However, it appears to be that females are better at some spatial tasks than males, leading to the possible conclusion that males and females take different approaches to assessing and solving spatial problems. The differences that are seen between males and females do vary with previous exposure to spatial data. An iterative process can result, where those people exposed to events with high spatial components (e.g. LEGO-construction, exploratory games and team sports) improve their spatial skills, while those who are not exposed to such activities do not improve their spatial orientation skills and so subsequently do not seek out these activities (Coluccia and Louse, 2004). Gender differences vary for different aspects of sketch mapping, with males approaching the map from a global perspective and females focusing on local landmarks (Coluccia et al., 2007). Some research has suggested that the gender difference is reducing over time, possibly as a result of a more equal culture (Self et al., 1992). While differences have definitely been seen, the reasons for these differences are not fully understood, partly because the effect is confounded with many other variables including experience of spatial data. Additionally, the area of ‘spatial tasks’ covers a wide range of skills and uses so it is likely that the evaluations may not be asking the right questions to assess this difference. The differences recorded in a laboratory setting have not been duplicated in the real world and many evaluations use undergraduate university students which is not a representative group of the wider population (Self et al., 1992).

People's spatial processing abilities will be affected by their previous experiences as well as the current situation (Kitchin, 1996). Users in an experiment run by Gluck et al. (1999) performed well in the first part of the task, which was to seriate (arrange in order) a series of variables in a structured evaluation, including options of using sound to represent some of the data. The final task was a less structured exercise that encouraged data exploration, using the same tools as the earlier task. The users had trouble using the tools for the unstructured task, particularly the sound. The authors suggested it was the result of them having little experience with GI and not knowing how to use the tools. Unfortunately the evaluation did not include GI users as well as non-GI users, so no comparison can be drawn. There is also an important age difference, with older subjects performing less well than younger subjects in spatial tasks, but the reasons for this are relatively unknown (Liu et al., 2011). There are also potential interactions with sound and age but these are not well explored.

Significant gender differences are reported in angle estimation of spatial data, but there is no clear difference in distance estimation with the literature divided on whether there is a difference or not (Holding and Holding, 1989; Galea and Kimura, 1993). Fisher and Orf (1991) showed that people’s familiarity with a particular location did not have a significant impact on perception of distance within that environment when asked whether buildings were ‘near’ or not.

As mentioned before, previous spatial experience will have an impact on current spatial ability. Additionally, confidence in one’s own spatial ability (or lack of confidence) will also have an impact (Casey et al., 1997; Ramirez et al., 2012). Spatial ability covers a wide variety of skills and there are many different tests aimed at quantifying this. People’s ability to estimate their own spatial skills are generally reasonable, with the literature showing a range of between 50% and

60% correlation between a person's estimate of their spatial skill and their measured spatial skill (Hegarty et al., 2002; Ackerman and Wolman, 2007) . However there are some differences and some studies measure sense of direction rather than more general spatial ability.

2.1.4 Non visual map users

Even though the visual sense is very dominant in navigation, around 90% of input compared to 10% for other senses (Bujacz and Strumiłło, 2007), visually impaired users (those with either no or low vision) have the same requirement to be able to use and understand spatial data as everyone else. The visual dominance in spatial data is something that a person is exposed to from birth and reinforced daily by immersion in our visual culture. This can make accessing spatial data represented visually difficult for those unable to use vision effectively and also predisposes people to visual representations of spatial data over other methods (Rice et al., 2005). Indeed a range of studies found that only between 30% and 40% of blind people surveyed make independent journeys outside their own home and 20% of respondents had not left their home at all in the preceding week because of their limited access to spatial information (Jacobson, 1998). This restriction on being able to navigate independently has a great impact on a person's quality of life (Golledge, 1993; Jacobson, 1998). When considering blind users, there may be cognitive differences between how those who were born blind and those who become blind in later life interpret spatial data (Kitchin et al., 1997).

Several studies have investigated the creation of navigation aids for visually impaired users with a variety of solutions from vacuum formed tactile maps to GPS-enabled navigational aids (e.g. Loomis et al., 1998; Bujacz and Strumiłło, 2007). However, with a significant proportion of this research there has been no or very little engagement with the visually impaired community, meaning that while these technically advanced tools are created, they are rarely used because the user community does not find them helpful (Perkins, 2002). When designing a new tool or solution for a particular user group or task, it is very important to include that group in the process and ensure you understand their needs and requirements so the tool that is created is beneficial to the user group.

Echo location techniques in blind participants are possible but are difficult to learn. One tool (Kaspa) uses a stereo sonar signal to act like bat 'echo location' for blind users, but it takes many months to learn how to use the system and ideally the training should be done as a child for it to be most effective (Bujacz and Strumiłło, 2007). Blades et al. (1999) observe that blind users can gain effective access to spatial data though tactile maps, but blind users do not always use the best information encoding method to allow them to get the most from the map. Therefore training in the use of tactile maps is important.

More recent research has focused on how blind participants ‘see’ the world (i.e. their cognitive spatial processes) with an aim to use this knowledge to create truly useful navigational aids for them. Specific studies have looked at pedestrian navigation in Belfast (Kitchin et al., 1998) and Singapore (Pow, 2002), considering the problems that the visually impaired face and how these may usefully be addressed. Pasqualotto and Newell (2007) compared spatial memory abilities of participants with different vision levels and found that sighted and late blind individuals performed better at the task than congenitally blind participants.

One area of research particularly related to this thesis is using sound to create maps for blind or partially sighted people. This is when the visual sense is replaced by the sonic one, as shown by Zhao et al. (2008) and Medvedev (2011) where sound is used to represent specific elements of interest in a visually based spatial display. This very much limits the amount of information that can be communicated, as it is difficult for sound to represent spatial data over an area, as generally vision works over space and sound works over time. When hearing the sound, the user has to 'visualise' the data they are hearing, which is probably not the best way of utilising sound. Sound on its own can be used to communicate lots of information (e.g. conversation, narration, film sound track), but this gives more of a general overview than a specific picture.

2.1.5 Big data and visual saturation

While visualisation is the main tool used to represent spatial data, there are limits on the quantity of data that can be displayed visually. These limits are going to become an increasing problem, because there is an increase in the number of people using spatial data (as mentioned earlier) as well as an increase in the amount of spatial data that are available. This is referred to as ‘big data’ (Batty, 2012) or the ‘data deluge’ (Porter et al., 2012) and is something that concerns nearly every academic area.

Batty (2012) talks in very general terms about the amounts and types of data that we will be faced with handling in the near future. The size of these data sets will be beyond anything we currently have and new ways of handling and interacting with these data will be required. Undoubtedly animation will be a way of dealing with some data sets (Openshaw et al., 1994; Radburn et al., 2010), but sonification will also open up new possibilities for representing, analysing and understanding the data. The source of these new data will vary depending on the area, but developments in electronics and miniaturisation have meant that large scale networks in the order of thousands or tens of thousands of sensors are becoming much more realistic (Porter et al., 2012). Big data are prevalent in many different disciplines and requires both new ways of collecting, storing and analysing data as well as new ways of representing these data (Keller et al., 2012).

Cheshire and Batty (2012) also comment that while the temporal aspect of big data is not that new (as we have been representing temporal data using animation and other methods for a number of years), these types of data are starting to become available in real time, which requires new and novel methods of representation, of which sonification may be one. Additionally the increasing availability of uncertainty data is another strand that adds to the amount of data that is available to us.

Along with the development of technology which has driven many of the advancements in geographical information science over the last 50 years, the development of the Internet and easily accessible GPS enabled mobile phones have resulted in an influx of VGI (Volunteered Geographic Information). VGI is a form of citizen science, where volunteers provide information on the spatial location of different objects, including time sensitive information such as road traffic congestion (Haklay et al., 2008; Haklay, 2012). This is both increasing the amount of spatial data available, but it also brings extra problems with it to the geographic information user. This is because different users provide different information, including information of different quality and it is not always immediately obvious whether the data provided is of sufficient quality for a particular use.

2.2 How does Sound Influence the Interpretation of Data?

2.2.1 How do we use sound to understand the world?

Sound is one of the most important senses after vision (Fortin et al., 2007), something we use in our everyday lives, often without explicitly realising it. We use sound to tell when the kettle has boiled, to help us know when it is safe to cross the road, to know someone is at the front door or that our phone is ringing. Sound is also used in less explicit ways, a common one being in films. While the first silent movies had a pianist accompaniment, soon sound became integral to the film and has progressed from the first "talkies" (with relatively low quality mono sound track, often just the speech of the actors) to today's modern films with surround sound, effects, speech and background music all combined together to enhance the film experience. As well as the obvious speech that sound provides, it also adds atmosphere to the scene. Watching a film without its soundtrack (but still with speech) is a very different experience to watching a film with its soundtrack; it is surprising the amount of information that is lost.

2.2.2 Why non-speech sounds?

While sonification is by definition use of non-speech audio, there are instances where speech is useful. However, when using sound to represent data, speech can be very limiting because it can take many words to describe what may be a relatively simple pattern in the data. Equally, speech suffers from many of the disadvantages that text based systems suffered from, prior to the introduction of graphical displays (Barker and Manji, 1989).

Additionally, sonification is designed to make the best use of the human auditory system, exploiting the ability to hear subtle differences in the patterns and the timing of sound, which synthesised speech cannot currently exploit. As humans we hear some of these subtle changes in real speech, but currently there is no way to artificially create these subtle changes in synthesised speech (Brewster, 1994).

2.2.3 Background to sonification

The research field of sonification is the result of collaboration between academics from a wide variety of backgrounds including computer science, psychology, engineering and music with the wish to represent data using sound in a systematic, formal manner (Walker and Nees, 2011). The concept is not new (see below) but the development of personal computers that can generate their own sound has led to a wide range of new possibilities over the last 20 years. Within its area, it is maturing as an application method, but it has yet to develop significantly beyond expert users. A number of reviews have been conducted (Kramer, 1994; Barrass and Kramer, 1999; Flowers, 2005; Hermann, 2008), but there is not yet an established methodology for creating sonifications accessible for newcomers to the field. Towards the end of the research for this thesis, a new book was published, ‘The Sonification Handbook’ (Hermann et al., 2011), which is designed as an introduction for new users as well as a consolidation of current practice.

As mentioned already, sound is the next most powerful sense in the body after vision in terms of information provision (Fortin et al., 2007). While the process of measuring the amount of information captured by different senses is complex, one source claims that the eyes can process 50 billion bits per second and the ears can process 1 billion bits of information per second (Kurzweil, 1992 in Lange, 2005). While this would say that vision is 50 times as powerful as sound, 1 billion bits per second is still a significant amount of information (~119 MB/s, megabytes per second). Nørretranders (1999) also attempts to quantify the amount of information humans can receive and estimates vision can receive 12 MB/s, touch 1 MB/s and hearing, smell & taste combined 1 MB/s. While the numbers differ, the relative proportion of visual to audio is roughly the same. However there is no doubt that that the sound-brain system can receive a large amount of information and is very good at hearing and storing non-speech sound. Sound can also signal background activities that may not be visible and it does not require the user to be orientated toward the object in question. The brain can use sound to perceive highly transitory events, can hear many forms of repetition and correlation within sound and may be better at comprehending time-varying and multidimensional data sets using sonic rather than non sonic methods (Shepherd and Bleasdale-Shepherd, 2008). Specifically, selective attention is possible using the auditory system, where the listeners 'tune-out' certain sounds and only hear them again if the sound changes. This could be very useful for listening to datasets that contain too much noise to be viewed visually (Eldridge, 2006). The majority of people do not fully utilise the sonic comprehension they have, and a very good example of how some people have utilised their sonic comprehension much more effectively than the majority is the blind community. When a sense is taken away, the remaining senses become more acute. A number of congenially blind people (those who were born blind) have made exceptional use of their hearing and can navigate their way around space using an echo location method similar to that of bats (Gardiner and Perkins, 2005).

The concept of using sound to represent information has been around for a very long time, with the first recorded instance potentially being around 3500 BC in Ancient Egypt. Two independent records were made of the transfers of grain in and out of the grain stores, which were then read out simultaneously in front of the Pharaoh, who could tell instantly if there were any discrepancies (Worrall, 2009). Another example, although not quite as old, is that of the Geiger counter. Nearly everyone is familiar with the concept of this radiation detector, with more frequent beeps representing a higher radiation level.

Pollack and Flicks (1954) were one of the earliest who considered the possibility of representing a data set using sound. They performed some exploratory tests, and concluded that sound could be used, but faced some limitations in terms of the amount of information that could be represented. Yeung (1980) was one of the earliest examples of evaluating sound as a means of representing multivariate data (sonification as we would call it now). He used a relatively complex representation of sound (seven different variables represented using seven different parameters of sound including pitch, loudness, damping, direction, duration and silence). The participants in his study (professional chemists) were able to extract patterns from the data with a 90% accuracy rate without training and a 98% accuracy rate with training. This is a very good example of how expert users can quickly adapt to a sonification representation of familiar data. However, Yeung does not compare different types of sonic methods so it is not known which sonic variable of the seven he used was the most effective.

More recently, sonification has been formally defined as 'use of non-speech audio to represent data' as a part of the formation of ICAD (International Conference on Auditory Display) in 1994 (Kramer, 1994). There have been a number of reviews of the history of sonification (Bishop and Cates, 2001; Worrall, 2009; Dubus and Bresin, 2011) many of which comment on the fact that sonification/sound has often been ignored in interface design despite the great potential it has.

This is partly due to the lack of a mature, general purpose sonification tool (Flowers, 2005) and a limited theoretical framework (Frauenberger and Stockman, 2009).

Whilst visualisation may appear the obvious method of display for all spatial data, there are times when a static visual display (such as a printed map) is not sufficient or suitable for use. Animation of a series of these static maps works well for temporal data, showing changes over time (Tobler, 1970; Dorling and Openshaw, 1992; Unwin, 1994). However, very quick changes or variations, particularly for complex patterns, are not handled well in the visual system. Auditory displays exploit the ability of the human auditory system to recognise temporal changes and patterns which may be missed by the visual system (Flowers and Hauer, 1995; Walker and Nees, 2011) and can also reduce visual saturation and workload by representing some data using sound instead of vision (Brewster, 1997).

Flowers et al. (1997) looked at the possibilities relatively early on, and observed at this point that it was necessary to start with simple sonifications until the tools and methods could be evaluated and then become more complex. They also highlighted the technology at the time as a limit on the sonifications they could create. The use of audio in general has had very little development in the wider community and there remains doubt about how useful audio can be. This is partly due to a lack of guidelines and tools for designers to use (Frauenberger et al., 2007).

The first major publication on sonification was as a result of a meeting in 1994 and the publication itself (a book) is very hard to access (Kramer, 1994). Finding material to reference is a particular issue in the community, and is something that the community is trying to address (Stockman 2012, pers. comm.). The Kramer book is seen as the source of original information, and the new Sonification Handbook (Hermann et al., 2011) is seen as an important follow up to Kramer's book. It acts as a good introduction to the area and a way of bringing practitioners up to speed with current thinking and new developments. The Handbook may begin to act as a guide covering the entire breadth of sonification, but before it was published the work in the field tended to be fragmented across conference papers and journals in a variety of different fields. These could be diverse as computer music, human computer interaction, multimodal interaction, civil engineering, geography, experimental psychology, artificial intelligence, cognitive psychology and many others.

The ICAD conference series is one of the main academic groups using examples of sonification in their research. ICAD is relatively new (the first conference was 1994) but has grown rapidly and now showcases a wide variety of applications using sonification. While very few of the examples in ICAD consider spatial data, the research is relevant because there is a lot of information on which sounds and which sonification methods have been used and how effective they are. A number of the papers from the conference series also look at the background to sonification and provide a review of the thinking behind the concept.

The terms ‘auditory display’ and ‘sonification’ are sometimes used interchangeably and are sometimes used to refer to different, but related concepts. Sonification is the term generally used in this thesis. Auditory display is often considered to be the physical output of the sonification – including the hardware such as speakers/headphones (Hermann, 2008).

2.2.4 Theory of sonification

While sonification has developed dramatically over the last 20 years, there is a very limited theoretical framework surrounding sonification. The need to develop a framework has been mentioned many times, both from a theoretical point of view and a practical one (Barrass, 2005; Flowers, 2005; Eldridge, 2006). However the sheer diversity of the community may limit this, at least until a broad inclusive solution is found. Indeed, Frauenberger and Barrass indentified seven different design approaches in a workshop at ICAD 2008 (Frauenberger and Barrass, 2008) so an over arching, all encompassing framework is likely to be difficult to develop, partly because of its dispersal over such a wide range of academic areas. Frauenberger and Stockman comment that “much of the sound produced by today’s technology fails to exploit the sophisticated abilities of human hearing" (Frauenberger and Stockman, 2009, p.907). They make specific reference to two sonification map examples (Horowitz, 2007; MacVeigh and Jacobson, 2007), neither of which make any reference to the design process, consideration of other options or justification for the decisions chosen. They propose a new design framework to help novice auditory display designers. However the framework is relatively complex and all of the participants who evaluated it said it was time intensive to complete and some aspects of the process were confusing. Some also said they found it heavily cluttered with domain specific jargon which limited its application. While there is a need to develop a framework, Eldridge (2006) observes that this need should not stand in the way of developing specific applications as it would stifle development.

One reason for the limited overall framework may be a result of the wide variation in applications of sonification and the lack of overlap between them. In a survey of sonification experts (people who have had sonification research published multiple times at ICAD, the conference for sonification) 4 out of 13 said that the properties of the information involved drove some of the fundamental design properties (Frauenberger and Stockman, 2009). Additionally, one expert from the same study reflected a view that was common amongst the participants:

“Furthermore, a majority of the knowledge base specific to auditory display has been generated with a focus upon only narrowly contrived, highly specific applications. Usually no attempt is made

to refer to, draw upon, or contribute to any greater theoretical framework...."
(Frauenberger and Stockman, 2009, p.914).

There are a number of overarching conceptual models on sounds themselves and how people relate to these sounds, but very few on how these sounds can be used to represent data. This is very different to the visual field, where many models have been developed (see comparison in Nesbitt 2006 and Section 1.2 of this thesis). Nesbitt looks at sound in two different ways – firstly the musical properties of a sound (e.g. pitch, loudness etc.) and secondly the everyday properties (how we actually hear the sound, e.g. what creates the sound, the force, material, configuration of it etc.). Figure 11 shows the musical properties of a sound, including aspects of pitch, loudness, timbre and so on. This is by far the most common approach taken in the sonification community, and this method is very easy to understand and use, although some of the different elements (such as timbre) do not have very well established definitions.

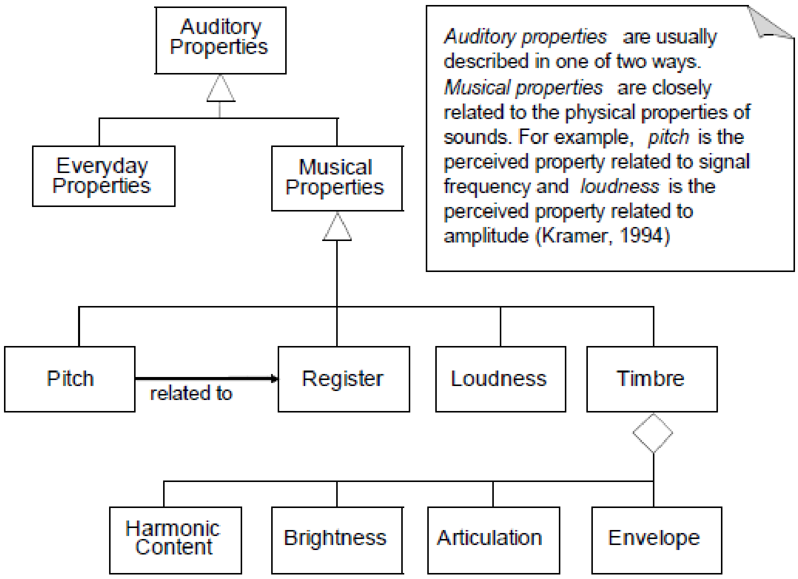


Figure 11. Auditory properties described using musical properties, the most common method used in the sonification community (Nesbitt, 2006, fig. 3).

The second approach (Figure 12) looks at the everyday proprieties of sounds, or in other words, the objects used to create the sounds. This is very useful in a practical way, but also may provide a different way of categorising 'similar' sounds and 'different' sounds. This aspect of Nesbitt's research is based on work by Gaver (1993) and is something that has been used by the sonification community, but is not often used as a theoretical basis.

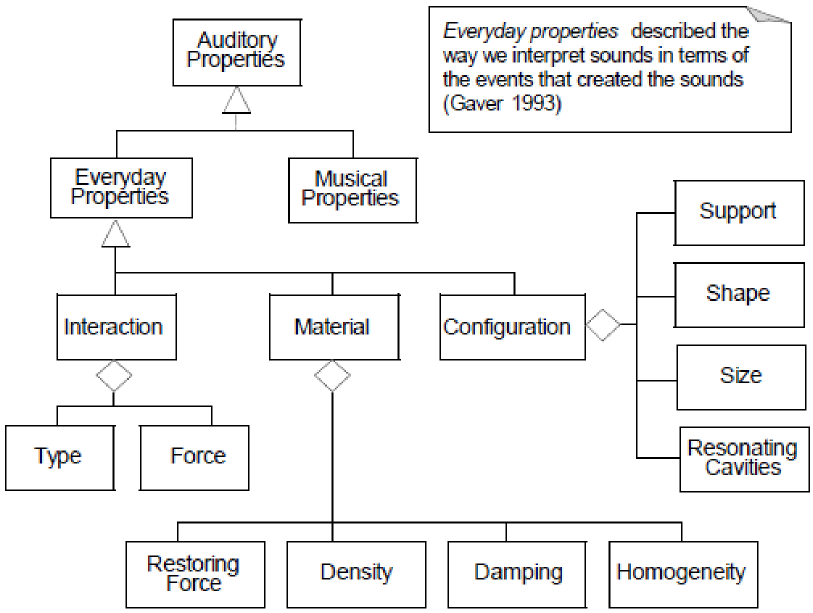


Figure 12. A different approach to describing auditory properties, based on Gaver (1993), considering the issue from the point of view of the objects that actually create the sounds (Nesbitt, 2006, fig. 4).

While neither of these approaches is uniformly better or worse than the other one, it is worth remembering that the terms that are often used to define sounds (pitch, volume, timbre, etc.) are not the only way of measuring sound.

Diniz et al. (2012) propose a different approach to sonifications, where the mapping of data to sound is an interactive one which the user can alter as required. They compare the traditional, 'static' approach (Figure 13) to a 'interactive' approach (Figure 14) where the user can adapt the sonification mapping to allow more effective exploration of the data.

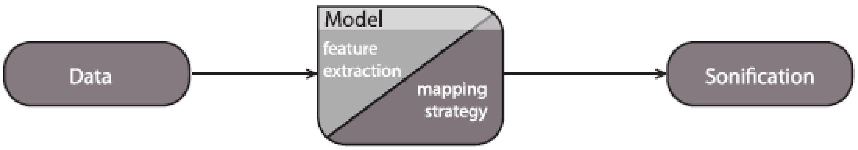


Figure 13. One-way or static sonification. The model (feature extraction and mapping strategy) is fixed and cannot be changed by the user, limiting their interaction with the data (Diniz et al., 2012, fig. 1).

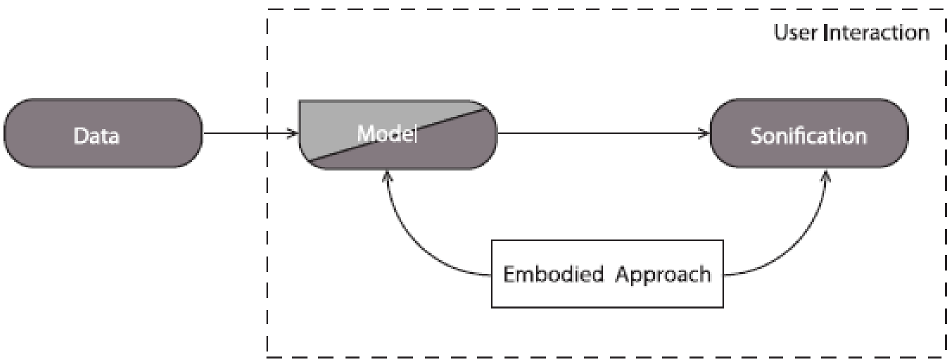


Figure 14. Two-way or interactive sonification. In this case, the model is flexible and can be altered by the user to get the most information out of the data (Diniz et al., 2012, fig. 2).

This interaction in the model and data mapping is likely to benefit the user by enabling them to make more effective use of the sonification. Additionally, this is a very similar idea to model based sonification (see Section 2.2.6.4). The inclusion of the user in the conceptual view of the sonification is very important. Interactivity in a general sense is important (as discussed earlier), but it is also important for the user to be in control of their interaction with the sonification and this allows them to get much more information out of the experience than if they were just passively listening to the sonification (Hunt et al., 2004).

Figure 15 (Saue, 2000) shows very neatly the different aspects that go towards creating a sonification. The two main aspects are the data that are being sonified and the auditory display, or the sound used to sonify the data.

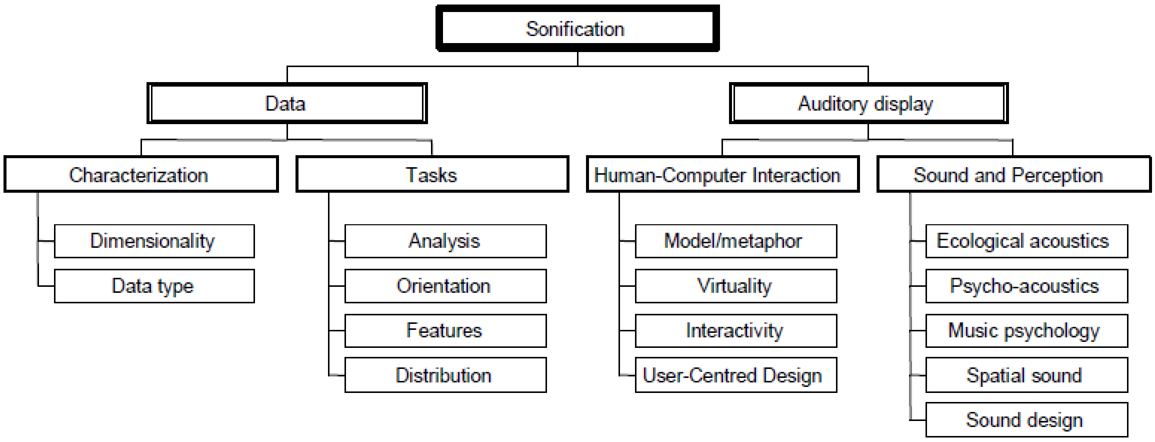


Figure 15. The different elements involved in a sonification (Saue, 2000, fig. 1).

The data has two separate elements, the characterisation and the tasks. Characterisation includes many practical aspects of the data (such as format, type, number of dimensions) as well as the conceptual model underpinning the data. The tasks element is the aim of the outcome of the sonification, i.e., what task (or tasks) will the sonification help the user complete? The auditory display element consists of the interaction by the user (human-computer interaction) and practical aspects of the sound (sound and perception). Human computer interaction includes issues related to the interactivity of the sonification and any models used in the explanation. Sound and perception includes the actual sounds used for the sonification, but also user specific issues of how they hear the sound and the ecological (i.e. environmental) issues of where the sound is played.

While there are a wide range of decisions to be made regarding which sounds to use and how to use them to represent the data, Flowers et al. (1996) amongst others make the recommendation to start with a simple sonification rather than a very complex one, as it is much easier to explain to novice users.

2.2.5 Functions of sonifications

The next two sections describe the different functions and types of sonifications. There are many ways of defining ‘sonification’ and categorising different types of sonification, such as by function, type or technology. The categories used here are by no means fixed and many of the examples mentioned may cover more than one category. There are many different ways sound could be used to represent a simple piece of information, which makes sonification a very flexible technique. However as there are many different ways to achieve the same outcome, it can also be very complex (Grond and Berger, 2011). For a more detailed overview of functions and types of sonifications, the reader is referred to the chapter by Walker and Nees (2011) in The Sonification Handbook.

Functions can be groups under five headings: alarms, monitoring, data exploration, entertainment / sport and art / public engagement. These groups are selected for ease of explanation, but not all examples fit neatly into these groups. There are other examples which will be mentioned when relevant and illustrate the wide variety of applications that sonification can be applied to.

2.2.5.1 Alarms, alerts and warnings

Alerts are one of the more traditional types of sounds used to provide information. However, they generally do not provide very much information; for example a door bell will indicate that someone is at the door, but not who or why. Alarms and warnings tend to provide some additional information, for example a fire alarm contains the information that the alarm is urgent and requires immediate evaluation. A well designed alarm will allow the user to understand what the warning is and how quickly they need to respond to it and this is particularly important if there are a number of alarms that may activate in any particular environment (Edworthy and Hellier, 2000).

2.2.5.2 Status, process and monitoring messages

These types of sound convey some information about the amount of a task that is complete or monitor a particular set of data to alert the user when it changes. These often make use of the ability of the listener to detect small changes in auditory events (Walker and Nees, 2011). The most common example of these is patient monitors in hospitals, where they can either alert the staff implicitly by reflecting a change in the patients status at an anaesthesiologist’s workstation (Kramer, 1994) or explicitly by activating an alarm when a patient’s blood pressure or other vital signs cross a certain threshold (Watson, 2006).

2.2.5.3 Data exploration

Sonifications for data exploration are one of the more common forms of sonification and attempt to represent a whole data set (or the most relevant parts of it) in a holistic manner, as opposed to the alarms and status messaging examples above, which condense the information as much as possible. When the term ‘sonification’ is used, it is often data exploration that is being referred to.

Data exploration is often thought of as the most ‘scientific’ of sonifications and usually makes use of a type of sonification called parameter based sonification (see Section 2.2.6.3 for details). These can vary from tools designed for a number of specialists users (e.g. MacVeigh and Jacobson, 2007) to those which allow an individual to explore a data set in order to discover more details and information about the data (e.g. Goddard, 2012). There have been attempts to create easy to use tools for applications in these areas, but often these are of limited use because of the variety of data sets that need to be presented and the wide range of file types people may use. This may explain the reasons for the fact that this was one of the earliest types of sonification implemented, but has progressed relatively little over the last 20 years of sonification.

2.2.5.4 Entertainment, sports and leisure

There are a wide range of uses of audio both in terms of audio-only examples and audio combined with visual stimuli. Audio-only games are of great use to the blind and partially sighted community and include simple examples such as the Towers of Hanoi (Winberg and Hellstroem, 2001) and more complex arcade games such as space invaders (McCrindle and Symons, 2000) and role-playing games (Liljedahl et al., 2007). A system for an ‘accessible aquarium’ has also been developed providing a sonic representation of the locations, movement and types of fish in an aquarium (Bruce and Walker, 2009). Audio is also obviously important in terms of music and radio as well as accompanying visuals in film sound tracks.

Sonification is also used in sport development including applications such as Sofirow (Schaffert et al., 2011) where rowers’ movements are sonified to help them gain a better understanding of their rowing stroke and to help improve performance. This sonification made a significant difference to the mean boat velocity and is being used by the German national junior rowing team to improve their performance. A similar piece of work with similar results has been conducted with speed skaters (Godbout and Boyd, 2010) and another study is in progress with swimmers (Hermann et al., 2012). A separate side to the body movement group is that for limb rehabilitation patients where stroke victims have to undergo specific training to improve their limb movement. The sonification gives them immediate feedback on their limb movement and they can compare their sonification to a ‘correct’ movement to help them work out which specific aspect they need to improve (Bruckner et al., 2011).

2.2.5.5 Art and public engagement

Art and public engagement utilises the novelty aspect of sonification to engage audiences who may not normally be interested in specific areas of scientific research. While the scientific knowledge communicated using some of these methods can vary from a reasonable amount to negligible, the sonifications generally result in an increased level of interest in the area and sometimes can be used to help explain complex scientific concepts. Fernstrom and Taylor (2010) used sonification to represent different levels of water toxicity and allowed users to sonify different water samples to compare the differences. This type of sonification has not been done before so they faced a wide variety of technical challenges, but by the end they had a working and usable installation. The discovery of the Higgs Boson particle created a wave of publicity, part of which included a sonification of the data from the Large Hadron Collider (LHC) at Cern that showed the existence of the Higgs Boson (BBC News, 2012). As well as public engagement, the team at LHC plan to use sonification to analyse some of the large data sets generated by the collider (Ghosh, 2010).

The art community have considered how sound can be used in artistic environments, such as sound sculptures and soundscapes (Sinclair, 2012). Their focus is on how the sound impacts the understanding of the art work and the viewers’ emotional state, rather than how it can be used to explicitly communicate data. A good example of this is a study by Carles et al. (1999) who asked participants to rate the beauty of a number of landscapes. They played different sounds with different landscapes and the sonic element had a significant influence on how participants rated the landscapes. ‘Soundscapes’ are also explored by Schafer (1994) who defines them as being any of the sounds that surround us in a particular environment and observes that while they may be less obvious than any other pollutants that can enter our bodies, sound pollution is no less important. A soundscape as part of an immersive virtual environment also relates to some of the work completed on using sound in virtual reality settings (see Section 2.3.3 for details). Sonification can also be used as an alternative representation for art. A group of Chinese scientists used sonification to allow people unfamiliar with ancient Chinese to be able to experience classical Chinese poetry though a non-speech medium (Huang et al., 2012).

2.2.6 Types of sonification

As well as categorising sonifications by their function, they can also be grouped by the sonification technique used. There are four main techniques used in sonification and each has its own conceptually different approach to how data is related to the appropriate sonification. None is necessarily better than the others, but they each have their own advantages and disadvantages and can also be used in combination with each other for particular applications (Hermann et al., 2011).

2.2.6.1 Audification

Audification is what might be called 'direct sonification' where waveforms of time based data are directly translated into sound (Kramer, 1994; Dombois and Eckel, 2011). The waveforms are often frequency or time-shifted into the audible range for humans. The aesthetic values of this output are often 'interesting' to say the least and can be difficult to listen to, but are useful to specialised groups of users. Seismic data are suited very well to this technique and experts are able to interpret the sound with up to 90% accuracy (Speeth, 1961; Dombois, 2002) and this can also be applied to well log data (Stewart and Brough, 2008).

2.2.6.2 Auditory icons and earcons

Auditory icons and earcons are similar in a number of respects, but the sounds used for auditory icons directly relate to the piece of information they are trying to represent, whilst earcons use a more abstract sound. Auditory icons were originally devised as the auditory equivalent to visual icons and examples include running water (to represent water) or a car starting to represent a car (Brazil and Fernström, 2011). Earcons often use musical notes to represent information, and exist in a structured hierarchy to allow different earcons to be related to one another (McGookin and Brewster, 2011).

"Earcons can be thought of as short, structured musical messages, where different musical properties of sound are associated with different parameters of the data being communicated." (McGookin and Brewster, 2011, p.339)

Earcons have gone through a series of developments, increasing their complexity and uses. They use a hierarchical technique to create a hierarchy of messages (such as a general computer error and different sub-types of computer error) and in conjunction with this a hierarchy of sounds. Figure 16 shows an example of this.

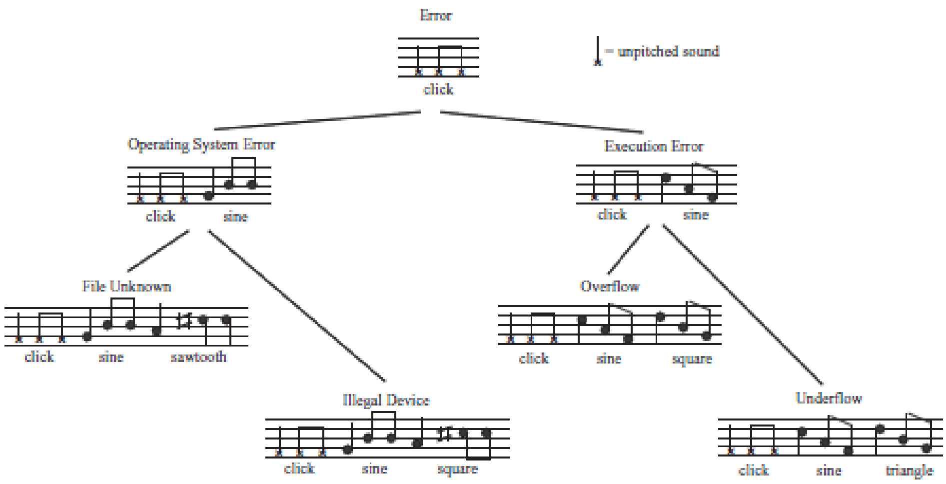


Figure 16. A hierarchy of sound applied in a sequential manner to indicate different forms of computer error. (Blattner et al., 1989, fig. 12; McGookin and Brewster, 2011, fig. 14.3).

These hierarchical earcons are significantly easier to learn than their non hierarchical counterparts (McGookin and Brewster, 2011) and in a particular instance when they were used to augment the visual interface, the participants made significantly fewer errors than when using non-hierarchical earcons (Leplatre and Brewster, 2000). They were also able to infer positions of unlearned earcons based on the hierarchical nature of the sound (Leplatre and Brewster, 2000).

Brewster (1996) also expanded the use of auditory icons from the use of error messages and menu items to buttons, scroll-bars and windows. The aim of this was to enable non-sound designers to add-in sound to their existing applications with the aim of improving performance for sighted and non-sighted users. This did reduce the number of errors the users made, but did not produce any broader insights on the usefulness of hierarchical earcons more generally, possibly because of study design limitations (Brewster and Clarke, 1997).

2.2.6.3 Parameter mapping sonification

Parameter Mapping Sonification (PMSon) is one of the more widely used types of sonification and involves the association of information with auditory parameters for the purpose of data display (Grond and Berger, 2011). How the data are mapped to the sound is very flexible and so this can be a very powerful technique, but it can also be very difficult to implement. This is because of the huge range of options that are possible using this technique, there are many decisions to be made about which particular aspects of data to sonify, which sounds to use to represent them, how that sound will change with the data and how (if at all) the user may interact with this.

PMSon is also one of the oldest types of sonification and many of the early examples mentioned in this thesis are this type of sonification. Some of the more detailed aspects (such as the sound synthesis and parameter mapping) will be discussed in more detail later in the thesis (see Section 2.2.7). This method is the best way of representing more complex data and is a technique that will be used in this study.

2.2.6.4 Model based sonification

Model based sonification takes Parameter Mapping one step further and builds an intuitively dynamic sonification model. The user then explores the data by interacting with this model, and this interaction process creates the sonification. Often the model is an instrument that the user can ‘play’ to explore the data. The user can also manipulate the mapping of the data to sound, both to explore different aspects of the data, but also to optimise the sonification model to their specific requirements. Very often these sonifications use data with a large number of dimensions and/or data points (Hermann, 2011; Walker and Nees, 2011). This area has significant potential but is currently relatively underdeveloped as there are few examples that utilise Model Based Sonification to its full extent.

2.2.7 Sound variables

In a very similar way to visual variables, there are a number of different parameters that can be varied in a sound. These sound variables include pitch, tempo and timbre and they can be used in a comparable way to colour, size, shape and so on in the visual domain. However, because of the relatively new nature of sonification, much less is known about these factors than their visual counterparts. The following section includes an overview of the most commonly used sound variables and some information about them. There are no doubt others that could be used, but the relatively limited literature on their use limits the amount that is known about them. Closely related to the sound variables themselves is how these variables are mapped onto data, which is considered in the next section.

The three most commonly used sound variables used in sonification are pitch, loudness and timbre (Neuhoff, 2011). However, all of these variables have some limitations, both practical and theoretical. There are also many different ways to think about variations of sound (as mentioned in Section 2.2.4) many of which are still being developed and explored.

The pitch (or frequency) of a sound is a fundamental feature and is often characterised as ‘high’ or ‘low’. The human auditory system is capable of detecting changes of pitch of less than 1Hz at a frequency of 100 Hz with a typical range of 20 Hz to 20 kHz (20,000 Hz) with higher frequencies the first to be effected by age related hearing loss (Carlile, 2011). Western musical notation uses pitch to distinguish notes and many sonifications utilise musical notes as their basis for data representation. The average person can distinguish between 48 and 60 different pitches over four or five octaves, which makes it ideal for representing a range of data (Yeung, 1980; Krygier, 1994). The use of different notes is a good way of making use of an existing framework and it is something that the majority of users are likely to be familiar with. However it does require fitting the data into the existing musical framework, which can limit the types of sound variation that can be used. Pitch is a very fundamental parameter of a sound, as when a listener hears different pitches, they are represented directly in the brain. Levitin (2007) placed electrodes on the brain and he could observe from the electrodes readings which pitches were being played to the subject.

Loudness (or volume) is another dimension that is frequently used, but it has limited success for four main reasons. Firstly, while differences between different volumes can be easily discriminated, there is a much lower resolution (i.e. number of discernible levels) than for pitch. Secondly, memory for volume is extremely poor. Thirdly, background noise and the sound reproduction equipment can result in the volume varying considerably, depending on the users' environment (Flowers, 2005). Finally, there are no existing cognitive structures for loudness, unlike pitch or timbre (Neuhoff, 2011).

Timbre is a relatively easy concept to understand (often explained as the quality of the sound, or how we can tell a trumpet and a piano apart when they are played at the same frequency with the same loudness), but it lacks a true, useful definition (beyond the explanation above) (Neuhoff, 2011). The easiest way to use timbre would be to use different instruments, which has the potential to work but timbre is not suitable for an ordered scale. This type of scale may be more complex to explain to the user and some people also have trouble telling certain instruments apart.

These different sound variables could be used in conjunction to show variations in two or more different variables, but the dimensions are not completely independent of each other. For example, an increase in pitch (but no change in volume) can be heard by a listener as an increase in volume (Neuhoff et al., 2002). While this effect can be used beneficially, by providing redundancy in using volume and tempo to represent the same information, using the two together to show different variables is unlikely to work.

There are a number of other ways of categorising sounds, with an often used one being natural and non-natural sounds (Susini et al., 2012) or abstract and realistic (Krygier, 1994). Some of these groups work better as a continuum than categorical groups but they can also be very subjective as one participant's definition of natural is likely to be different to another. Sometimes natural sounds are seen as more pleasant and useful than non-natural sounds (Susini et al., 2012), but this does vary.

2.2.8 Mapping data to sound

The process of choosing which sound to represent the data in a sonification is crucial in creating a useful and effective sonification. The mapping will also depend on the aim of the sonification, for example whether there are particular aspects of the data that are more important (or the designer sees as more important) than others (Grond and Berger, 2011). There is only a small amount of research on the process of mapping data to sound, but this does include some guidelines.

Figure 17 shows some of the process of defining and applying a sonification mapping. Many aspects of the data features section are dependent on the data that are being sonified, as well as the overall aim of the sonification. The sound synthesis area gives some information on the wide range of sound variables that could be used, but also how they relate to one another through a range of different 'levels': signal domain, sound object and auditory scene. These levels relate to where the transformation is applied to the data. The signal domain related to the data itself is the stream of numbers that can be manipulated in terms of frequency and time. The sound object consists of aspects of the individual sound that can be altered (e.g. pitch, volume etc.) and the auditory scene consists of how the individual sounds relate to one another when played in the final output. Choices at one stage will have an impact on the other stages, and it is this level of complexity that allows sonification to be flexible, but also means there is still a significant amount of research to be done in this area.

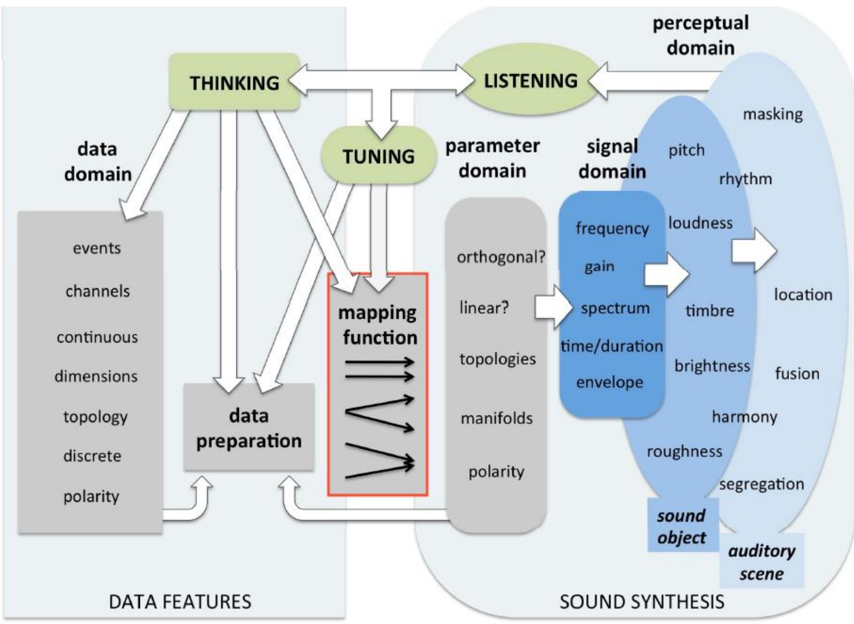


Figure 17. The process of converting data into sound involves a number of different processes and factors (Grond and Berger, 2011, fig. 15.1).

One of the earliest examples of mappings is that of the Geiger counter, where a faster tempo of clicks represents a higher level of radiation (Dubus and Bresin, 2011). The polarity of this mapping was a result of the process used to create the sound (more radiation leading to more electrical impulses leading to more clicks) rather than an explicit decision, but it does work very effectively, possibly because it is directly related to the feature it is representing.

The actual process of picking a specific mapping for a data set is best done through experimentation, with reference to the data set, sonification setting and user requirements (see Walker and Kramer, 1996, 2005 for a good overview of the issue). While the inputs to choosing the sound used in a sonification may be complex, it is also very important to make the correct choice. While the correct choice may not be immediately obvious, a badly chosen sound or mapping will have a substantial impact on the effectiveness of the sonification (Bennett and Edwards, 1998; Jeong and Gluck, 2003; Zhao et al., 2008).

Walker has completed a significant amount of research in this area looking both at what sonic variables to match to which types of data, and which polarity to use in the mapping. Polarity is the orientation of the data to the sound, so for example whether a high pitch represents a high temperature or a low temperature. Polarity can have a significant performance impact if the polarity of the mapping is not what the user expects (Walker and Nees, 2011). Of the research that has been completed in this area, a number of consistent patterns have emerged (see Table 1), with an increase in pitch representing an increase in temperature effectively, but an increase in size is best represented by a decrease in pitch (Walker, 2002). These are findings from Walker's research, but they are not unanimous across all users. It is very dependent on the mental models that users create (Walker, 2007) and users with visual impairments sometimes chose the opposite polarity to those without visual impairment (Walker and Lane, 2001). Pitch is a well used variable, but its meaning can vary depending on the context. Rigas and Alty (2005) used pitch to help visually impaired users understand graphics information and found that a rising pitch successfully communicated higher values.

Data	Best Sound Variable	Best Polarity
Temperature	Pitch	Increase Pitch = Increase Temp
Size	Pitch	Increase Pitch = Decrease Size
Pressure	(no good variable)	NA
Velocity	Pitch and Tempo	Increase Pitch / Tempo = Increase Velocity
Dollars	(no good variable)	NA

Table 1. Summary of Walker's research into polarity mappings (Walker, 2002; Walker and Nees, 2011).

When choosing the sound variable to map to a specific dataset, it is important to evaluate the mapping with the target audience to ensure it is the best mapping to get the best performance out of the sonification. Walker (2002) suggested a minimum of 20 participants to test a new mapping, but this will depend on the data and context of use. It has also been observed that the naturalness of a sound influences the perceived usability and pleasantness of sonic feedback (Susini et al., 2012).

Dubus and Bresin (2011) reviewed 64 different examples of sonification and categorised the type of data and the type of sound variable used to represent that data. They found preference for pitch above other variables (as mentioned earlier) but also found trends for specific sound variables (e.g. spatialization, or spatial sound) to be used for specific types of variables (e.g. kinematic quantities). They did not have a large enough sample to draw any conclusions, but did make some interesting observations.

2.2.9 Variation in understanding of sound

The sections above have discussed some of the different sounds (and differences in sounds) people can hear. Much like interpretation of visual symbols, this does vary between users. Some of these aspects (such as the loss of ability to hear high frequencies with age related hearing loss, as mentioned earlier) are relatively well understood. Some of the aspects that are dependent on cognitive understanding (as opposed to physical changes in the ear) are more complex and have not been researched so extensively.

One issue to be aware of is amusia, which is when suffers cannot tell the difference between different pitches. Amusia is not well understood, but a survey of 600 participants in the UK found approximately 4% who could not near anomalous pitches inserted into popular melodadies (Kalmus and Fry, 1980). A similar proportion (5%) was obatained in the US with a similar test (Peretz and Hyde, 2003). It is unlikely that a complex sonification of any sort would be understandable to sufferers of amusia, whether it uses musical notes or not. Many people claim they are ‘tone deaf’, but this is not amusia they are referring to, as they can still fully hear and enjoy music. What they are likely to be suffering from is an inability to sing (i.e. ability to reproduce specific notes on demand), which will not affect their ability to use a sonification.

Training is an important aspect, both in terms of general training with sonification as well as specific training on a particular sonification. Extensive training for visual data representation (e.g. graphs) is included in the school curriculum, but training for sonified data is not, which is likely to impact effectiveness (Davison, 2011, pers. comm.). Training on a specific sonification is also important, whether this is training before using the sonification or improvement as a result of repetition. Blades et al. (1999) observe that blind users who were not taught how to use tactile maps performed less effectively with them than those who were taught how to use them.

Additionally, Flowers et al. (1996) found that user performance with repeated exposure to a particular sonification improved performance.

It is often suggested that participants with greater musical experience or ability will be more effective than those without this experience. However, this has not been demonstrated in the vast majority of studies (Walker and Nees, 2011). There are a number of issues with how to measure musical experience and ability, which is likely to confound this relationship (Edwards et al., 2000). It has been observed (anecdotally) that blind users with musical experience expected to be more effective at using a sonification than blind users without musical experience. However, when this was tested, no difference was found between the two groups performance (Bujacz, 2012, pers. comm.).

Age related hearing loss effects the abilities of the listener to hear higher frequencies, as discussed earlier. A neurobiology study found that people who play musical instruments reinforce certain neural pathways which result in a delay to this age related hearing loss. This study was only related to perception of speech, but highlighted some interesting findings (Parbery-Clark et al., 2012).

Different users may prefer different sounds in a sonification and additionally different sounds may be more or less appropriate in different locations (Bonebright and Nees, 2009; Bonebright and Flowers, 2011). Therefore it may be useful to have a number of different preset options which the user can switch between until they find a combination of sounds that works well for them.

Given that repeated exposure to a technique (sonification, visualisation or otherwise) leads to an increase in performance of using that technique (Flowers et al., 1996) it may be that a user's learning style would have an impact on their performance. Therefore users with a predisposition to sonifications would perform more effectively with sonifications. Learning styles or cognitive styles are described as a preference for a particular way people learn and different people have different learning styles. The classification of styles varies greatly (Coffield et al., 2004), but often feature visual, auditory, reflective and interactive methods. There is a wide range of literature on learning styles ranging from complete endorsement to complete dismissal and everything in between (Kolb, 1981; Felder and Silverman, 1988; Riding and Cheema, 1991; Healey and Jenkins, 2000; Coffield et al., 2004; Healey et al., 2005; Hawk and Shah, 2007; Pashler et al., 2008). Academic development of this area has been patchy, mainly due to the fact that many different academics have been working on different groups of learning styles and have very infrequently referred to each other's work (Riding and Cheema, 1991). However, even with this and the criticism mentioned above, both sides agree it is a sizable commercial industry and popular with both teaching and businesses alike.

2.2.10 Why is sonification useful?

Sonification is useful as a way of reducing visual overload in interface design. Visual overload results from the ever growing demand to display more and more data (Frauenberger and Stockman, 2009). Auditory displays can help to overcome some of the limitations of visual displays and reduce cogitative overload by using sonification to transfer some of the information from the visual element of the display to the auditory element of the display. This multimodal interaction results in better performance from users, particularly when faced with more complex tasks (Oviatt et al., 2004). Sonification can also be used to replace visual interfaces when they are not appropriate, such as devices that are too small to have a visual display or devices that are used in situations where a visual display is not appropriate. Audio shows great potential in this area, with the ability to supplement a smaller visual display than normally possible (Brewster, 2002). It is also key to making technology more accessible to the visually impaired. While the technology exists in most mobile phones, portable music players and kitchen appliances to make pleasant and helpful sounds, they very rarely use these sounds. Due to this, Frauenberger and Stockman (2009) argue that this is a design problem rather than a technical problem.

One big advantage of sonification is its ability to exploit some of the advantages of the human hearing system, which can perceive highly transitory events and can hear some forms of repetition and correlation (Shepherd and Bleasdale-Shepherd, 2008). Additionally, listeners can 'tune-out' certain sounds in a sonification and will only hear them if they vary (Eldridge, 2006). These techniques make sonification more useful for some specific data sets than visualization, because they exploit the human sensory system in a more effective way. Some sonification techniques are quicker to learn compared to other representation methods. For example, when using auditory graphs, after 20 repetitions discrimination accuracy had risen from 81.3% to 87.2%, whereas for tactile graphs this figure rose from 81.3% to 81.9% (Flowers et al., 1996).

One use of auditory displays is to allow some tasks that would normally occupy the visual sense to be moved to the auditory system to reduce visual overload. This could either be achieved by moving one task completely (so it is just represented using audio and not visually) or by creating redundancy by representing the data visually and sonically at the same time. Transferring the task would allow the user to use their visual perception for a different task and redundancy may increase performance. However, results are mixed in the effectiveness of this approach. Seagull et al. (2001) compared visual only, audio only and redundant methods (both visual and auditory) and found that the auditory only interface was slower than visual or redundant. This may be because of the complexity of the sound or the lack of specialist skills and experience of the evaluators for the task.

Sonification can also be used to reinforce data shown visually, rather than moving it to a different sense. This makes use of dual-coding theory which says that if a variable is reinforced in an interface it will be more easily remembered. In fact, this is already done using two visual variables to show one piece of information to great success. Using two different sound variables to the same effect is conceptually possible, but a sonification using this process is more difficult to understand (Flowers, 2005). Sonification is also useful in a range of environments as well as for the representation of spatial data (see Section 2.2.5).

2.2.11 Limits of sonification

While sonification has many advantages, it also has some limitations. One major difference between sound and vision is how they are interpreted by the user. Spatial data shown visually can easily be referred to multiple times by the user, whereas sonic data cannot easily be examined in the same way. If the sonification is a recording, then the clip would need to be played again. If the sound is generated from moving a mouse over a map, then it is easier to access than a sound clip. However, the mouse still has to be physically moved to the point of interest as it cannot just be glanced at, like the visual medium (Rice et al., 2005; Bonebright and Flowers, 2011).

One way around this with spatial data would be to use eye tracking hardware to locate where the user was looking and use this to drive the sonification. However, eye tracking hardware is expensive and can be difficult to set up and it is unlikely to be available to the end users of this tool. Evreinov (2001) uses eye tracking to record user's eye movements and employs this to drive a sonification of an image that is being viewed. He experimented with different sound mappings, but ultimately struggled with the complexity of the task he was completing and did not provide any evaluation of the technique.

It is often thought that using an auditory display will prevent the user from being able to listen and respond to other aural inputs, both speech and non-speech (Bonebright and Nees, 2009), however this is not necessarily the case, but there has been very limited testing in this area. When listening to speech it is true that for each extra voice added on top of the current voice(s) played they become harder to understand. Performance is surprisingly accurate with up to five concurrent speech signals, but deteriorates rapidly after this (Bonebright and Nees, 2009). Even when represented with an unusual task involving auditory graphs, Peres and Lane (2005) found that participants' performance, whilst initially poor, soon returned to an acceptable level (after around 25 task repetitions).

2.3 How do Vision and Sound Influence the Interpretation of Spatial Data?

2.3.1 Theoretical background

Sonification has many different applications over a wide range of areas. One of these is using sonification with spatial data, which has had relatively little work compared to some of the other applications of sonification. Krygier (1994) was one of the first to write a review of the potential for sonification of spatial data, and he freely admits the possible clash between the very visually based world of cartography and GIS and the mainly non-visual world of sonification. Krygier created a list of nine different sound variables, or abstract sounds as he calls them. He also observes whether they are suitable to represent ordinal or nominal data. Most of the variables (including volume and pitch) can be clearly ordered and so can be used to represent ordinal data. The only one he lists as something that is best suited to nominal data is timbre (see Figure 18).

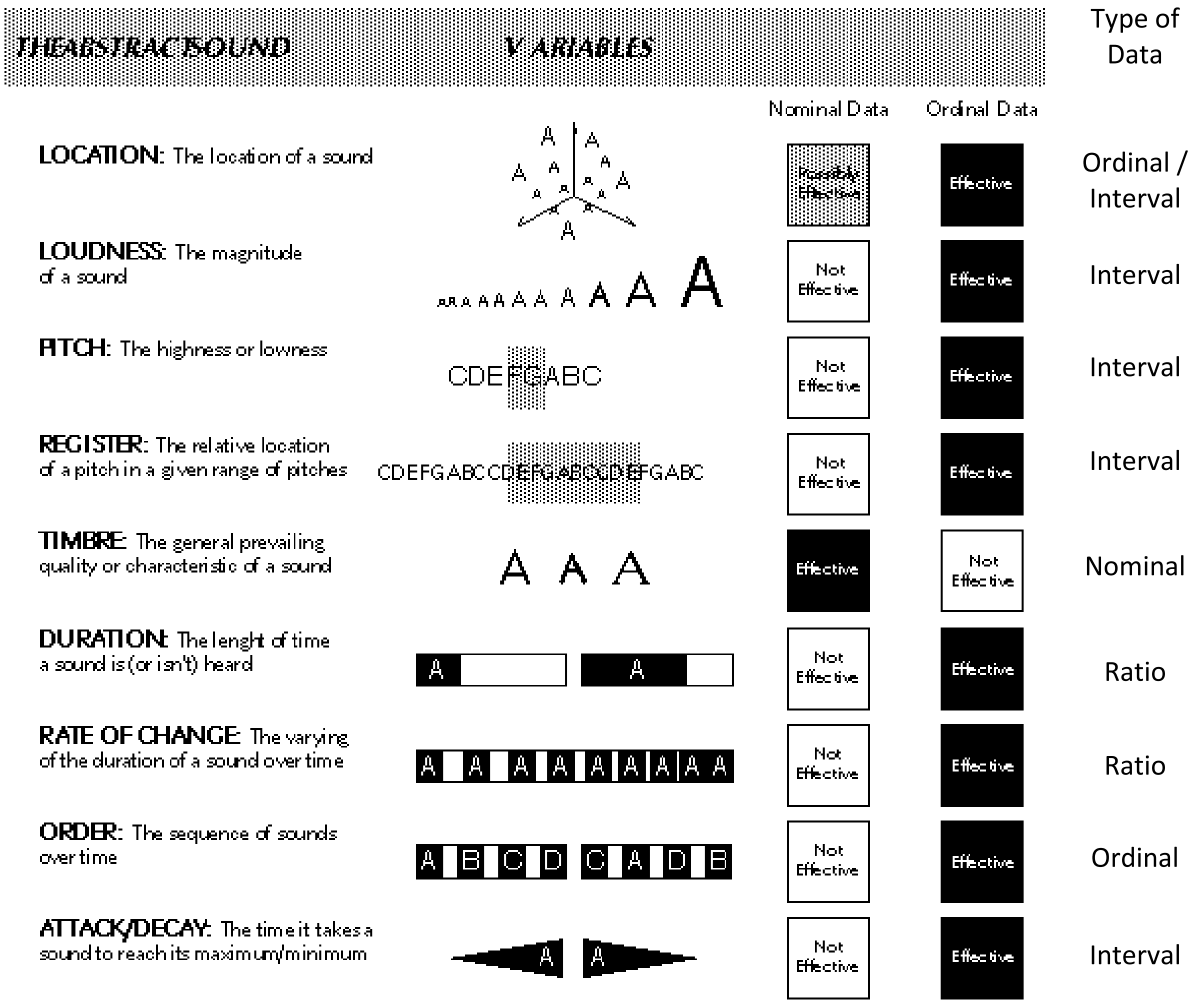


Figure 18. The nine sound variables Krygier identifies and the type of data that can be represented using them (after Krygier, 1994, fig. 1). The far right-hand column lists the type of data that can be representing using that type of sound.

The location of a sound refers to the location of a sound in 2D or 3D space. Left/right panning of sound is relatively easy to do (using stereo) and basic forward / back is possible using a home cinema setup (5.1 surround sound). However, more complex location is difficult to represent. Humans often become unconscious of a sound with a constant volume, but can very easily notice a variation in volume. For example a computer's fan soon becomes inaudible after switching the computer on, but even a slight variation in the fan speed will be noticed very easily (Buxton et al., 1985). This can be exploited by sonifying data which is full of complex variations which are easier to distinguish sonically than visually. Krygier views pitch as one of the most effective ways of differentiating order with sound and according to a survey by Dubus (2012) it is one of the most frequently used. Pitch is also directly represented in the brain which reduces the cognitive load when it is processed (for more details, see Section 2.2.7). Register is the relative location of a pitch within a given range of pitches and would most likely be used to represent a variable in combination with pitch. Timbre is defined as the difference between different instruments when they are played at the same pitch and volume and is difficult to quantify. It has no inherent order within it, so works well with nominal data. Attack has been found to be much more successful in conveying information than decay (Lunney and Morrison, 1990). Attack could be used to represent the spread of values, with a long attack representing a large spread of values. However, it would take time for each individual value to be played, particularly if they had a relatively large attack.

Some of these variables are more frequently used than others in sonifications and different variables are more appropriate in different circumstances (as listed in Figure 18). Krygier (1994) categorises many of the sound variables as suitable for ordinal data, but some of them are appropriate for ratio data as well, depending on the user. For example, if a user can tell what pitch a particular note is, or at least how far along a scale the note is then this does give them some knowledge of a ratio scale. It would only be the case that they are limited to an ordinal scale if the user can tell whether a pitch is higher or lower, but not how much higher or lower. This is known as relative pitch, and is present among all trained musicians and a significant proportion of the population in general (Levitin and Rogers, 2005). However, this would not work for people suffering from amusia (see Section 2.2.9). A similar issue applies to location, where the user's ability to judge the location will have an effect on the type of data that can be displayed using location.

2.3.2 Spatial data examples

There are a wide range of sonification examples, however relatively few of these use spatial data. This section performs a review of the studies that use sonification in conjunction with spatial data.

The first section covers the pieces of work that are just theoretical discussions and prototypes. The tools these pieces of work created have not been tested in any formal way, so it is not known how effective the sonifications are. However, they do have some useful contributions in terms of sonification design. The second section looks at those which do perform a user evaluation and draws out the relevant results from them.

Cassettari and Parsons (1993) looked at the concept of using sound in a GIS, both as a generic way of representing data, but also as a way of representing sound data in a GIS environment, such as noise pollution. They also make the point that using sound in this way is a very complex process. Schröder (1995) included audio as an option in a scientific visualisation tool and made the point that customisation is very important in a data visualisation environment (both customisation of the visual and audio elements) and it allows much more effective and more flexible use than if the customisation was not present. Medvedev (2011) attempted to address the issue of the limited amount of information that could be communicated using sound. She used a process called additive synthesis to create a number of different representations of a visual photograph, and then combined them into a complex sound form. It is unclear whether this was actually effective though, as no evaluations were completed. Smith et al. (1990) and Erbacher et al. (1995) discuss the Exvis project which combines advanced (for its time) visual and auditory variables to display multidimensional data sets. The project had great potential in terms of novel sonification, but they did not complete any evaluations with it so it is not known how successful this method of sonification was.

Fisher (1994) created an example using sound to represent the likelihood of correct classification of each pixel of a satellite image. The area of interest could be highlighted with the keyboard arrows, or an 'overview' of the whole data set could be played, which would play the sound for the whole data set, passing over each pixel in turn. However, this was very limited by the computer and sound technology available at the time, restricting both the sounds that could be used and the data that could be represented. Veregin et al. (1993) developed a similar tool with a more limited implementation of the sonic aspects within a wider data exploration tool. Lodha et al. (1997) created a tool called LISTEN, a modular system which allowed different data sets to be represented sonically. The modular system made it possible to use LISTEN to create a wide range of different sonifications with almost any dataset, suitable for any individual. However, this flexibility also made it very complex to use and no other authors appeared to use it to sonify their own data (see also Lodha, Wilson, and Sheehan 1996; Wilson and Lodha 1996). Yinnan et al. (2008) implemented a form of sonification in a CAD (computer aided design) environment, but again it was not clear if this was successful as there was no evaluation. Meijer (1992) used a different approach and attempted to sonify visual images for blind users. The resolution was limited to 64x64 pixels with 16 grey tones per pixel and the approach suffered from needing to transmit too much information for the sonification process to be effective. Paolino et al. (2009) used sound to reinforce information shown visually on a mobile phone, allowing the user to have access to the information even if they could not use the screen at a particular time. Flowers et al. (2001) sonified a large data set of daily weather information, using MIDI stringed instruments to represent the data. They allowed the user to choose what rate to hear the sonification at (e.g. two days per second) and this was important to ensure the user could understand the data. A more recent example is MacVeigh and Jacobson (2007), who created a sonified map in Java, with three different types of land cover (sea, land and harbour) and used waves, birds and a ships horn respectively to represent these. This study was a proof of concept and was not formally evaluated as it was only tested by two of the authors' colleagues. They did specifically highlight the need for a sonification tool that works within a standard GIS environment to make user evaluation easier.

Table 2 summarises the literature evaluating the use of sound to help data presentation in a traditional desktop GIS setting, excluding literature that did not undertake evaluations. The studies cover a wide range of applications, which does limit comparability, however some findings can be generalised to other sonifications. Three of the examples are tools designed for blind or partially sighted users, but some of their findings are relevant to a wider audience.

Many of the studies make reference to the limited user evaluations completed in this area, with Jeong and Gluck (2003) making specific reference to the importance of completing more user evaluations. The studies in Table 2 do complete formal user evaluation, but only four of them have more than 12 participants. Evaluations with smaller sample sizes (e.g. 12 or less) do collect very useful information, but the small numbers limit the statistical validity of most significance tests. Having larger sample sizes addresses this issue and enables easier detection of significant associations or differences.

Within each evaluation, a training session was often important to familiarise users with the concepts and implementation of the sonification (Blades et al., 2002; Nasir, 2009). The concept of redundancy in representation has been discussed already (see Section 1.2.3) and is shown to significantly improve performance (Jacobson and Kitchin, 1997; Constantinescu and Schultz, 2011). While Jeong and Gluck (2003) found that the option with redundancy was not the most effective, the participants in their study did prefer it over the other two unimodal options. However, a common finding among the literature is the need to ensure that the sonification is not too complex, as users often struggled with examples using complex sonification techniques including one complex implementation of spatial sound (Gluck et al., 1999; Nasir, 2009). Complexity in itself is not bad as Pauletto and Hunt's complex sonification of EEG data was the most effective in their evaluation. However, more complex sonifications can have more expensive

Study author and overview	Participants	Sound	Key findings
Jeong and Gluck (2003) Identify highest / middle value census data among 9 US states using sound	51 students, on-campus, visually able	Volume of unspecified musical clip	User preferred combined option. Haptic alone more effective Musical background may have impact for tempo factor (pilot) User testing is very important and more needs to be done
* Jeong and Jacobson (2002) Performance of multimodal interfaces, asked users to order values	visually able	Musical note (piano)	Musical performance (being in an orchestra) significant Musical competency (play a musical instrument) effects time
Harding and Souleyrette (2010) Sound and haptic engineering application for highway planning	12 civil engineering students – vision normal	Real world sounds, pitch, volume, pitch & volume together (MIDI, type of sound not specified)	Allowing user to choose sound mapping is important Two distinct groups, but both equally successful Users preferred to look at more layers visually
Gluck et al. (1999) Ordering and analysis task of spatial risk data	10 non-GI users	Tone, volume, timbre, number, duration (type of sound unspecified)	More sound options would be better Users found the unstructured task difficult
Nasir (2009) Sonification of contour data using spatial sound	45 normal vision	Spatial and non-spatial sound (piano)	Training session very important Simpler method was easier to understand Complex method was too difficult for many users
Constantinescu and Schultz (2011) Looks at redundancy within auditory displays	7 participants, 5 non-blind, 2 blind	Pan, pitch and tempo (piano & guitar)	Redundancy resulted in better performance Participants had different approaches, but both successful Repetition leads to improvement for some participants
Pauletto and Hunt (2009) A biofeedback example of sonification	21 participants	Sin oscillator, uses different frequencies for different muscles	Interaction improves usability

Study author and overview	Participants	Sound	Key findings
Zhao et al. (2008) Compare Excel and sonified map of population data for US states	7 blind users	5 violin pitches and speech panning	Multiple levels of data are important Interactivity is an important part of the tool
Blades et al. (2002) Visually impaired learning novel walking routes	38 visually impaired	Verbalisation	Repeating evaluations means the user retains more knowledge Creating a physical model improved performance Training session is important
Loomis et al. (1998) Navigation system for the blind	10 visually impaired	Speech that appears to be coming from specific locations	Virtual acoustic display worked best
Jacobson and Kitchin (1997) GIS for visually impaired and low vision	Blind students	Audio tactile maps (verbal)	Sound and tactile combined most effective
Milios et al. (2003) Locate horizontal and vertical directions	10 normal vision 2 blind	MIDI, piano notes	Repeated use of user interface improves performance More evaluations required
Tzelgov et al. (1987) Different methods of showing radiation level	12 subjects	Visual, audio and combined methods	Audio only was most effective for search task

Table 2. A table summarising the user evaluations completed using sonification of spatial data. *(Task I only, Task II is Jeong and Gluck, 2003)

hardware requirements, which can be a limitation (Loomis et al., 1998; Horowitz, 2007) and can be confusing for the user if they are not implemented effectively.

The literature shows that allowing the user to interact with the sonification in a dynamic way (Pauletto and Hunt, 2009) and being able to navigate through the data using multiple levels is important (Zhao et al., 2008) as it allows the user to make the most effective use of the sonification. Not all users necessarily use the same approach when using a sonification and it is important to give them flexibility to choose or alter the interaction method to suit themselves even if the reasons for the different approaches are not always clear (Gluck et al., 1999; Harding and Souleyrette, 2010; Constantinescu and Schultz, 2011). One commonly evaluated factor that may influence this and/or general performance (measured through performance or completion time) is musical background (Jeong and Jacobson, 2002; Jeong and Gluck, 2003). Another possible influence is age, because of its effect on age related hearing loss although there were not enough participants in the study by Constantinescu and Schultz (2011) to be confident of this. It is also commonly seen in the literature that participants who used the same sonification a number of times increased their performance, although this learning effect may be limited (Blades et al., 2002; Milios et al., 2003; Constantinescu and Schultz, 2011).

2.3.3 Virtual environments and gaming

An academic context is not the only use of sonification. While it is not termed as 'sonification', computer gaming has made extensive use of sound in its immersive environments. Sound has also been explored to a lesser degree in virtual environments, but with greater penetration in the commercial sector than the academic one. Video games and GIS are also related, with a spatial representation of a particular area important to many video games (Shepherd and Bleasdale-Shepherd, 2009). Video games have been thought to have an impact on spatial cognitive skills for a number of years, with Gagnon (1985) showing a relationship between the two, with increased video game exposure correlating with an increase in spatial cognitive skills.

Sensory feedback is vitally important in video games, and interactivity is crucial to this, both in video games, virtual environments and scientific visualisations (Haase et al., 1994). There are many sensory feedbacks in games, with a significant number using multiple senses (e.g. visual & audio or audio & kinesthetic) to represent one action. However, within each sense category, the feedbacks used are often relatively simple, a) to limit the amount of information presented to the player and b) from practical limitations in terms of available technology (i.e. not all technologies are available on all platforms, so for multi-platform games, the technology used has to be present across 3 or 4 platforms (Shepherd and Bleasdale-Shepherd, 2008)). While there are a variety of examples in videogames of multi modal interaction, there is no formalised framework for multimodal interaction in videogames. This is because the available resources for game development do not allow time to be spent on activities such which are not directly related to the game itself. Additionally games need to sell to a wide audience to be cost effective, which means that the games have to work on a wide range of platforms. Developing a new multi modal technique that relied on a specific piece of hardware would be difficult to make a commercial success unless the hardware already had a significant market penetration. Additionally, often the visual and audio designers on a game work separately, with their work not being integrated until the final stages, which limits the interaction between the two areas. While the examples of multi-modal interaction in video games work very well and may appear to be relatively advanced, the range of signals used is actually relatively limited, and there is no theoretical background to them which limits their transfer to academic research (Shepherd and Bleasdale-Shepherd, 2008).

In a computer game environment, ease of use is very important. If the player cannot work out how to perform a particular move or task, they are likely to give up on the game very quickly (Pagulayan et al. 2003). A simple, easy to learn technique is better (at least in the initial stages) than a complex, difficult to learn technique. Complex techniques can be used, however it would be best to introduce more complex techniques after users have mastered the basics. This is the approach computer game designers have employed for a long time, and the same approach should be used with sonification; start simple, then become more complex.

There is a significant overlap between computer games / virtual environments and the discipline of geography, particularly with the potential benefit for field trips. A computer simulation of field trip locations (often created using a gaming engine, which allows greater detail and faster creation times than traditional virtual reality software) allows students to explore the environment before the actual field trip (Ashfield et al., 2010). Incorporating other senses (such as sound) into a virtual environment would improve the realism, but implementing it in a convincing way is complex (Anderson and Casey, 1997). Research has shown that the presence of high-quality audio information alongside a high quality visual display increases the perception of quality compared to the visual stream alone (Storms and Zyda, 2000). Visualisation has reached a level of maturity where the easiest way to improve the realism of a visualisation is not to simply increase the resolution, colours, textures or level of detail. We need to think more about what we are visualising and how this impacts the user (Lange, 2011) and this may result in visualisation experts exploring the other senses to see how they can contribute to the immersive experience.

Shepherd and Bleasdale-Shepherd (2008) consider how data could be represented using a multi-sensory method rather than the current predominately visual method often employed in GIS systems and VR (virtual reality) environments. They use the video game industry as an example of how audio and haptic interfaces can be exploited, in addition to visual methods, to enhance the immersion of the player. While the aims of GIS and VR systems are very different to the entertainment and immersion that games are trying to achieve, there is overlap between the different areas. For more details about multi-sensory representation in general, see Section 2.3.4.

Hetherington et al. (1993) compared different representations of a waterfall, with and without sound. The presence of sound in the representation made the observers view the representation as significantly more realistic. This effect was greater than the difference seen between a static and dynamic image of the waterfall. While virtual environments can provide a very realistic environment for entertainment or academic use, it is important to be able to navigate around within these environments. Perception of distance and orientation is often complex in these environments, significantly harder than in the real world (Riecke and Wiener, 2007). Markers and colour references can assist with addressing these navigation issues (Rousell et al., 2008), but it has also been observed that spatial auditory information improves navigation (Grohn and Lokki, 2003). The work by Rousell et al. (2008) involved a comparison of audio, visual and audio-visual navigation techniques around a 3D molecule in a virtual environment, so may have limited applicability to geospatial data. However, it certainly shows that using audio as a supplement to visual navigation may improve performance in a virtual environment (Grohn and Lokki, 2003).

2.3.4 Using vision and sound together

The previous sections have looked at representing spatial data using vision and using sound. This section will consider how both vision and sound can be used in combination to represent spatial data. Traditionally spatial data are represented visually, but there are benefits from using more than one sense in combination. There are two main reasons why this would be useful; firstly sensory substitution or replacement and secondly sensory combination for reinforcement or enhanced data representation.

Substitution or replacement may be useful to provide an alternative sensory source for impaired users (including examples such as blind or visually impaired users) or to circumvent poor operation conditions (such as low light levels or loud environmental noise). At the simplest level sound can be substituted for vision (or vice-versa), but this often is not effective and has severe limitations in terms of the amount of information that can be represented. For more effective use the interface (and often the data) needs to be redesigned for a different sense, and there has been a reasonable amount of success with navigation tools for blind or partially sighted users (see for example Jacobson and Kitchin, 1997; Loomis et al., 1998; Bujacz and Strumiłło, 2007) as well as other applications.

Sensory combination is more complex than substitution, but can also be much more powerful. Reinforcement allows the importance of a particular variable to be emphasised (Shepherd and Bleasdale-Shepherd, 2008) and redundancy allows user performance to be enhanced by reinforcing the presentation of a particular variable (Jeong, 2005). Enhanced data representation is feasible, which would either allow more variables to be represented than practical using one sense, or reduce the cognitive load on a user by splitting existing variables between different senses (Shepherd and Bleasdale-Shepherd, 2008).

There are different implementation approaches to using different senses; they can be used in conjunction at the same time (such as described above) or either switching between senses rapidly to allow the user to 'see' the data in as many different ways as possible, or switching between senses on demand so the user can explore which sense is best used with which data. No single option is best in all circumstances, but different ones will be more or less appropriate and useful at different times (Shepherd and Bleasdale-Shepherd, 2008). Careful design of the combination of two different stimuli is required, as since if the visual and auditory display conflict with each other, they can confuse the user, resulting in poorer performance (Jeung et al., 1997). Many studies only evaluate one particular combination of these options (because of very reasonable time and cost limitations), but this does mean it can be difficult to compare different pieces of work and relate them to one another (Frauenberger and Stockman, 2009).

When a fundamentally different method of data representation is used, methods and guidelines for using this new method need to be developed (Shepherd, 1995). Shepherd discusses the advantages and disadvantages of using animation (or dynamic visualisation as he calls it) to represent spatial data with temporal information. There are many advantages, but also many questions about how best to use this new (at the time) tool, and this requires much more research to work out which methods work and which do not (Shepherd, 1995). If this is not handled well then "complexity and ambiguity threaten the viewer where there is undisciplined use of new visualisation techniques" (Shepherd, 1995, p.169). Sonification is at a similarly new point and requires a similar exploration of the possibilities, without overwhelming the user with the possible options.

2.4 How does Uncertainty Influence the Interpretation of Spatial Data?

As already discussed, understanding the world around us is crucial to enable us to operate effectively. Being able to talk about the world and direct people where to go is very important to being able to work together in groups and achieve more for the group than would be possible individually. Being able to represent location visually, for example by using a map, is an important part of this.

2.4.1 Uncertainty in spatial data

Maps have been around probably as long as humans have been communicating about places, developing from the first hand drawn lines in the earth. In any map, whether it is a simple line drawing or a complex full colour map, there is a limit on the amount of information that can be displayed. It is not possible to represent every real world phenomena in full detail on a map – there is simply not the space available. The process of spatial data collection allows the geographical environment to be reduced and simplified to allow it to be represented on a map. How the information is reduced and simplified depends on the aim of the map, which can vary substantially between users (Robinson, 1995). Figure 19 shows the flow of information from the geographical environment (the real world) through to the map image, and how information is lost at each stage.

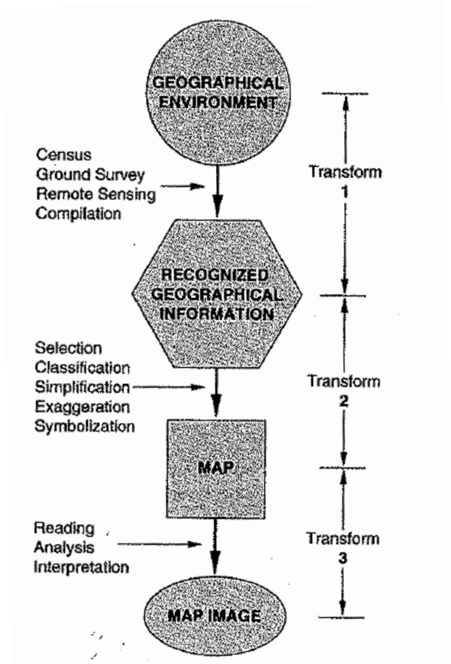


Figure 19. The process spatial information goes through from the real world to a mental representation of a map (Robinson, 1995, fig. 2.10).

As not every detail of information can be included, there is always some uncertainty in the resulting geographic information. There is also unlikely to be consensus about which details should be included and which should not; therefore the map (or spatial data) output is a product of the person doing the mapping, who makes decisions about what should and should not be included.

As implied above, the process of conceptualising, recording and analysing spatial information is a multistage process and uncertainty can occur at any of the stages. As shown in Figure 20 different types of uncertainty can occur at particular stages of the process.

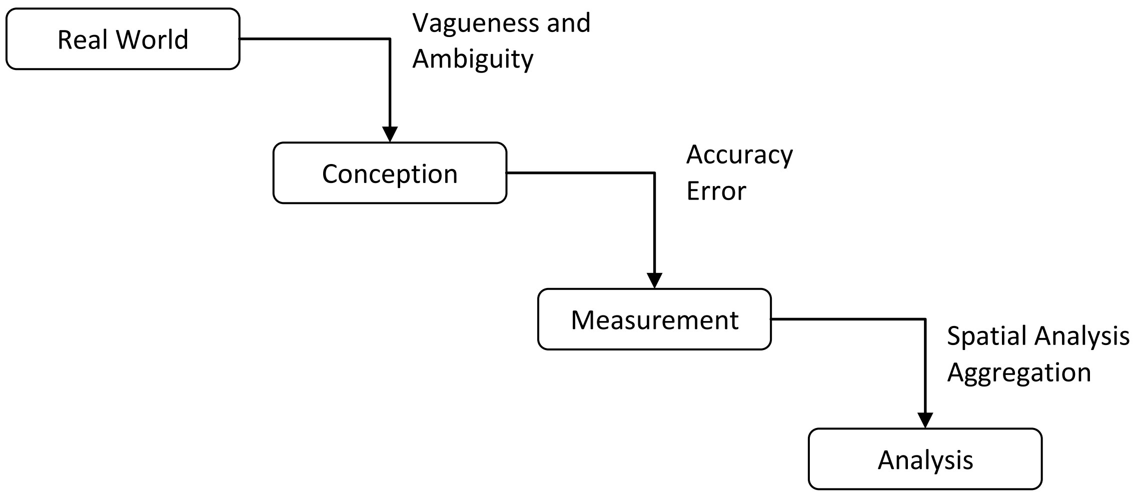


Figure 20. A conceptual view of uncertainty, showing the different places uncertainty can enter geographic data (after Longley et al., 2001, fig. 6.1).

2.4.1.1 Uncertainty in the conception of geographic phenomena

There are two fundamentally different approaches to conceptualise geographic information within a GIS. These approaches either use objects or fields as a data model to capture geographic information (Heuvelink, 1998). The object approach takes a relative perspective and considers spatial data to be a collection of simple objects (i.e. points, lines or areas) that are characterised by their spatial and non-spatial attributes. Examples of objects include land parcels, buildings and roads with examples of their attributes being ownership of the land parcel, current use of a building or name of a road. In contrast, the field approach conceptualises the world from an absolute perspective as a continuous surface of attribute data without explicitly defining abstract geographical objects. Examples include elevation, concentration of pollutants in soil and air pollution. Each data model has advantages and disadvantages which need to be taken into account and also each data model will have different ways of quantifying uncertainties associated with geographic information. A very good example of this is the consideration of uncertainty in a remote sensing context. Remote sensing uses almost exclusively the field data model and the statistical underpinning of remote sensing allows error to be considered relatively easily (Foody and Atkkinson, 2002). However ambiguity and vagueness have not been discussed as much because they will not readily fit into to the existing statistical framework. New data models (such as fuzzy sets) are being developed which allows for increasing use and discussion of inclusion of ambiguity and vagueness within remote sensing data sets (Fisher and Pathirana, 1990). A full discussion of the differences between data models and the impact this has on uncertainty is beyond the scope of this thesis, for more details see Robinson (1995), Zeiler (2000) or Longley et al. (2001).

In order for data regarding geographic phenomena to be recorded, the real world has to be conceptualised before it can be recorded in the chosen data model. There is not always an obvious natural unit of geographic analysis, for example what proportion of an area of trees should be oak if an area is to be designated oak woodland? Homogenous areas of just one species do not occur naturally, so is a proportion of 75% sufficient, or 80% or 90%? Whoever is recording the area and creating the map sets this limit, which should be formally declared in the ontology and made clear in any end product using the spatial data. Ontologies are used to formally delineate these types of definitions, so spatial data have a formal conceptual understanding (Hunter, 2002). This information can then be used to inform the user of the uncertainties within the data, and also to allow comparison between datasets that use different ontologies. Unfortunately, often this type of declaration is not included in the intermediate spatial data or the final end product (Longley et al., 2001).

Any uncertainties at this stage can be categorised as vagueness or ambiguity. Vagueness is when a category cannot be precisely defined – as in our example of oak trees above. This can be addressed by fuzzy set theory, where membership of the category is defined by the degree of membership – e.g. our woodland could be 75% oak, 25% some other species (Fisher, 2000). Ambiguity is when the category is well defined, but there are disagreements about how to apply it because of different perceptions of it (Fisher, 1999). This is a particular issue across languages and cultural groups (Longley et al., 2001).

2.4.1.2 Uncertainty in the measurement of geographic phenomena

After the world has been conceptualised, it has to be measured using a variety of tools in order to collect the data to represent the world on a map. The tools can create error in the data recorded, depending on how accurate the tools are and how well they are used. Error is used to refer to a situation when the difference between the recorded value and the true value can be measured (i.e. the true value is known) and uncertainty is used to refer to a situation when the difference between the recorded value and the true value cannot be measured, because the true value is not known. The measurements may be made using different data models (e.g. object-based or continuous) depending on the intended usage of the data. There will always be differences between the recorded measurements we make and the truth because it is not possible to store a fully complete model and it is very unlikely that full geographical complexity can be reduced to models with perfect accuracy (Zhang and Goodchild, 2002). This element of uncertainty is the one that is most fully understood with many different approaches to quantifying the error in the data set compared to the real world, such as a confusion matrix (see Section 2.4.4).

A number of authors highlight some of the ontological and perceptual issues regarding the presence or absence of uncertainty in data and the terms used to describe it. MacEachren et al. (2005) say that data producers seem keener to report accuracy than uncertainty, possibly because 'accuracy' would be seen as a benefit to the data, whereas 'uncertainty' would be seen as a detraction. Duckham and McCreadie (2002) also mentioned issues associated with the negative connotations' of the word ‘error’.

2.4.1.3 Uncertainty as a result of analysis of geographic phenomena

As there are a variety of uncertainties in the base geographic data, any analysis of this data may propagate these uncertainties. This could include performing an overlay to find the area where two polygons coincide. If the polygons had uncertainties associated with them (for example an error in the spatial extent), then the overlay operation would propagate these errors in the spatial extent into the output polygon. This is also an issue in modelling, where a number of different data sets are used for a complex model (such as that used in hydrological modelling). The data sets may well have different uncertainties within them and the uncertainties can propagate throughout the model (Benke et al., 2011). Problems also include making inappropriate inferences from aggregate data, such as assuming all people in a census output area have the mean characteristics (known as an ecological fallacy) or the fact that the relationship between aggregate variables depends on the size and shape of the aggregation (known as the Modifiable Areal Unit Problem, MAUP) (Openshaw, 1984; Longley et al., 2001).

2.4.1.4 Fitness for use

An additional aspect of uncertainty of spatial data is that of ‘fitness for use’. For example, when a user wants to use a set (or sets) of spatial data for a decision making exercise, they need to be able to assess whether the data are appropriate to use for their task. This is a difficult issue to assess, and requires an expert user to look at the available metadata for the data sets involved (including the aspects mentioned above) to make a decision (Devillers et al., 2007). With the increasing number of people using spatial data no longer restricted to expert users, there is a need for a tool to assist users with this process. Unfortunately this type of tool is beyond the currently available technology and will be for the near future. However, Devillers et al. (2007) developed a tool allowing metadata to be aggregated from a number of different data sets, which is a step in this direction.

There are very few standards to assess such 'fitness of use'. There are a significant number of standards for data quality (such as ISO and ICA, see Devillers et al., 2007 for details), which mainly focus on data accuracy (i.e. how closely the data are to the real world objects on the ground). There are five different elements common to most of the standards, known as 'the famous five'; positional accuracy, attribute accuracy, temporal accuracy, logical consistency and completeness (Devillers et al., 2007). While these cover data accuracy reasonably well, they do not address 'fitness for use'. The issue of fitness for use is significantly more complex, because the answer to the question 'Can I use this data to do X?' very much depends on what 'X' is. While these issues are very important for the ever expanding use of GI, a more in-depth discussion is beyond this thesis.

2.4.2 Uncertainty in this research

The most well researched area of uncertainty is that of accuracy, which is a measure of how well geographical information in a GIS represents the real world objects it was recorded from. Highly accurate data will have a low level of error within it. It is information on accuracy that is most frequently available for spatial data sets. Positional accuracy and attribute accuracy are the two different types of accuracy that will be used in this research.

Spatial data inherently have a position and the error within the position is termed positional accuracy. Positional accuracy is usually measured in terms of how far the representation on the map could be from its true location in the real world (Goodchild and Hunter, 1997). The level of accuracy of this position has a significant impact on the number, extent and type of analyses that can be performed on a specific set of data and the usefulness of the output of any analysis (Beekhuizen et al., 2011). Recording the positional accuracy of a data set, or a feature within that data set is often supposed to be addressed in the metadata that comes with the data set when it is provided. However, when this data is provided there remains the question of how best to represent the accuracy. In addition, often metadata are provided at the data set level rather than the feature level which makes it impossible to display a positional accuracy for each feature, because the data are not available.

Attribute accuracy is a particular issue in models predicting future variables, such as climate prediction models. In a typical model, there is a vast amount of information available about the accuracy of different elements within the model and how these inaccuracies may propagate through the model (Zhang and Goodchild, 2002). While the model designer may be aware of these, or even have recorded this in the documentation, most models try to reduce and/or eliminate the error which may be appropriate in some cases. However, in cases of deep uncertainty, where future research may never be able to eliminate the uncertainties seen in the models (Kasperson, 2008), it may be best to represent this uncertainty explicitly to the user. Recently available data have started to include such measure of uncertainty, as it has reached a point where it cannot be reduced any further, given the current scientific knowledge. When this limit has been reached it is important that this information is communicated to the users of the data. Communicating this information on the uncertainty of the variables in the data set can also allow buildings and infrastructure to be built to particular limits, or ranges of weather, depending on the organisations approach to risk.

This thesis will not consider the details of the different types of uncertainty any further, but is concerned with how uncertainty is represented rather than how it is defined and recorded. Indeed, very few data sets have any measure of object level uncertainty included with them which limited the data that could be used in the case studies.

2.4.3 Why is uncertainty important?

Spatial data is a crucial part of many decision making processes and uncertainties within the spatial data may have a major impact on the decision. If information on the uncertainty of spatial data is not included, it can either lead to unfounded assumptions about the data, or an inability to understand the data properly. This latter issue is a common occurrence with data relating to climate change (Pidgeon, 2008; Gosling et al., 2012). Additionally, if uncertainty is not understood correctly or not represented in an effective way, its presence can undermine trust in the operator using the data or results generated from the data with potential for an entire piece of work to be abandoned because of the perception of the impact of uncertainty. Equally, confidence in results is crucial for them to be accepted and for the public to recognise (and take seriously) the risks being addressed (Zerger, 2002).

Uncertainty in definitions used in data collection can affect the comparability of statistics, for example in the US between 2003 and 2004 the method of recording terrorist activities changed significantly, meaning the two years cannot be compared in any meaningful way (Fisher, 2009). Census statistics can have substantial regional differences because of reporting techniques and have been known to be manipulated for political aims (Simkhada et al., 2009). China is a good example of this as between 1949 and the early 1980s, statistical data were used as a propaganda tool and lead to some embarrassing mistakes in the evaluation of China's development in the west (Banister 1987; Heilig 1999). Such problems are not necessarily limited to the developing world or politicians. The UK Jewish population has been historically underreported in the UK Census because Jews generally report themselves as ‘Jewish’ in the ethnicity section rather than religious section of the census form (Graham and Waterman, 2005). This is a mismatch between the census designers (who see Judaism as a religion) and the Jews themselves (who see it as an ethnicity). There are also very good scientific reasons for altering recording methods (for example to improve the amount of information captured or to make it more appropriate for a particular use) that can make calculating change between data sets challenging. A prime example of this is the Land Cover Map 1990 and Land Cover Map 2000 data sets, which were designed and made for very different reasons and so changes in land cover between 1990 and 2000 cannot be assessed easily (Comber et al., 2003).

Additional issues from inconsistencies in definitions include political ones, such as the mean sea level datums for Belgium and the Netherlands being about 1 metre different. There is a good argument for all countries using the same datum for sea level, but which country's sea level do you choose? Land based international boundaries (such as the one between Ireland and Northern Ireland) can also cause problems when trying to combine data sets to produce whole island maps because of incompatibilities (Kitchin et al., 2007).

Many more spatial data sets are being shared between different organisations and people than even 4-5 years ago, which in some ways is good because it means spatial data are now much more accessible. However, there are two main down sides to this; firstly the fact that when spatial data are collected they are usually collected for a specific purpose and subsequent uses may not be suitable for the data, and secondly the average technical expertise of users using spatial data is decreasing because of the increased range of people using spatial data.

Over the last 4-5 years there has been an increase in geospatial data being made available to the general public in the UK such as Ordnance Survey's 'OpenData' initiative (Ordnance Survey, 2012c). This reuse of data is important as it generates a significant funding stream, both in academic and commercial areas. Additionally many research councils in the UK provide a number of 'data sources' or 'data repositories' to allow collected data to be used by as many people as possible (including the ESRC Data Archive, http://www.data-archive.ac.uk; NERC Environmental Information Centre, http://www.nerc.ac.uk/research/sites/data/terrestrial.asp and British Atmospheric Data Centre (BADC), http://badc.nerc.ac.uk). It is also increasingly common for there to be a section in funding proposal forms on dissemination, with a requirement to distribute the findings of the research (including data, if relevant) to as wide an audience as possible (Weber et al., 2012). However, given that data are initially collected to a specific level of detail for a specific purpose, the question then is how does the user know that the data are appropriate for what they want to use it for? This is where metadata can provide a significant benefit, providing information on how the data were collected, which can be used to work out whether it is appropriate or not for a particular use. There are a number of standards addressing data quality, as discussed previously, but take up of these is very limited (Fairbairn et al. 2001; Devillers et al. 2007), because of the complexity involved, as shown in the quote below.

"Data quality is a messy issue, and the average GIS user would probably far rather ignore it than confront it" Goodchild (1995, p.424).

The EU INSPIRE directive makes specific provision for inclusion of uncertainty information in metadata (Comber et al. 2006) and its implementation may result in greater use of metadata to communicate this.

Secondly, due to this expansion of availability of spatial data, many users of the data have less knowledge about the features and limitations of such data than 10 or 20 years ago (Devillers et al., 2005). Goodchild (1995, p.424) sums this up nicely by saying "GIS is its own worst enemy: by inviting people to find new uses for data, it also invites them to be irresponsible in their use". If the users are not aware of the limitations of the data and base important decisions on the data, it could potentially result in negative social, political or economical consequences (Monmonier 1994).

2.4.4 How is uncertainty estimated?

Given the importance of uncertainty, it needs to be quantified or estimated before it can be communicated. There are a number of ways of quantifying uncertainty within a geographic data set, which can vary with the type of geographic data set that is being sonified. Many uncertainty calculations evolve from statistics and give a summary of the error of the data set. The most common of these is the RMSE (root mean square error) which is used to show positional accuracy, usually represented as a distance (Wood et al., 1994). The RMSE is usually reported on a data set rather than on an object level basis, which does provide some information, but assumes the error is distributed homogeneously over the data set. This limits the use of the error information, because while the data set level information is useful for error as a result of instrument imprecision (or other systematic errors that affect every element within the data set) it does not show whether the uncertainty varies between different elements in the data set (Lee et al., 1992). A confusion matrix works in a similar way, in that it provides a summary of the error for the whole data set, but it can also illustrate different errors in different categories which provides some additional information. However, it still assumes homogeneity within each particular category (Aspinall and Pearson, 1995).

GIS software has often lacked the appropriate tools to estimate, understand and display uncertainty (Unwin, 1995; Duckham and McCreadie, 2002; Devillers et al., 2005, 2007). This lack of development of uncertainty capability is due to a combination of limited theoretical progress on understanding the issues uncertainty creates and some of the negative perceptions around the concept of data having uncertainty. Progress is being made, with raster GIS solutions tending to have more advanced progress than vector ones (Unwin, 1995) because of the developments from remote sensing, as outlined earlier. Duckham and McCreadie (2002) illustrated some options for a vector-based error aware GIS, but this was a custom built solution that could not easily be incorporated in to a commercial GIS such as ArcGIS.

Research has been conducted on object level uncertainty, although the focus is usually on the presentation method of the uncertainty information rather than the generation of uncertainty data itself. Duckham and McCreadie (2002) highlight the amount of information that is implicitly stated on paper maps. These include the scale and definition of the paper map, from which the Minimum Mapping Unit (see Fisher 2009) and smallest possible object (e.g. 0.5mm line) can be deduced. Duckham and McCreadie (2002) designed a conceptual model for an error-aware GIS, which highlights the amount of work that would need to be done to create one, given the very limited starting point of current commercial GIS. Similar work was conducted by Devillers et al. (2005; 2007) who created an outline for an active error aware GIS which was able to provide advice on whether particular data are suitable for the work the user is trying to complete. They create a prototype to see how this might work in practice, but the appropriate data (with sufficient metadata) and infrastructure (a well used GIS system with this capability) are not currently in place. There has also been a significant amount of research on modelling uncertainty to generate data to display, with a number of different techniques such as Monte Carlo simulations (Ehlschlaeger and Shortridge, 1996) and fuzzy sets theory (Fisher and Pathirana, 1990). While the generation of the uncertainty data is very important to a successful method to communicate the uncertainty data, it will not be covered extensively in this thesis.

2.4.5 How is uncertainty represented visually?

Drummond (1995) makes the important point that however the accuracy is calculated, it needs to be communicated to the user for it to have any effect. The representation of uncertainty has had significant coverage within the literature and often visual methods are effective where the uncertainty information is represented alongside the original spatial data (Ehlschlaeger et al., 1997). To represent object level data, an additional dimension is required to show the supplementary information, as the uncertainty information will cover the same spatial area as the base data. While this is easily done on simple data such as using error bars on a graph, with a more complex data set (such as a set of heights) this is more difficult, because it covers a much larger area. The design needs to avoid obscuring any of the existing information, and so details of uncertainty are often omitted from spatial data presentations such as landscape visualisations (Appleton et al., 2004) potentially because of this issue of obscuring.

There are many different visual methods some of which are more effective than others and some lend themselves to certain data sets. Hunter and Goodchild (1995) provide a good overview of a number of different methods of quantifying and representing uncertainty in spatial data. The more traditional ones include epsilon bands (Blakemore, 1984), map reliability diagrams and variability diagrams (MacLean et al., 1993; Wieczorek et al., 2004). Schweizer and Goodchild (1992) experimented with using colour to represent uncertainty, and found that uncertainty data was treated in the same way as any other spatial data represented on a map. They found significant problems with the users understanding of uncertainty, which may have impacted the results. Benke et al. (2011) use a combination of spatial and non-spatial methods to represent uncertainty in output from a hydrological model. They use a cross-sectional profile to show the uncertainty in the hydrology of an area. Other methods take a more inclusive approach and try to show a range of uncertainty; these include fuzzy logic (Fisher and Pathirana, 1990; Leung et al., 1992) and probability surfaces or simulation techniques (Ehlschlaeger and Shortridge, 1996). These use the same visual methods as the more traditional techniques, but change them to show the range of uncertainty variables. There are limitations in these methods, for instance if there are too many objects then there is not enough space on the map to show the uncertainty and the object clearly.

More radical methods do not stick to the traditional visual methods, but utilise small multiples or animation (Tufte, 1990; Aerts et al., 2003). These can be quite effective, but often lend themselves to one particular data set and have limited wider applicability. Appleton et al. (2004) discuss various methods of representing uncertainty visually including colour, shading, blurring, text labels and multiple maps; however these methods can obscure underlying data or limit the amount of information shown. Kardos et al. (2005) use a quad-tree overlay technique to represent uncertainty in New Zealand census data. The overlay consists of different densities of hatching, depending on the area of the census data block and the uncertainty it has. Many of these more novel aspects have only had limited development and may only have limited applicability. Spear et al. (1996) compares a number of methods and finds that their usefulness varies between people. Ehlschlaeger et al. (1997) use the example of animation to show a series of possible options for downscaled elevation data. Animation is a popular option and it has some very good advantages for highlighting changes between different possible options or outcomes. They also highlight how complex the process of creating animations is and how many different options have to be chosen, any of which could affect the effectiveness of communicating the uncertainty information. Research is beginning to be done on optimising animations to show data in the most effective way (Goldsberry and Battersby, 2009), but there are still a significant number of areas that are not fully understood.

There are many different visual methods of showing uncertainty data, but many of them have the potential to obscure or cover the underlying data. Even if they are included as some sort of overlay (such as colour or hatching), if they are not designed particularly well there is still potential for them to be too complex or distracting for the user to use them effectively (MacEachren et al. 2005). Animations can communicate a large amount of data, but they need to be very well designed to avoid over whelming the user (Ehlschlaeger et al. 1997). However there are limitations with representing this information visually, often because of the limited amount of information that can be displayed visually. For example a typical method of representing temperature for many different sites over the UK would be to use a blue-red choropleth map with blues representing cooler temperatures and reds representing warmer temperatures. However if these temperatures are future temperatures for a specific date, and we also have some information on the confidence (or uncertainty) of these temperatures how might we show the confidence as well as the temperatures? The uncertainty could be shown using hatching or shading of some type, but for a large number of different sites over the UK this would get very crowded and difficult to read. Additionally, a darker blue with a light shading may look very similar to a lighter blue with a dark shading, making the display potentially confusing. This is an example of the situation when the uncertainty information could not be displayed visually on the same map and a different solution would be required.

Tufte (1997, 2001) gives a very good overview of presenting scientific data in an effective visual way and many of his approaches can be used in a GI context. There are many conventions and recognised ways of representing particular objects on a map that have become accepted standards, such as rivers being blue, forest being green and symbols for pubs, post offices, view points and the like. However, representing uncertainty data (of any type) is significantly more complex because of the area the data can cover and the fact that there are not many examples of this, so there are no accepted standards. Whenever uncertainty is represented in spatial data, the user has to go through a learning phase to ensure they understand what they are being shown, which significantly adds to the time taken to gain useful insight into the material (Veregin et al., 1993; Fisher, 1994; Ehlschlaeger et al., 1997).

Different approaches to communicating uncertainty are required for different audiences. For example, expert staff in meteorological offices around the world are happy using probabilistic data in a range of forms, but when the weather is presented to the general public, probabilities have to be used very carefully to ensure they are understood correctly (Environment Agency, 2009; Lunnon, 2012). The use of percentages or frequencies is the most effective way of representing probabilities to the general public and putting the forecast event in context of a recently experienced event is also useful (Environment Agency, 2009). Participants' previous experience of handing uncertainty data will influence their performance when using uncertainty data, as will participants understanding of science. Those who understand science as a search for absolute truth will struggle understanding the concept and implications of uncertainty, whereas those who understand science as a debate will be happier with uncertainty data (Rabinovich and Morton, 2012). When considering how to communicate uncertainty, the context of the communication and the audience must be carefully considered.

2.4.6 Potential non-visual options

A number of studies have considered how non-visual methods could be used to represent uncertainty, but these are less common than the number of studies evaluating visual methods. Smell and taste would be very difficult senses to use in a way which is easy to control the sense (to allow it to represent something) and be non intrusive to the user. A very limited amount of research has been done into using smell as an interaction and selection method, similar to vision or sound. There is significantly less theoretical information on how smell in the human body works and there are very severe limits on switching between different smells quickly. As such, it is not a suitable medium for this research. Brewster et al. (2006) used smell to get participants to tag photos in photo collections. There were significant implementation issues, but the principle did work. It is unlikely this could be easily applied to spatial data, because of the amount of information involved, but potentially a map could be used as an input device, with different smells being released depending on the location the user selected on the map.

Tindall-Ford et al. (1997) and Oviatt et al. (2004) explored the potential of combining two senses together (for example vision and sound or vision and touch), but they found the participants were overloaded with the volume of information. Nesbitt (2006) commented that while there was a significant exploration of the visual sense, very little had been written about the use of sound. A significant amount of work has been done with haptic (touch) interfaces (see Jeong and Jacobson, 2002; Golledge et al., 2005; Roberts and Paneels, 2007; Harding and Souleyrette, 2010 for some examples), but it often requires expensive equipment to recreate and training to use the touch environment (Golledge et al., 1995; Rice et al., 2005).

Relatively little research has been conducted on using sound to represent any aspects of spatial data. Fairbairn et al. (2001) mentioned sound in their Research Agenda for Cartography and talked about the many possibilities of using sound (such as to explore multivariate data, to narrate animated maps and to provide additional realism to virtual reality environments). They also mentioned the small amount of research that has been completed in this area, as well as that of haptics (touch) and they commented that these specialised representation methods have great potential for particular data sets. Harrower (2007) and Bunch and Lloyd (2006) looked more specifically at animated maps and they considered how sound could be used with these. From a cognitive load theory point of view, using sound to transfer some of the information away from the visual channel to the audio channel worked very well and their main focus is on using narration to provide contextual information in animated maps. They mentioned using sound to represent data and stated that this is a more complex area that they feel has potential, but it requires significantly more research. A number of studies have considered the idea of multimodal maps (i.e. maps using more than the visual sense) and one study (Bunch and Lloyd, 2006) considered that effective multimedia maps can be created, but they have to be designed very carefully to avoid overloading the user with too much information. Also these are a relatively untested concept, so there are few previous examples on which to base an effective design.

2.5 Developing a Conceptual Model of the Sonification Process

This section develops a conceptual model of the sonification process. Initially the main stages of the sonification process are outlined, followed by a synthesis of the literature relevant to the main stages. This is then developed into an extended conceptual model which identifies the main factors that influence the effectiveness of a sonification. The conceptual model will be used to identify areas for further investigation in this research and the results will be used to improve the conceptual model in Chapter 7.

Based on the HCI literature mentioned previously (see Section 1.3), there are six main stages for a sonification (A-F), as shown in Figure 21. These are split into three different sections which cover the Design, the Use and the Evaluation of the sonification.

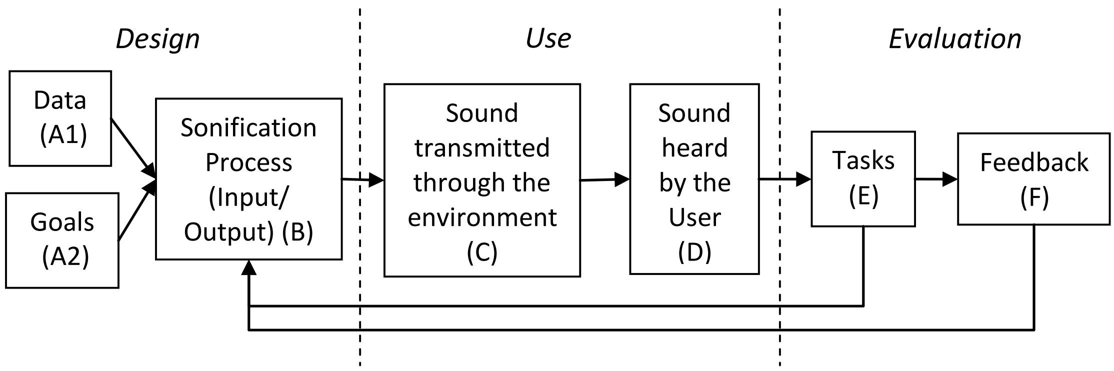


Figure 21. The six main stages of developing a sonification, split into design, use and evaluation sections.

The six elements shown in Figure 21 are the main stages of the sonification process. These form the core flow process for the majority of sonifications and include the areas that are likely to impact the effectiveness of a sonification. There are three main sections of the sonification, the design, the use and the evaluation. The design section covers the process of choosing how the data are represented using sound and is very dependent on both the data itself (A1) and the goals of the sonification (A2). The data are then processed (B) with reference to the goals to generate the sounds for the sonification. The use section covers how these sounds are transmitted through the environment (C) before they are heard by the user (D). The next section covers the evaluation aspects of the sonification. Specifically the user will have a particular task to complete (E) which can have varying levels of success. If the sonification allows the user to modify the sonification process (for example by altering the data to sound mapping) they may do this to improve their experience (iterative loop E -> B). They may also feedback their views to the sonification designer (F) who may also elect to improve the sonification process (iterative loop F -> B). These six stages form the main stages of a sonification. The next two sections use the conceptual model as a basis for a synthesis of the literature linked to each stage and an extended conceptual model of the overall process including the main factors that are likely to impact a sonification.

2.5.1 Synthesis of the literature

A wide range of literature has been reviewed that is relevant to sonification including both theoretical and practical work. The next figure (Figure 22) builds on the six main stages depicted in Figure 21 by showing examples from the literature that are relevant to each element that may impact a sonification. This synthesis of the literature also cross references each element to the relevant sections of the thesis.

The data (A1) is crucial to the sonification and there are many characteristics that will impact the sonification. If it is uncertainty data that is being represented, then it is very important to understand how the data producers conceptualised the uncertainty and how they then encoded this in the data (see Longley et al., 2001; Zhang and Goodchild, 2002 and Section 2.4 of this thesis for more details). The goals of the sonification (A2) will also influence the sonification design because different sonification methods suit different goals. For example searching for a temporal pattern in the data will require a different sonification technique to searching for a spatial pattern.

The sonification process (B) has a significant body of literature from a range of sources, mainly in terms of how sound can be used to represent different information as well as which sounds work best to represent which data sets. Multimodality is when sound is used to show a different variable (or variables) than vision to reduce the load on the visual sense. Often it is the best use of the sonic variable when it represents something that is directly related to the visual variable which will make it easier to understand (Oviatt et al., 2004). Redundancy (or dual coding or reinforcement) is when one data set is encoded in two (or more) variables. It is well known in the visual mapping area, often with colour and size or colour and shape both used to represent the same information, and this results in better performance than just coding using one variable. It is

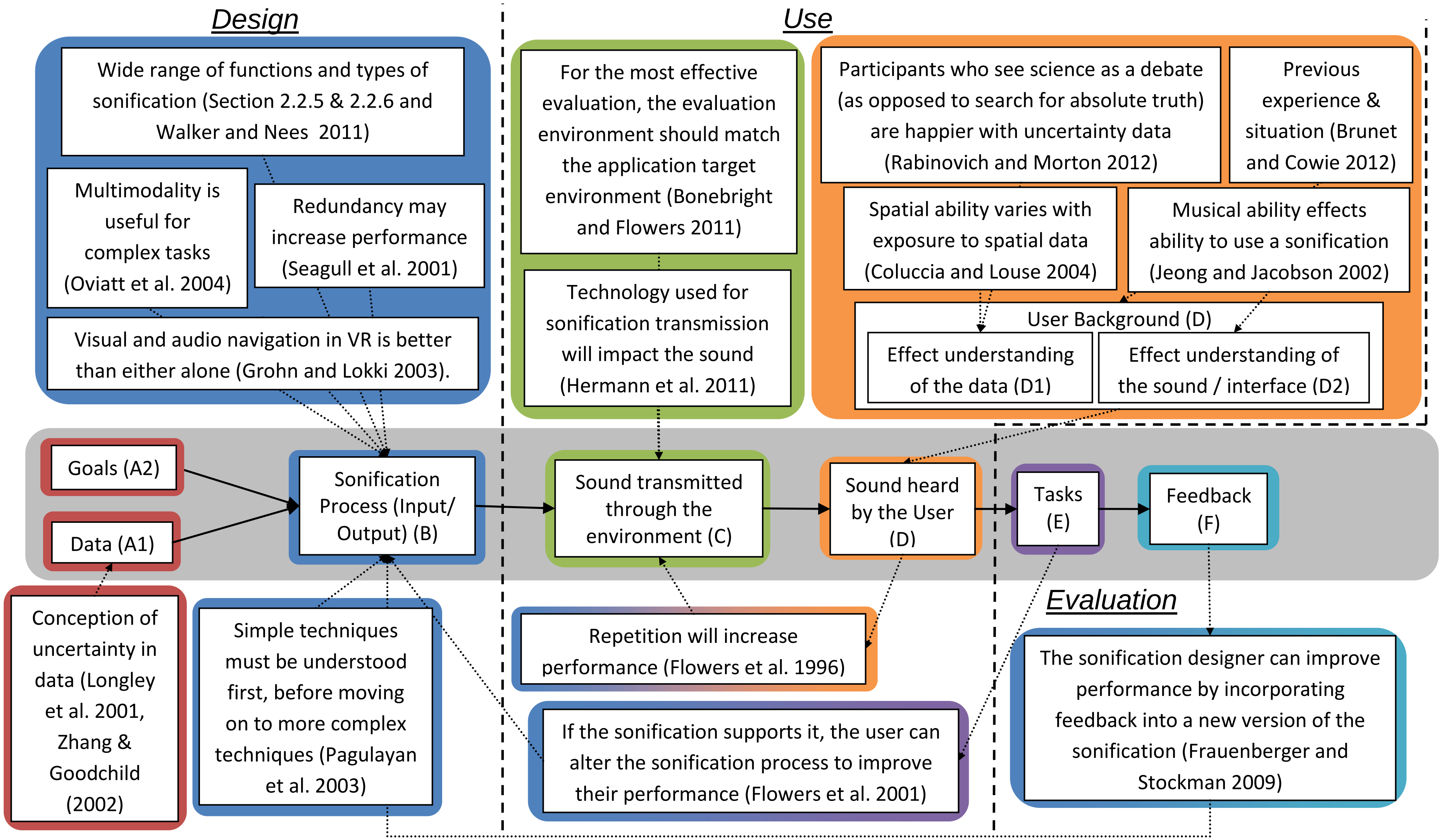


Figure 22. Based on the six main stages of a sonification (Figure 21) this figure shows examples from the literature and how they contribute to the different aspects of the model.

also used in a sonification context, and where a sonic variable and a visual variable are used to represent one piece of information (e.g. pitch and colour). This also results in better performance than using one alone (Seagull et al., 2001). Navigation in virtual environments benefits from using vision and sound together (see Grohn and Lokki, 2003 and Section 2.3.3 in this thesis). Additionally, the principle from the computer games community of starting with a simple interface and gradually becoming more complex is very important, as if the initial interface is too complex, this will put off the users (Pagulayan et al., 2003).

Once the sound has been generated it needs to be transmitted to the user (C). The factors that can affect this step include the technology used in the transmission (e.g. headphones / speakers, including technical differences between different types of headphones and speakers) as well as the environment (Hermann et al., 2011). Ideally the environment for the intended use of the sonification should be the same as the evaluation environment otherwise the evaluation will only provide limited information on the usefulness of the sonification. For example, a tool for sonifying stock market data worked well in a quiet testing lab, but could not be distinguished from background noise on the trading floor (Bonebright and Flowers, 2011).

Other important insights from the literature consider how characteristics of the user will influence the success of the sonification (D). The user's background will affect both the understanding of the data that is being presented to them (D1) and their understanding of the sound in the sonification (D2). Participants' previous experience of uncertainty in data will impact their ability to make effective use of it, and participants who understand science as a debate (as opposed to search for absolute truth) are happier with uncertainty data (Rabinovich and Morton, 2012). The user's ability with spatial data will depend on how much exposure to such data they have had previously (Coluccia and Louse, 2004). Whether or not musical experience has a significant impact on the ability to understand the sound in a sonification is much debated in the literature because of the limitations of measuring musical experience (Edwards et al., 2000). It has been observed in some studies (Jeong and Jacobson, 2002), but not in others (Walker and Nees, 2011). Individual differences, the situation and context will influence both the understanding of the data and the sonification, as identified by Brunet and Cowie (2012).

While sonifications can be very new to the majority of users there is a learning effect which is often seen, when repetitions of a task using the same sonification (C -> D -> C) results in improved performance (Flowers et al., 1996). The sound heard by the user will either allow the user to achieve the goal they set out to achieve, or not, and if the sonification supports it the user can alter the sonification process (B) to help them achieve the tasks (E -> B) (Flowers et al., 2001). Additionally, they can also provide feedback to the sonification designer (F) who may be able to improve the sonification for all users (F -> B) as described in the iterative improvement phase of *paco* (Frauenberger and Stockman, 2009).

The different aspects discussed above have a number of contributions to make to this study. The whole area of sonification is relatively new and undeveloped and as such there are no over arching conceptual models of sonification readily available. Most likely due to its origins in HCI and psychology, the area of sonification tends to favour practical implementations and evaluations of these implementations over theoretical discussions of the concepts and principles (Supper, 2012). However, the more theoretical aspects are beginning to be addressed, both in items of practically locating the disparate elements of sonification (as in the Introduction to the Sonification Handbook, Hermann (2011)) and theoretically, such as Frauenberger and Stockman (2009).

2.5.2 Extended conceptual model

This section uses the previous conceptual model and literature synthesis to construct a flowchart outlining the processes that are required to create a successful sonification of uncertainty in spatial data, and the factors that will influence this (Figure 23). The model may also be applicable to sonification of other multi-dimensional data or other aspects of spatial data.

The first input is the data (A1) that will be sonified and there are two main factors about the data that will influence how it will be sonified and therefore the success of the sonification. Firstly, the format of the data (A1a), both in terms of technical format and how it is constructed as a model of reality (see Section 2.4). This also includes the resolution, file format and so on. Different types of spatial data (e.g. discrete / object based or continuous) are likely to require different sonification methods to be effectively represented. Secondly, specific to this study is the conception of uncertainty and how this is shown within the data, which is an outcome of the data provider's choices (A1b). There are different ways to represent uncertainty and this will impact how it can be presented. Uncertainty can be a complex and difficult concept to communicate, whether this is done visually or sonically. The goals of the sonification (A2) also have a major impact on the sonification design and will vary depending on the specific sonification application.

Secondly is the sonification process (B), which is also called 'input/output' (based on Ibrahim and Hunt's 2007 HCI model, see Section 1.3) where the data is the input and the sound is the output. This is dependent on the particular sonification technique chosen from the range of potential options (see Section 2.2.6). Sonification processes are very reliant on technology and certain sonification techniques require specific technologies which may not be available. The sonification that is produced also needs to be aesthetically pleasing, or at least aesthetic enough so the user is

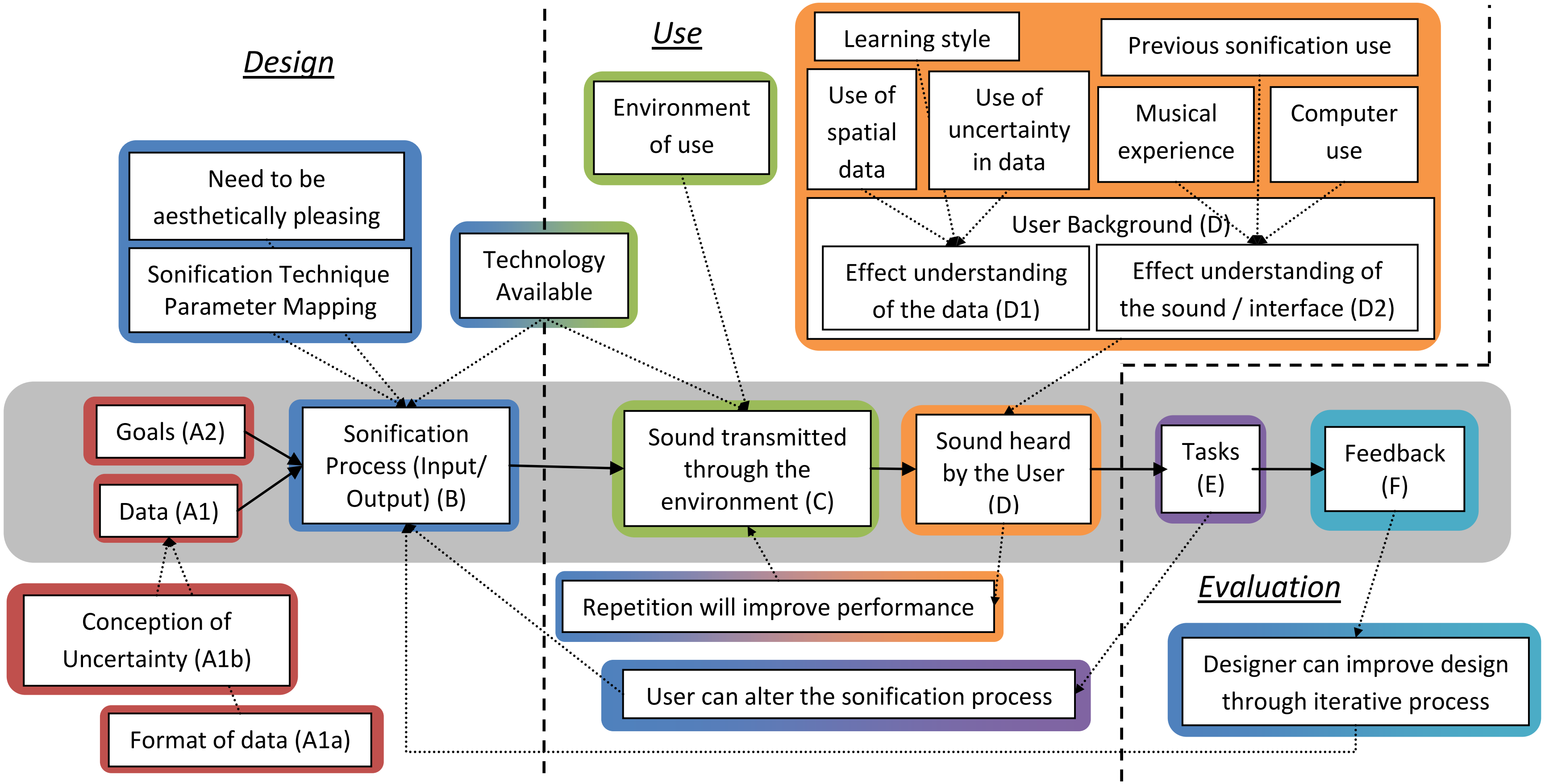


Figure 23. An extended conceptual model showing the overall process for sonifying uncertainty in spatial data, and the factors that are likely to influence the effectiveness of the communication process.

able to listen to it for a reasonable period of time to complete the task. This is a similar issue to that of acceptable design in maps (Robinson, 1995) and the standards for sonification are beginning to be developed (for example Herman, 2011).

Sonifications are very environment dependent and the environment of use (C) will impact how the sound is heard. A sonification that is designed to work in a quiet office environment over headphones will be very different to one designed to work in a busier, nosier environment, such as on a device like a mobile phone. The technology available will also impact the sound through how it is output to the user, for example over different types of speakers or headphones.

For many years, visualisation was limited by what the available technology could achieve. This is now much less of a constraint (Lange, 2011), but sonification is currently going through a similar process and it will be a number of years before we are no longer limited by the technology. During this process, conventions will be created, with particular sounds associated with particular meanings. Some of these may be logical associations, such as the roar of a car engine representing traffic. Other associations may just be based on arbitrary decisions, such as the visual decision to represent motorways as blue on maps, because blue was chosen for motorway signs. Equally, some sonic variables (e.g. volume, tempo, duration, etc.) will work best with specific data sets (nominal, ordinal, etc.) and these will be adopted as convention. If these associations are successful, it will reduce the need to look at the legend on a regular basis, particularly for frequent users.

After the sound is created (B) and transmitted through the environment (C) it is then heard by the user (D). A number of user characteristics will affect the performance of the sonification. These can be grouped into two separate areas; firstly areas that effect the understanding of the data (D1) and secondly those that effect the understanding of the sound and interface (D2). The factors that affect the understanding of the data concern the user’s previous knowledge, with greater knowledge about an issue resulting in a better performance. The factors that affect the understanding of the sound and interface include musical experience, learning style and previous sonification use. The expectations of the user will also vary depending on who is the user. For example, a GIS user will expect different things to a climate scientist because of their previous experience of using spatial data.

Both the parameter mapping and user characteristics will have an impact on the effectiveness of the sonification. They are also important in a visual interface (a map, or any other visual interface), but the impacts of these will not be the same in both sound and vision, so the two cannot be directly swapped. If the interface is a dynamic one driven by user movement (such as moving a mouse around on a map to generate a sound for different locations) the user can then move their mouse to a different location to hear the sound for that location.

The user will either be able to achieve the tasks they intended to or not (E), and if necessary this can feedback to the sonification process to alter the mapping used to improve the sonification for that user (E -> B). The users can also feedback to the sonification designer (F), who may decide to alter the sonification for the benefit of the majority of users, or a specific target population (F -> B). This improvement through iterative development is a common design process in this research. The next section in this investigation outlines the aspects of the conceptual model that will be evaluated. The results from this investigation will be used to update and improve the conceptual model in Chapter 7.

2.6 Overall Aims and Hypotheses

The overall aim of this research is to evaluate whether sound can be used to help represent characteristics of spatial data. Examples of uncertainty and distance will be used and while the results will relate to these, they may have wider applicability to other data. Uncertainty is used as an example because it is a set of characteristics that all spatial data have to some degree, and it is very rarely represented in an effective manner because GIS software generally do not have the capability of displaying associated uncertainty (Unwin, 1995; Devillers et al., 2005). With the use of GI increasing across all sectors, more demands are being placed on spatial data and more information on the appropriateness of spatial data for specific uses is being requested. This is in addition to the fact that for some spatial data the uncertainty element of the data is crucial to understanding the data properly.

This thesis will not be able to assess all of the factors outlined in the conceptual model because of time limitations. One main area that will be evaluated is the impact of participant background on their performance when using sonification. More specifically, the participant’s learning style and musical experience will be assessed and evaluated. The participant’s use of spatial data and awareness of uncertainty in spatial data will be explored and defined for inclusion in the revised conceptual model. The impact of the data will also be discussed throughout the case studies, showing how different producers of data approach and encode uncertainty data in different ways.

A number of studies have created prototypes of tools that can use sound to represent an aspect of spatial data, but it is very rare for these to have a conceptual model associated with them. Additionally very few of the prototypes have any significant user evaluation, with the majority of the evaluations being by two or three of the authors’ colleagues. This thesis will provide both a conceptual model and a user evaluation consisting of GIS specialists and those likely to make use of this technology in the real world. With larger sample sizes (~50 participants per case study) the results from the evaluations will be more statistically valid and be applicable to a wider variety of sonifications of data.

The research focused on a number of specific elements of the user, including learning style, previous sonification use, musical experience, use of spatial data and awareness of uncertainty in spatial data. Musical experience was chosen because it is widely debated in the literature whether it has a significant impact on sonification use. Learning style as a concept is contested, but has not been evaluated before in relation to sonification use. Previous experience has been highlighted as a potential impact on effectiveness (Brunet and Cowie, 2012) but it has had limited discussion in the literature. This research will evaluate these elements in detail, providing a new contribution of knowledge.

2.6.1 Specific research aims

The research aims can be split into three different areas, each building on the previous. The specific research aims are:

1. What form does uncertainty information need to be in for it to be represented using sound?
2. How effective are different means of representing uncertainty in spatial data?
3. Which types of sound variation best represent spatial data uncertainty to users?

These will be considered in each case study, and will be reviewed in Chapter 7, drawing together the findings of this thesis.

2.6.2 Research hypotheses

This section outlines the six research hypotheses that will be tested by this research. They cover aspects of the design, use and evaluation sections of the conceptual model. The hypotheses in the sonification design section (H1 and H2) will be more difficult to test than those in the sonification use and evaluation section (H3 to H6) because the study design is limited by the currently available data and technology.

2.6.2.1 Sonification design

The way uncertainty is conceptualised is important because this has a major impact on the amount of uncertainty information that is available within the data set. Additionally uncertainty can occur at different stages of processing of geographic data (see Section 2.4.1) and this will affect the type of the uncertainty and therefore the impact the uncertainty will have on the data. All types of uncertainty data are important, although data on some types of uncertainty (e.g. accuracy and error) are much more common than others (e.g. vagueness and ambiguity). The impact of uncertainty is increasingly important because of the growing reuse of public data, which can often be used for a different task than that for which it was initially collected.

H1. The way uncertainty is conceptualised within the spatial data set will impact how it can be represented in a sonification.

The field of sonification is relatively underdeveloped and many of the changes and advancements in it are driven by developments in technology. This is no different to early developments in GIS and visualisation, and to some degree both visualisation and sonification are, at times, more driven by the technological improvements than improvements in theoretical understanding. The specific technology used in this research will probably be redundant in a few years and the current implementations of sonifications are limited by the technology available. The available technology will impact the design, implementation and results of the sonification.

H2. The technology available to design, implement and evaluate the sonification of uncertainty in spatial data will impact the findings.

2.6.2.2 Sonification use and evaluation

The conceptual model outlines many factors that may influence the user's ability to make effective use of a sonification. It is not possible to evaluate all of them in this work with the resources that are available, so a number of factors are prioritised, including the user's previous knowledge of the data set being sonified and their previous experience of music and sonifications. These were chosen because these factors are listed in the literature as being likely to have a significant influence (see Section 2.5.1 for details).

H3. The ability of the user to make effective use of the sonification will be effected by a variety of factors including their knowledge and previous experience.

The use of sound and vision together has had relatively little previous investigation (see Section 2.3) and there are two theories (cognitive load and dual coding, see Section 1.2.3) that provide new approaches to using the two senses together to communicate more information to the end user, or to allow the end user to understand information more effectively. Measuring comprehension of the information is a vital part of the evaluation and if there are methods that could be used to improve comprehension of the information then this is useful for a wide variety of applications. Dual coding has a good basis in the research literature and it is hypothesised that the effect of dual coding the spatial information visually and sonically will improve performance.

H4. Providing redundancy (dual-coding) in a task by showing the same information both visually and sonically will improve performance.

Supplying additional information to the end user is important with the increasing amounts of information that are available. It is also important that this information is understood effectively to ensure the user gets the most out of it that they can. It is expected that representing uncertainty sonically (as opposed to visually) will improve performance by allowing the user to access the data multi-modally rather than unimodally (see Section 1.2.3 for details).

H5. Transferring the representation of uncertainty information from the visual to the auditory domain to address visual saturation will reduce cognitive load and therefore improve performance.

There are a wide variety of sounds that can be used and it is important to know whether user preferences vary between different sounds, and if they do whether this has an impact on performance. If there is improved performance with a preferred sound then it is important to know whether the preferred sound can be predicted based on factors from the individual, or whether users need to be given freedom to experiment with different sounds to select the one they prefer the most.

H6. Preference for one sound from an option of two will result in better performance with the preferred sound than performance with the other sound.

These hypotheses will be evaluated using the results from the three case studies in the discussion chapter at the end of this thesis. The next chapter outlines the common methodology used throughout all three case studies, followed by the individual case study chapters.

Chapter 3. Methodology

This chapter outlines the methodology adopted for the case studies in this thesis. Factors common to all three case studies will be discussed in this chapter, with the elements specific to each case study mentioned in the appropriate case study chapter. Three case studies have been used to allow the evaluation of sonification to represent uncertainty with a variety of different spatial data types. Each case study focused on a particular data set and the evaluation consisted of a number of expert users familiar with that data set as well as GI experts who were not regular users of that specific data set. This chapter is split into a number of sections, covering a justification of the overall methodology, the data used in the evaluation, the sounds, the technology, the participants, the task and the evaluation itself. Each section will evaluate methods used in the literature and discuss those used in this thesis.

3.1 Justification of Methodology

A number of major methodological decisions had to be made in the early stages of this research, including the number of different studies to undertake, the nature of the evaluation and the type of sonification to be used. This section will explain and justify the decisions made for this thesis.

3.1.1 Multiple case studies or one case study?

One of the first decisions to be made was whether to perform one larger study or split the available research time over three, smaller, case studies. This required a balance to be found between the amount of information that could be collected in a study and the number of different types of sonification stimuli that could be evaluated. A single study would have allowed more in-depth information to be collected about one particular sonification implementation or one specific data set. An approach that used multiple case studies would have allowed different types of sonification and data to be compared and enabled results from previous case studies to inform subsequent ones. The two different approaches would have had different advantages and disadvantages in terms of the amount and type of data that could have been collected.

Sonification is an area that has had relatively little research and so would benefit from an approach that allowed a wider range of information and variables to be collected. In-depth results from one study would also have been useful for sonification research, but the greater breadth of results available from an approach using multiple case studies was thought to be more appropriate at this time because of the current limited understanding of sonification. Therefore, it was decided to use a multiple case study approach for this thesis. This allowed the use of sonification to be evaluated with a number of different spatial data sets and also enabled results from previous case studies to be used to inform subsequent ones. Specifically this allowed elements that appeared to have an impact on the effectiveness of the sonification in earlier case studies to be investigated in more detail in later research.

3.1.2 One-to-one or group evaluations?

The next issue that was addressed was whether to run the evaluation sessions on a one-to-one or group basis. A single session that involved a group of participants would have allowed many more cases to be collected than running sessions on a one-to-one basis, given the same level of resources. However, a one-to-one evaluation would have allowed the collection of much more detailed data from each participant. It would also be crucial for evaluation methods such as the talk-aloud protocol where the participant would have provided information on their current actions and their thoughts and intentions surrounding them. Sonification is an area which has had relatively little formal evaluation, particularly using examples of spatial data. Both group and individual evaluations would have added useful information to the literature and it was decided to complete group evaluations because they would provide increased statistical validity for the results (as data from more users would be collected). It also allowed more generalisations to be made about the usability of sonifications with different types of users, because a wider range of users could be included in the evaluation. A discussion session was included at the end of the evaluation which allowed more qualitative results to be collected from the participants. While this was obviously not a substitute for a much more in-depth one-to-one evaluation, it allowed additional information to be collected that would have been missed in an exclusively questionnaire-based group evaluation.

3.1.3 Passive or interactive sonification?

The final major decision made was whether to have a passive or interactive sonification. A passive sonification would be one where the participants listened to an audio clip a set number of times and then answered questions on their experience of the sonification. An interactive sonification would be one where the participant could interact with the sonification on a one-to-one basis, with the sonification being driven by their actions. An interactive sonification would have been much more complicated and time consuming to create, whereas a passive sonification would have been relatively quick and easy in comparision. Interactive sonifications would have provided much more engagement for the user (see Section 2.2.4) and would generally have allowed much more engagement with the data than passive sonifications. It was decided to create a series of interactive sonifications which would allow for a higher level of user engagement. The higher level of user engagement would also provide much more in-depth results as the user would be more likely to give critical thought to the issues involved. This choice meant that the time to develop the interactive sonifications had to be well scheduled to ensure that the work required could be completed on time.

3.2 Data

Sonification has a wide variety of uses, but relatively few examples that used geospatial data. The few geospatial data examples have tended to use census or population data for non specialist users (see Section 2.3.2). Census and population data are relatively accessible and are data that non specialists can easily understand, which is important for the evaluations to be effective. However, this type of data often does not contain explicit uncertainty data, which means they would be of limited use in a sonification of uncertainty in spatial data. The uncertainty element of the data could be fabricated, but the use of fabricated data often leads to unreliable results. This is a particular issue for specialised environments where users are familiar with the trends they expect to see in the data (Genov et al., 2009). If the trends they are expecting to see are not visible in the data set, they can be distracted by thinking about why the trends are not there rather than focusing on the evaluation.

While the type of data used will impact the sonification design, differences can also be seen in studies that uses census or population data. The design of the sonification is very dependent on the data but also on the task that the user is trying to achieve. For example, tasks asking the users to find the highest or middle value will be different to tasks that involve the users understanding the spatial pattern of the data.

This research is specifically interested in using sound to represent uncertainty data because with the increased use of spatial data there is a major need and opportunity for this additional data to be presented effectively. Whichever data set is going to be used for the evaluation there needs to be uncertainty included in the data set, either explicitly or implicitly. Explicit uncertainty data has the uncertainty values for the objects already in the dataset (e.g. as a value within a field) so with this type of data it is relatively easy to sonify the uncertainty. Uncertainty data that is implicit is not stated as a separate field within the data set, therefore the uncertainty element of the data needs to be deduced from other available information. For example, if a road dataset contained a field for 'date last updated' for each road, it could be assumed that the roads updated longer ago have more uncertainty than the roads updated more recently. Implicit uncertainty data can require a significant amount of work to actually quantify the uncertainty before it can be sonified, whereas data with explicit uncertainty can be used without changes. While synthesised uncertainty data could be added to existing real data, it may be confusing for the users particularly if they are expecting to see certain trends within the data because of their previous experience with similar data.

Geographic data are usually stored in either continuous (e.g. raster) or object (e.g. vector) data structure forms. Continuous data covers a whole area contiguously, without any gaps whereas object data just stores information about the specific objects in question, so there can sometimes be gaps in between the objects with no information. It is likely that these different data structures need different sonification approaches because the user will interact with the data in a different way. Therefore it was decided to use different types of data as examples to see whether there were any differences or similarities between them in terms of the most effective interaction method.

3.3 Sounds

The sounds used to represent information are fundamental to any sonification and the sonification studies reviewed in this thesis use a wide range of different sounds and variations of sounds to represent data (see Section 2.3.2). Many of the evaluations in the literature use musical instruments to generate sounds, with varying the volume and/or pitch very common. Some applications are more suited to variation in pitch than others (e.g. navigation) and occasionally non-musical computer generated sounds are used. Spatial sound (i.e. sound appearing to come from different locations around the user) is common, but can only be done easily using a setup with multiple speakers.

The use of musical instruments has advantages, which include the ease of implementation of a sonification using MIDI technology (see below), the fact that musical instruments all have a scale of notes (low to high) to which values in a dataset can be easily mapped and a number of the more common instruments and the sounds they make are familiar to the vast majority of users. However, the use of instruments does limit the number of different possible notes and limits the degree of customisation of the sound. Using non musical sounds can provide much more flexibility (e.g. Sofirow, which allows the sound to be manipulated according to the motion of an oar), but are much more complex to program and use. They are also more difficult for participants to learn quickly. Spatial sounds (using multiple speaker setups) provide very easy ways to locate different sounds by allowing the operator to play sounds through different speakers. However, they require more hardware (e.g. speakers and mixing desks) than a stereo sonification and can be complex to setup, depending on size. A similar output of spatial sound can be achieved using stereo headphones and a piece of information known as the HRFT (head related transfer function). This allows the sound to be modelled so the user can identify the location that the sound is coming from. However, HRFTs are person specific and it is a lengthy and complex process to record the information for each person who wishes to use the sonification (Bujacz et al., 2011).

3.3.1 MIDI

MIDI (Musical Instrument Digital Interface) is a technology often used in computer programs to create basic musical sounds. Technically, MIDI is the standard for communicating between different computerised musical instruments, such as synthesisers and digital musical instruments, rather than a way of creating sounds itself. However, many computers are supplied with basic built in synthesisers that can generate a range of notes from different musical instruments, controlled using MIDI. When evaluation projects say they 'use MIDI to generate sounds' it is this in-built synthesiser they are referring to.

The use of MIDI as a technology to create and control sound is very common, with many studies using it as a quick and simple way of getting a sound output. In a recent review (Bearman and Brown, 2012) the most popular sound generation technique for non sonification specialists (non-ICAD in Figure 24) was Built-in MIDI. If a musical instrument is chosen for the sound, MIDI is an easy option as all the information that is required is a code to represent the instrument and a code to represent the note that need to be played. There are many add-ons for a variety of programming languages (including JavaScript, Visual Basic, Java and many others) that can take the parameter input of instrument and note and then play the sound. This works very well for relatively simple sounds, but cannot create very complex sonifications. For these, more advanced tools are required such as SuperCollider or Puredata and these are often used by sonification experts because they need the more advanced capabilities than those available using MIDI.

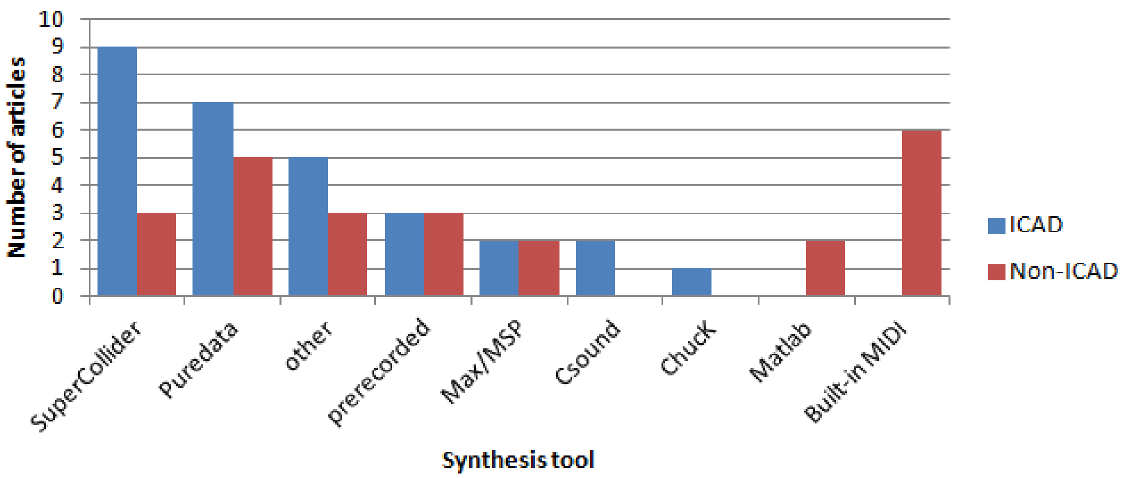


Figure 24. Graph of different sound synthesis tools used by sonification experts (ICAD) and non-experts (non-ICAD) (after Bearman and Brown, 2012, fig. 1).

Many studies, particularly those by non-experts, cited simplicity and ease of use as the reasons for using MIDI (unpublished data from Bearman and Brown, 2012). Additionally, Milios et al. (2003) said that they used MIDI because MIDI generated a more pleasing sound than pure tones. MIDI is not appropriate for all uses and for example in the application of sonification for human movement, more information needs to be shown than can be represented with MIDI notes. For example, Sofirow (Schaffert et al., 2011) uses a specific waveform of sound constructed from the rowers' movements which allowed the rowers to pick out different features in their rowing stroke. This allows them to target training to the specific element of the stroke that needs improvement and this level of detailed information would not have been possible to sonify fully using MIDI.

For this research it was decided to use MIDI as the main technology for producing sounds because it was the simplest technology available and integrated well with the tools that were used to create the sonification evaluations.

3.3.2 Instruments and pitches

None of the data used in the evaluations in this thesis were nominal so a set of sounds with some type of order was required. No nominal data were used because much of the spatial data that are likely to be represented using sound are not nominal. It was decided to use piano notes because the vast majority of people are familiar with a piano and there is a very clear range of notes and a clear scale, with a 'high' end and a 'low' end. This scale is also very easy to visualise for a non musician (as the notes are laid out in a linear fashion on the instrument) compared to instruments such as a guitar or clarinet (where the notes are laid out in a much more complex fashion). Pitch is the most frequently used sound variable in sonifications (Dubus and Bresin, 2011) and is easily understood by the vast majority of participants. Pitch is also directly represented in the brain (Levitin, 2007) which means that it will be easier for participants to understand than other sonic variables. Many other sonic variables are available, but pitch is the most well explored variable from a cognitive point of view and has a larger number of different possible values than other variables such as volume.

Related to the sound variable of pitch is the important concept of chords. Josselin (2011) discusses examples of chords (a combination of 2 or more musical notes played simultaneously), which could be used to show binary data very easily and easily extended to bivariate data. Chords can either be harmonious (the individual notes sound 'nice' when played together) or inharmonious / discordant (the individual notes sound 'bad' when played together) and the specific combinations of notes for chords are well known and have been used in music for a long time. There is a sound mathematical explanation of which notes work well together and which do not (Burrus, 2009) and this could be exploited in a sonification. However, the implementation of chords adds another level of complexity to the programming and the use of the sonification, so this was not pursued.

3.3.3 Data to sound mapping

The process of linking the data to the sound (mapping the data to the sound) is important, as it is crucial to the users' understanding of the data. While Walker (2002, 2007) has completed a range of evaluations on different mappings of sound to data, many evaluations only provide limited information about the mappings they have used in their work. For example, many studies provide the sound variable(s) that are mapped to the data, but do not provide the polarity, i.e. whether a high pitch (or volume, etc.) represents a high data value or a low data value. Dubus and Bresin (2011) completed an interesting review of sonifications, finding that 86% of studies used pitch, 62% used sound level (volume), 57% used spatialization (left/right, front/back etc.) and 43% used a combination of timbre and rhythmic time scale. Pitch was the most frequent variable of sound used, and many studies used a combination of variables.

When using pitch the general approach is to use higher pitches to represent higher values, although as Walker (2002) observed, the appropriateness of this varies with the data that the sound represents. The mapping of higher pitch to higher values was used as a starting point for the case studies in this research, with each mapping refined for the individual case studies during the pilot testing phase. Details of these steps are reported in the individual case study chapters.

3.3.4 Multimodal approaches

Many of the evaluations used sound in combination with another sense in a multimodal capacity, either reinforcing the information shown or supplementing the information shown with additional information represented using a different sense.

It is possible for users to make sense of sound and visual stimuli at the same time, even if sound represents one set of data and vision represents a different set of data. A study by Bonebright and Nees (2009) showed that it was possible for participants to interpret data represented using sound and vision data at the same time. The participants listened to auditory earcons that asked them to select one of four regions on a screen and also listened to a set of spoken passages and answered questions about them. Mousavi et al. (1995) used a geometry task to compare multimodal visual/auditory interfaces to a unimodal interface and they found that the multimodal interface was significantly more effective.

In addition, allowing users a choice of modes to perform a task can increase their accuracy. Oviatt et al. (2004) compared tasks of different levels of complexity and showed that as the task became more complex, the participants were more likely to move to a multi-modal method than a unimodal one in order to complete the task.

The multimodal method holds potential to increase the amount of data that can be communicated and it was decided to include a multimodal option in each case study. There are two possibilities of exploiting multimodal interfaces, reinforcement and augmentation. Both have certain advantages and disadvantages, and it was decided to implement them both. Reinforcement results in better performance than using a single variable alone, but does not offer the opportunity to represent additional data. Augmentation allows additional data to be shown, but there may be interference between the two variables and it does not see the performance boosts that reinforcement offers.

While using vision and sound together has a number of advantages, it can be difficult to combine the two in an effective way (Nesbitt, 2006). It is also important to pick the methods of display very carefully as if they are confusing to the user or contradictory when the different methods of representation are combined, the user's performance will suffer (Jeong and Gluck, 2003). One of the most important aspects is to combine representations of similar types of data. For example, when using the visual sense, size works well to represent ratio data, so this would work well with pitch or volume (which also works well with ratio data), but not with timbre (which only works with nominal data). For more information on the different dimensions of sounds and which data they can represent most effectively, see Section 2.3). Additionally, the sonic variables and methods that are used to represent the data must be explained to the user, as the users will not be familiar with using sound as stimuli in this way.

In addition to the multimodal methods described above, it was also decided to include a visual only method and a sonic only method of representing the data to allow for a comparison. The visual only method was intended to be a broadly equivalent method to that typically used for this type of work. While the typical methods used did vary, visual only was employed as a baseline for comparison. See the individual case study chapters for details.

3.4 Technology

All of the examples referenced in the literature review that used sonification were custom made prototypes, which required specific programs and libraries to work. This worked well for the evaluations that have been conducted, but severely restricted the ability to roll out the technique of sonification to a wider audience, particularly those examples that used specialist tools and programming. MacVeigh & Jacobson (2007) observed this and highlighted the need for a sonification extension to an industry standard GIS (e.g. ESRI's ArcGIS). This would be much easier to use and more accessible than a custom coded prototype and would also allow for easier evaluation. This study therefore created a new series of sonification tools which would work within existing industry standard software.

3.4.1 General purpose sonification tools

There are a wide range of general purpose sonification tools freely available for research use. A typical example is the Sonification Sandbox[1] created by the Sonification Lab in the Department of Psychology at Georgia Institute of Technology, USA (Walker and Lowey, 2003; Davison and Walker, 2007). It is a custom coded tool that allows the import of CSV data into the interface which can then be sonified using a number of MIDI tools. The tool was designed to be easy to use to allow novice users to input some data and create a sonification. However, it is limited in complexity. This is one of a range of tools that perform relatively similar roles, including SoniPy[2] (for use with Python), AesSon[3] (for use with Max/MSP), and the Interactive Sonification Toolkit[4] (a stand-alone sonification tool). However, these tools are all relatively basic and for instance, none of them have the capacity to include a real-time interactive GIS element to allow navigation through spatial data. This inability for the sonification tools to be used with spatial data means that they were not appropriate for this research.

There are a number of more complex sonification tools, including SuperCollider[5] and Puredata[6] (McCartney, 1996; Bovermann et al., 2011). These are very powerful tools, but are also very complex to implement and contain more flexibility in the sounds that can be created than required for this thesis. They also do not have any inherent ability to represent spatial data, so if these tools were used, the ability to handle spatial data would need to be programmed in. Therefore these more advanced sonification tools were not appropriate for this thesis.

All of the spatial data examples included in this thesis used custom coded software to create their sonifications of spatial data. These work well, but required a significant investment of time and skills to develop. Indeed, creating a sonification tool for advanced use in multiple disciplines is a complex process however it is created (de Campo et al., 2006). The use of spatial data with sonifications in a way suitable for this thesis is an advanced area and the current general purpose tools are not advanced enough to support this.

There are a number of commercial and non-commercial GIS programs, but the use of sonification with spatial data is relatively new so none of the typically available GIS packages include this as an option. A number of GIS programs do include the option to add on existing code to their interfaces (for example the use of Visual Basic for Applications in ArcGIS). The approach used in this thesis is to take an existing GIS tool and add on the code required for sonification and

[1] Available to download at http://sonify.psych.gatech.edu/research/sonification_sandbox/index.html
[2] Available to download at http://www.sonification.com.au/sonipy/index.html
[3] Available to download at http://www.sense-aware.com/2009/10/aesson-toolkit.html
[4] For more details see Pauletto and Hunt (2004)
[5] Available to download at http://supercollider.sourceforge.net/
[6] Available to download at http://puredata.info/

evaluation. It is easier to add a sonification extension to a GIS package than it is to add a GIS extension to a sonification package and the GIS software used in this work (ArcGIS and Google Maps) are already industry standard. More details on the specific methods used are included in each case study.

3.4.2 Developing a tool using existing technologies

This study attempts to maximise the reusability of the sonification tools developed by using existing tools and software where available. Along with good documentation and commenting, this will allow for easier reuse of the code. This will also minimise the development required for others to complete evaluations and to enable the tools to be reused and/or recreated by other users.

The concept of reproducible research is well-understood (Brunsdon, 2011); from a published paper and publicly available information any capable academic should be able to duplicate a set of results from work in their field. When computer techniques are involved in any stage of research, the methods can become very difficult to reproduce at a later stage because of the rapid pace of computer technology development. File formats and scripting languages can be superseded, and methods from only a few years ago may not be easily transferable to currently-available software. This is an acknowledged problem given the growing necessity of archiving digital data, and is generally addressed via either emulation of obsolete technical environments or migration of digital objects to a currently accessible format (Brown, 2006).

These issues are particularly important with the research area in this thesis because of the types of technology used. They will date very quickly, and there may be significant issues with emulating the previous operation environments. In an effort to improve the reproducibility of the case studies, screenshots or videos of the evaluation process have been recorded, and are available on the DVD with a copy of the commented source code and flow charts of both the programmatic sequence behind the evaluation and the intended progress of the user through the evaluation.

3.5 Participants

The main focus of this research is the sonification of geospatial data, so regular geospatial data users formed a key part of the evaluation. Participants who had not used geospatial data before would have to learn many basic skills to take part in the evaluation, so were not approached for the evaluation. Geospatial experience was measured for each participant and tested during the analysis to see if geospatial experience had any relationship with the participant’s ability to use the sonifications. Each case study included participants who were familiar with the specific data set being sonified as well as participants who were not familiar with that particular data set. A participants' experience with the data set was recorded so it could be seen if there was a relationship between this and their performance with the sonification. In addition to experience with the data set, a range of other information about the participants was recorded, as explained later in this section.

When planning an evaluation with a limited amount of resources, there are two possible approaches. The first option is to run a very in-depth evaluation with a small number of participants and the second option is to run a less in-depth evaluation with a larger number of participants. The first option provides much more detailed information than the second option, but does have limited statistical validity because of the small sample size. A greater sample size improves the statistical validity of the results and can therefore result in better understanding of human perception in multimodal interfaces (Jeong and Gluck, 2003). In an ideal world it would be possible to perform in-depth evaluations with a large number of participants over the three case studies, but it is very rare for an experiment at this level to have the resources to complete such investigations. Additionally, regression analyses are used as the best way to see which factors have a significant influence on performance, and these require a minimum number of participants in each category so the sample size needs to be sufficiently large. Therefore it was decided to collect a larger sample size, using multiple choice and free test answers.

The participants recruited for the case studies were a combination of regular GIS users from UEA (staff and postgraduates) and expert users from particular fields for each case study. These included staff from the Ordnance Survey (OS), UK Climate Impacts Programme (UKCIP) and Local Authorities (LA) who used GIS on a regular basis and had specific knowledge about one of the data sets being used. These participants were approached through a gate keeper at each organisation who was aware of the research project. Evaluations are often run with users who do not know the aims of the research, in order to remove any possible bias (Bonebright and Flowers, 2011). However, this was not possible because of the nature of the work being completed and the specialised audience required. Interest generated from the novelty of the topic did help recruitment for the evaluations and a large sample would not have been possible without this.

Participants were either approached informally in person, through email or through a gate keeper at their organisation. Each participant was sent a Briefing Document a minimum of 24 hours before the evaluation took place, explaining what the research was and what they would be asked to do. They were also given a Participant Information Sheet and Consent Form to read and sign at the beginning of the evaluation. For more details on these and examples see the relevant case study chapters and appendices.

3.5.1 Information recorded about the participants

During the evaluations a number of pieces of information were recorded about the participants that were thought to have an impact on their ability to use the sonification. These factors were discussed in the conceptual model (see Section 2.5) and include learning style, GIS experience, data experience (specific to the individual case study), colour blindness and experience of using uncertainty data. Learning style and data experience are factors that have not been tested before and may have an impact on the participants ability to use the sonification. The results from the case studies will be used to update the conceptual model in the Discussion chapter.

Gender may have some influence on spatial ability, with males tending to perform more effectively, although the reasoning for this is not fully understood and may be an effect of the testing or participants upbringing than a purely biological difference (Self et al., 1992; Montello et al., 1999; Coluccia et al., 2007). As people's spatial processing abilities are affected by their previous experience (Kitchin, 1996) it is important to measure previous geospatial experience (in the form of area of education and level of education) and to have a measure of geospatial ability in the evaluations.

A very common method of testing spatial skill is to use a standard battery of tests (Pattison and Grieve, 1984; Casey et al., 1997; Ramirez et al., 2012). This allows easy comparison between different studies but only provides a specific set of information on the user's spatial skill. There are many different standard tests to measure spatial ability and they are often used to measure spatial ability in relation to gender. These generic tests would provide some useful information but this research was interested in the participant's ability to use a specific set of spatial data for a specific task rather than a generic measure of spatial skill. The specificity is important because a sonification has to be for a specific application and is not something that can be transferred between different applications without being redesigned. Therefore the general information provided by the test measuring spatial skill may not be specific enough to be relevant to the sonification used.

When measuring learning style, the research faced an additional complication in addition to the issue of how best to evaluate participants learning style without taking up too much of the evaluation time. Often, there is an 'auditory' (or similar) learning style but it is very rare for the 'auditory' learning styles to distinguish between speech audio and non-speech audio. By definition sonification is non-speech audio, so it may well be that even if a participant has a preference for auditory learning, they may prefer speech to non-speech or visa-versa. Therefore the impact of having an 'auditory learning style' may be positive or negative for participants. However the learning style of participants may still reveal some interesting insights.

As well as measuring participants' performance, it was also important to ask them about their experience of the evaluation (e.g. was the interface easy to use, how confident were you with your answer, etc.). This is because participants who were confident with their answer are more likely to be correct (Morse et al., 1996; Smith et al., 2008) as well as more likely to use the sonification tool in a real world environment. Additionally, the participants' preferred method may not be the same as the most effective method and this should be recorded in case it requires further investigation (Jeong and Gluck, 2003; Axon et al., 2012).

3.6 Task

It is important for a sonification of uncertainty in spatial data to be evaluated in a meaningful way. Therefore the sonification needs to be presented to the participants in a contextual setting with a specific task to complete, rather than just as an exercise with limited meaning. This allows the participants to evaluate the sonification in terms of whether (and how much) the sonification helps them complete the task given to them. The first part of this section will consider some theoretical aspects that are relevant to this issue, followed by examples of tasks in previous evaluations.

Participants' perception of audio is very different to their perception of visual data and this difference needs to be taken into account when designing the evaluation. The best method of interaction is an interface that the participant can control themselves (i.e. repeat the sonification at will) rather than one which follows a preset pattern (Hermann, 2008; Walker and Nees, 2011). The process of designing a sonification is one that is relatively complex. There are a number of methods from the design of visual interfaces that are worth applying to sonification. One of these is described in the following quote:

> *"A useful starting point for designing advanced graphical user interfaces is the Visual lnformation-Seeking Mantra: overview first, zoom and filter, then details on demand."*
> *(Shneiderman, 1996, p.336)*

Allowing the user to interact with the data is a prerequisite for this, and one example of this is Zhao et al. (2008) who created a map for the blind using a zoom-able interface using the numeric keypad. This allowed the user to orientate themselves within the data set. This example was a tool for visually impaired users, but is equally applicable to sighted users. It evaluates sounds as a visual replacement method, which is partially relevant to this research. Frauenberger and Stockman (2009) created a new way of designing sonification applications (*paco*), which is an iterative process designed for people who are new to the sonification field (for more details see

Section 1.3). However, due to the complexity of the process it was not possible to use it to help design the sonifications used in these evaluations.

The area of Social Signal Processing deals with how people interpret other people's feelings and attitudes whilst in conversation (Brunet and Cowie, 2012) and it raises three areas that are important for sonifications and evaluating sonifications. However, the paper is not specific to the area of sonification or spatial data, but the generic issues raised can be applied to this research. The macro context is very important in this field and it considers the factors that influence people but are outside of the direct context of the sonification. Figure 25 shows these three issues and the fact that they are all informed by culture. The individual differences are the factors in the conceptual model that will be recorded in the evaluation (i.e. GIS experience, data experience and so on). The other factors in Figure 25 (situation and other people) are aspects of the evaluation that will be kept as similar as possible between participants to reduce their influence on the evaluation results.

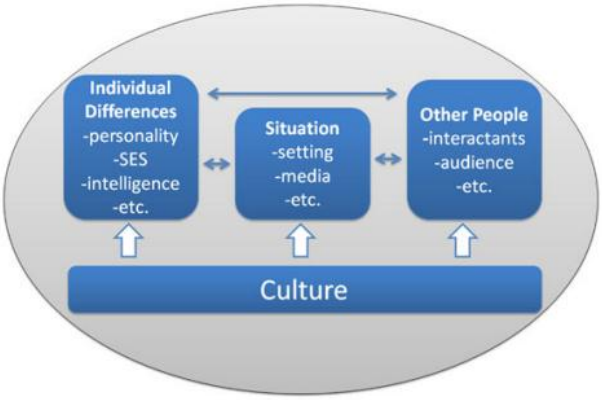


Figure 25. Issues that are likely to affect how a sonification is understood by the user, known as the Macro Context in the Social Signal Processing field (Brunet and Cowie, 2012, fig. 4).

3.6.1 Types of tasks in previous evaluations

Tasks that users have been asked to perform in previous literature can be categorised into three groups; ordering tasks, spatial data tasks and navigation tasks. Ordering and spatial data tasks can be both considered data exploration examples of sonifications (see Section 2.2.5.3). Navigation for the visually impaired is not usually classified as a sonification in the definition presented by Hermann (2008, see Section 2.2.5), but the area does contain relevant examples. One thing that all of the examples have in common is that they are interactive sonifications, i.e. the user is interacting with the computer on a one-to-one basis and the sonification changes with their actions. This reinforces the point made earlier that interactive sonifications are more effective than non interactive ones.

The simpler tasks users are asked to perform are ordering tasks, including ordering a set of values (from highest to lowest or lowest to highest) or identifying particular values (e.g. the middle value or the highest value). Often census or population data are used as examples for these types of exercises (Jeong and Jacobson, 2002; Jeong and Gluck, 2003) or other data such as spatial risk data (Gluck et al., 1999). Often the example has no relation to the skills of the participants and Gluck specifically used students without experience of risk assessment or geographic data to avoid any impact these skills might have. The advantage of this approach is that it ensures that the participants have no additional knowledge that influences the results, but does limit the complexity of the example data used in the evaluation. Zhao et al. (2008) used population data for US states for their sonified map and they also asked participants to complete the same task using traditional tools (Excel in this case) to see how it compared to the sonified interface. The comparison of sonic and non sonic tools is an advantage as it shows how sonification can be used to provide a better way of performing a specific analysis. However Zhao et al. (2008) also encountered issues of the same participants completing the task using both methods, resulting in them reusing information from one task for the other one. They solved this by making participants complete the Excel task first, but there could still have been some information transfer between the two methods. The tool developed by Zhao et al. (2008) is for blind users so the findings may have limited applicability to the aspect of this research that combines vision and audio stimuli because the interface is likely to be used in a different way by sighted users.

A more specialised example of a task requiring users to find specific values is that employed by Tzelgov et al. (1987) where they asked participants to identify high and low radiation areas using a tool like a Geiger counter. They did not use specialised users, but the tool is only applicable to one area which limits the findings to similar types of applications. The task of identifying high and low values works well for simple sonifications and is useful for comparing different sonic stimuli. However, in many cases it does not exploit all of the advantages that sonification can offer, particularly with complex or multi-dimensional data sets. These require a more specialised group of users with particular knowledge of the data sets and/or problems involved. The evaluation of a simple ordering or selection task can only produce a limited amount of results and the majority of spatial data are more complex than this, limiting the usefulness of results from these types of datasets.

More advanced spatial data tasks require the users to have some knowledge of the application area of the data set before they can make effective use of the sonification. One example of this is an evaluation by Harding and Souleyrette (2010) who created a sonification application for highway routing. They used degree level engineering students to evaluate the use of sound in an application to design the route of a new highway. This required the participants to have specialist knowledge (engineering in this case) and is therefore a sonification for a specialist audience (which requires a specialised testing audience) rather than for general users. This work does provide a detailed insight into the use of different senses in highway planning, but did not explore more than one sonic variable and findings from the study are only applicable to the highway planning area. Rabenhorst et al. (1990) created a multiple-window interactive data exploration tool, with sonification as a way of showing additional information. This was designed for experts in particle physics (based on the examples they provide in their work), but they did not complete any evaluation with users. Jacobson and Kitchin (1997) created a GIS for blind or low vision users who are also a specialist group. The interface was evaluated with participants who had experience of navigation with no or limited vision, which provided much more valuable feedback than if a similar evaluation was completed with users with normal vision. This highlights the need for and relevance of recruiting specialised users, which is applicable to applications of using sound to represent uncertainty in spatial data. Nasir (2009) used sonification to represent contour data to users without specialised skills in a particular area. The users successfully used a simple implementation, but the more complex one was very difficult for them to understand. The need for a specialist audience for these advanced sonifications is clear as non experts struggled to make use of the additional information sonification techniques can provide. Using a specialised audience with these types of advanced spatial data tasks allows for more detailed and useful results for the area of data being sonified. If these results can be replicated across a number of different data domains, then they are also more widely applicable. Participants are only willing to provide a limited amount of time for evaluations so the length of the evaluation must be considered when designing evaluations, particularly for the more advanced spatial data tasks because there is potential to ask the participants to perform many different tasks, which may be relatively time consuming.

The other main application of previous evaluations is navigation (Constantinescu and Schultz, 2011; Loomis et al., 1998; Blades et al., 2002). This is primarily for users who are blind or have low vision to assist them in real world navigation. These users could be considered specialist users who have experience of navigating familiar and unfamiliar areas with reduced or no vision. These specialist users provided very useful feedback for this application but their views, like all specialists users, only have applicability to that particular application. Milios et al. (2003) looked at a specific element of navigation, evaluating how blind or blindfolded users could interpret direction from spatial sound and different notes with the intention of integrating this in to a future navigation product. The principle of this technique worked, but participants said they felt they would have difficulty using the technique with larger amounts of data. Techniques of representing direction and distance from these pieces of work are useful, but the techniques need to be implementable with visual representations of spatial data in order to be applied to this research.

The final example is Pauletto and Hunt (2009) who used sonification to look at biofeedback. Their user group was participants who were used to understanding EEG output, either medical staff or patients who were undergoing rehabilitation and so had more knowledge about this area than non experts. They had specialist knowledge of this particular area and this helped the users understand the sonification output. However, the type of sounds and data used in this example are very different to this thesis.

All of the tasks used in these evaluations are limited because of constraints on how much time participants can be asked to give up for the study. These tasks do not fill the full range of sonifications listed in Section 2.2.5, but research in this area is unevenly distributed. The area of data exploration is one that has seen the majority of work in sonification because it is the area that is seen as having the most potential (Grond and Berger, 2011).

3.6.2 Tasks in this study

There are three case studies in this thesis, each using sonification with a different type of data. This section will briefly describe the data sets and tasks for each case study before describing the issues that influenced the design of the tasks. More details on each case study method design are provided in the relevant case study chapters. The first case study used Ordnance Survey's Address Layer 2 dataset and used piano notes to represent the uncertainty of the positional accuracy of each address location. The second case study used the UK Climate Impact Programme's UKCP09 (UK Climate Projections 2009) data set. The uncertainty of the projected future climate for locations over the UK is an inherent part of this data set and trumpet notes were used to represent the uncertainty element of the data. The final case study used piano notes and bird calls to represent distance of wind turbines in a landscape visualisation. The types of data chosen provide a range of different data to see what type of data is most effectively represented using sonification.

The tasks used in this thesis were designed to be advanced spatial data tasks which would be completed by participants with specialist knowledge of the data set used in each case study. An advanced spatial data task was chosen because the current need for sonification is a result of the increasing amount and complexity of spatial data available for users. Representing these complex data are difficult using current techniques and sonification may provide a new way of addressing this challenge. Each case study had a number of tasks for the participants to complete, with a within groups design used. A within groups design is a study where all participants complete all of the tasks and a between groups design is a study is where the participants are split into two (or more) groups and each group completes a different task (Bonebright and Flowers, 2011). The within groups design was used because there were not enough participants available for a between groups design to have sufficient statistical power. A score of percentage correct was calculated for each participant in each case study to show how effective the different representation methods were. This allowed the case studies to be compared and trends for different groups of participants analysed, as well as allowing the different representation methods to be compared within each case study and a broad comparison to be made between case studies.

When creating the tasks it was important to pitch the difficulty of the task at the correct level for the participants to avoid the 'ceiling effect' (where all participants get 100% correct answers) or the 'floor effect' (where all participants get 100% incorrect answers) (Lammers and Badia, 2005). The results needed to be a range of different values in order to see whether a particular group of participants were more or less effective than other groups. In addition to pitching the task at the correct difficulty level, the tasks also needed to be meaningful to the user in order to represent a reasonably realistic setting where sonification may be of use. The whole evaluation session could not take too long to complete because participants were only willing to give up a limited amount of time without formal payment and there were no resources for formal payment within this thesis. Additionally there needed to be enough participants involved in the evaluations for the results to be statistically viable because without this the amount of information gained from the evaluations would have been limited.

While the task that participants were asked to complete may have been the same, different participants may have taken different approaches to complete the tasks. This would affect the time it took to complete the task and/or the overall result for that user. The approach the participant used to answer a particular question was recorded during the evaluation so this could be taken into account in the analysis. One method of doing this would be to record the mouse movements while the user is interacting with the sonification.

In the explanation of the task that was given to the participants, it was important that participants were told that there was no correct answer for their views on how well the software works, as we were interested in their subjective feedback (Bonebright and Flowers, 2011) as well as the objective results from the task. It was also important to remember to tell the participants that they could take as long as they needed so they did not rush to complete the task in as short a time as possible, as this may have increased the number of mistakes they made.

As a part of the aim to make the results as widely applicable as possible, the data sets chosen were different to see how similar sonification techniques worked with different types of spatial data. Due to the differences in the data sets used, the amount the results can be generalised will be limited, but there may be some general trends that can be seen across the case studies.

In order to facilitate evaluation at a number of different locations, such as those of Ordnance Survey and UK Climate Impacts Programme, the technology used for the evaluations needed to be transferable between different locations relatively easily. The evaluations made use of existing technologies and software including ESRI ArcGIS, Google Maps and Microsoft PowerPoint which were all relatively easy to access Microsoft Windows applications that could be replicated in the evaluation environment in a variety of different locations. As well as the software, consideration needed to be given to the headphones used to generate the sound and the testing environment. A set of fifteen pairs of headphones were specifically bought and used for all of the evaluations.

3.6.3 Representation methods used in these tasks

While the tasks varied depending on the data sets used, the representation of the data needed to be similar between the different case studies. Each data set consisted of two main elements, the data itself and the uncertainty surrounding the data. These could either be represented visually or sonically using the techniques mentioned previously. The research was interested in how the uncertainty element of the data could be represented using sound and comparing this to more frequently used visual methods. It also sought to consider the impact of using sound to reinforce data shown visually. There were four main ways of representing the spatial data and associated uncertainty used in this thesis. They are visual only, sonic only, visual & sonic showing the same data and visual & sonic showing different data.

The visual only method of showing the spatial data and the uncertainty surrounding was used as a 'benchmark' and will be compared to the sonification method. It was intended that the visual only method was synonymous with the methods that are used currently to show uncertainty data. The sonic only method evaluated how spatial information could be represented using no visual stimuli, but did not form the main focus of this thesis. Methods using vision & sound were the more interesting representation methods and consisted of different ways of combining vision & sound to represent spatial information. Using vision & sound to show the same data was designed to see whether reinforcement of information shown visually with sound improved performance. The vision & sound method showing different data was used to show whether uncertainty information could be shown sonically in relation to the spatial information shown visually. Different sonic methods were compared to the visual only benchmark to see what improvement (if any) sonification provides. For the sonic representation methods, a sonic legend was used in a similar way to the visual legends for the visual representation methods, where a user considered a point on the map, visually checked it on the legend and then looked back it the map. This would not work in an audio only interface because it would be very difficult to re-establish orientation on the map (Rice et al., 2005). However, retaining the visual interface to the map allowed the user to establish their location which avoided this issue.

It was therefore decided to include each of these four representation methods in each case study, which allowed comparison between the different representation methods and between the different data sets. However because of the way the different case studies were implemented, not all of the results were directly comparable (see Table 3). The discussion chapter (Chapter 7) will discuss the comparability of the results from the three case studies.

	CS1 – Address Layer 2	**CS2 – UKCP09***	**CS3 – Distance of Wind Turbines**
Visual Only	Comparable across all three case studies		
Sonic Only	Comparable across all three case studies		
Visual & Sonic – reinforcement of the same data	Comparable across all three case studies		
Visual & Sonic – representation of additional data	CS1 and CS2 are comparable with each other. CS3 did not have any additional data, so cannot be compared to CS1 and CS2.		

Table 3. The comparability of the three different case studies and the four different representation methods. *The UKCP09 data were represented over two separate maps because of the complexity of the data, so there may be limitations on how CS2 can be compared to the other case studies.

The order of the representation methods was considered carefully to avoid the possibility of a learning effect by participants being able to apply information gained from one representation method to another. Within each case study, a number of different data sets were used to again avoid the possibility of a learning effect. More details on these aspects will be explained in the individual case studies, as the approaches used varied depending on the data set.

3.7 Evaluation Background and Setting

This section considers the importance of evaluating the representation methods. Once the method of representing uncertainty in spatial data as a sonification has been created, it needs to be evaluated. Poor methods of representing uncertainty data are likely to produce worse outcomes than not showing the data at all. Poorly designed visual displays will perform poorly (Sanders, 1992) and similarly poorly performed auditory displays will perform poorly (Walker and Kramer, 1996; Walker, 2002). Additionally, if the user is not familiar with sonification display in general, they may not realise they are using a poorly designed interface and think that the poor representation is the limit of sonification. If they do, this will damage the wider reputation of sonification. Therefore it is important to ensure the sonification method is designed correctly.

When a sonification is designed, it is important to be able to evaluate the effectiveness of the sonification to find out which aspects of the sonification are successful and which aspects need to be removed or improved. Often the type of evaluation is dependent on the type of task that is used with the sonification. There are a number of different sonification types (e.g. audification, auditory icons / earcons, parameter based mapping, model-based sonification, see Section 2.2.6 for more details), but often, particularly for data exploration tasks, the sonification implementation will be highly customised, which will not allow the use of a generic, out of the box evaluation technique (Ibrahim and Hunt, 2007).

The type of evaluation used will very much depend on the task the user is asked to perform. For the three groups of tasks identified earlier; ordering tasks, spatial data tasks and navigation tasks, each will have a different way of being evaluated. Basic ordering tasks are easy to evaluate in terms of whether the user has put the data in the correct order, or selected the correct option. How any incorrect options are marked, and what weighting they are given can vary, depending on the approach taken (Gluck et al., 1999). More advanced spatial data tasks can be more complex to evaluate, as there may not be a simple 'right' answer. However, in all cases there will also be a qualitative element to the results, whether it is in the form of a questionnaire or interview (Harding and Souleyrette, 2010), which will supplement the quantitative answer. Navigation tasks usually involve some form of real world navigation which can be marked on completion time, total distance travelled, number of incorrect turns and so on (e.g. Loomis et al., 1998). Very commonly all three types of task will be timed, with quicker times viewed as better in the evaluation of the tasks (Lloyd and Steinke, 1985).

Additionally we are trying to measure usability, which is not something that can be measured directly, but needs to be measured through usability parameters instead. There are two categories - subject user measures (such as how much users like the system) and objective performance measures (such as how capable users are of using the system) (Nielsen and Levy, 1994). These are also known as quality of experience and quality of service respectively (Wechsung et al., 2012). These measures can sometimes contradict each other, particularly when the user is given two static systems that cannot be changed. It is not uncommon for one to offer better service and the other to offer a better experience, so there are likely to be positives and negatives in both. Ideally, the positives from both systems could be combined to make a system that was productive and that the user preferred.

3.7.1 Evaluation environment

When conducting these evaluations, the environment they are completed in is important. Given that the evaluation was trying to establish the effectiveness of sonification as a solution to a real world problem, ideally the testing environment needed to be as similar as possible to the real world use environment. For example, a method of sonifying stock market data was tested in a quiet lab where it worked well, but when used in the real world on the trading floor, the sonification could not be heard and therefore was ineffective (Bonebright and Flowers, 2011). The tool created in this thesis were intended for use in an office environment which is likely to be relatively quiet (similar to the evaluation setting) but the noise level will vary for different offices.

As a result of expert users being approached for each case study, the evaluations were run in a number of different locations to enable greater participation. The main testing facility was the Experimental Lab in the ZICER building at UEA. This is designed as a testing environment for multiple users with restricted views so participants cannot see other participant's screens (see Figure 26). When the evaluations were completed off site, conditions were made as similar as possible to the testing environment, within the range of possible options using the facilities available. In order to maintain consistency, the same headphones were used for all the evaluations (BL-882 Stereo Headphones[7]).

Figure 26. Photo of the experimental lab where the evaluations took place. Each participants can see his/her computer screen, but none of the other participants' computer screens.

As well as the physical environment, consideration needs to be given to the environment experienced by the user. It has been shown that completing a task in a stressful environment (being subject to variable, unpredictable noise) effects participants ability to remember the locations of specific objects and put those locations in order (Evans et al., 1984).

[7] http://www.cclonline.com/product/36785/BL-882/Headphones/BL-882-Stereo-Headphones/SPK3224/

3.8 Evaluation and Data Collection Methods

During the development of the prototype for each case study, a pilot testing phase was used to test out the sonification method and the evaluation to ensure it would generate useful results. The tasks the participants were asked to complete needed to avoid floor and ceiling effects as discussed earlier, and the pilot testing was crucial to this to ensure a wide range of answers in the evaluations. These pilot sessions were also used to work out how many repetitions of the sonification could be performed within a timescale that was reasonable to ask participants to give up for the evaluation, as well as verifying that the evaluation questions were understood correctly. The details of the pilot studies varied for each case study, as they were using different evaluation tools, and more details are provided in each case study chapter.

An important point from the research literature is that it is important to have a training session to ensure that participants are familiar with the interface before the main part of the evaluation. Many of the studies in the literature review found the training element to be useful (such as Nasir, 2009) and this is particularly important with a topic where sonification is new to the majority of participants. The training element also takes up time in the evaluation so the need for training has to be balanced with the need to collect a sufficient amount of data from the participants while still remaining within an acceptable amount of time for the whole evaluation.

One data collection method that is not widely used in this area is verbal protocols (also known as think aloud protocols), where participants are required to speak aloud while they perform the evaluation, explaining why they are making specific decisions and what they are thinking about whilst doing this. The experimenter can also prompt them or ask specific questions as the evaluation is completed (Bonebright and Flowers, 2011). Using a verbal protocol data collection methodology would generate a significant amount of in-depth data for each user, but would limit the number of users that could be involved because of limits of time and money as the evaluations would need to be completed on a one-to-one basis. Additionally a recording of this activity does not always provide a complete representation of the evaluation as the user’s monologue and the application sounds may interfere with each other (Bonebright and Flowers, 2011). The alternative to these types of in-depth evaluation is to use methods such as multiple choice or free text answers, which would also allow more participants to complete the evaluation for a given level of resources. This study required larger sample sizes (see Section 3.5) so a verbal protocol was not used.

3.8.1 Data collection methods

For the data collection process, a variety of different types of questions were used. Multiple choice questions were used where possible as these are the easiest and quickest for participants to answer (Dillman et al., 2009). For a number of questions Likert scales were used; a bi-polar rating scale with between 4 and 7 options (Ware, 2004). At the extreme ends of the scale will be two polar opposites (such as ‘very interesting’ and ‘not interesting at all’) and the user can choose one of the options between the polar opposites to represent their view. The most common number of options is between 4 and 7, but occasionally evaluators used 8 or 9 options. There is a practice to use odd numbers to allow participants to give a ‘middle’ or ‘no preference’ answer, although opinions on this vary (Bonebright, 2012, pers. comm.). The number of options used in this research varied between questions, but where relevant, included an odd number to allow a ‘no preference’ or similar answer from the participants. Minor changes in the wording of a question can result in significantly different responses from the participants (Crawford, 1971; Flannery, 1971; Shortridge and Welch, 1980), so the instructions needed to be clear, consistent and piloted to check that they worked in a consistent manner.

Making use of both quantitative and qualitative data is important to cover the whole gamut of information available from participants using sonifications (Bonebright and Flowers, 2011) and this principle was applied to all three case studies. A very well developed field with very specific research aims would allow exclusive use of multiple choice quantitative questions, but this study also wanted to develop the field of sonification so the participant’s response to the sonification was unknown. Therefore the provision of qualitative questions was important to ensure as much information as possible was collected. Additionally, multiple methods in evaluations are often used to triangulate results, which increases the confidence in the conclusions (Li and Reeves, 1999).

Computerised evaluations were used as much as possible to reduce experimenter bias, so to ensure that all participants received exactly the same stimuli and the investigator did not unintentionally provide additional information to some participants and not to others. Also, standard documentation was provided to all participants. Where possible, participants’ responses were recorded electronically to reduce the workload in terms of inputting data for analysis and to minimise the potential of recording and transcription errors.

When conducting a usability study, the results are susceptible to experimenter and participant bias, including situations when the participant tries (consciously or unconsciously) to provide the ‘correct’ answer to the question, even if the experimenter assures them they want to hear their real view rather than the ‘correct’ answer (Bonebright and Flowers, 2011). One solution to this is to automate or use scripts for as much of the experimenter participant interaction as possible to reduce possible bias between groups. For these evaluations, the majority of the interaction was in the discussion phase, which was after the main data collection took place, so any influence here would have had less of an effect. One other technique is to hide the aims of the overall research from the participants so they do not know what the evaluation is trying to find. This does require a large testing community, which for this thesis would have needed a large number of individuals with GIS skills. Unfortunately due to the limitations of the resources for this thesis and the nature of the area, this was not possible because there are not enough available participants with sufficient GIS knowledge who were not already aware of the research. However, the interest in this area was used to recruit participants and hopefully resulted in greater engagement with the evaluation.

3.9 Evaluation Structure

The evaluation sessions were run in small groups and generally consisted of an introduction, a computer based evaluation and a group discussion at the end. The overall completion time was around one hour, which was around the maximum it was thought that participants could be asked to volunteer their time without formal compensation. The total time taken depended on the length of the discussion session, which varied with the participants. The style of the computer based evaluation varied with the methodology chosen; for more details see the relevant methodology sections in the case study chapters.

A general introduction to each session was given to participants, outlining some background to the evaluation and the tasks they were going to be asked to complete. A prompt sheet was used to ensure this was consistent between all sessions. Participants were also asked to read and sign the participant information sheet and consent form. They were also asked some introductory questions using the computer based interface before the main evaluation itself. The spatial data task followed the introduction. As well as the task itself, the participants were asked a number of questions about their experience of using the sonification and future preferences if they were to use a sonification tool again.

Given that there were a number of different iterations of the evaluation, each with different representation methods (i.e. visual only, visual & sound etc.) as well as different data types in some cases, it is ideal practice to randomise the order of these to avoid any learning effects (Bonebright and Flowers, 2011). This randomisation avoids the effect that completing the tasks in the same order may have on the results. The randomisation was approached in different ways for the different case studies because of particular limitations on the interfaces used.

Group discussion sessions formed a major part of the evaluation for all three case studies. Hopkins (2007) found that smaller focus groups (< ~6) worked, best as with larger numbers not everyone can make their views heard and people often end up speaking over one another. This was the case with this research and smaller groups (<8) were essential, although some groups of two or three tended to have quite stilted discussion. However, this experience did vary depending on the nature of the participants involved in the group.

A qualitative discussion session allowed the users to express any views they had about the sonification without restricting them to an interpretation set by the evaluation design. This was semi-structured to give some focus to the discussion, but if participants brought up an interesting point that merited further discussion, this was allowed to continue. The quantitative side was fulfilled by a task users were asked to perform using the sonification, where their accuracy and the time taken were measured. This allows for comparison of the different tasks in each case study (e.g. visual methods or sonic methods) as well as comparison between different user groups (e.g. experienced users of a data set against inexperienced users).

This chapter has outlined the methodology used for the three case studies in this thesis. Each case study chapter that follows includes more detailed information on the methodology used for that case study.

Chapter 4. Case Study 1: Using Sound to Represent Positional Accuracy of Address Locations

4.1 Introduction

This is the first of three case studies, each of which evaluates ways of representing spatial data using sound with different examples of spatial data. This chapter is split into a number of sections which consider possible data sets to use in this case study, the methodology used to collect the data, and the analysis and discussion of the results.

All spatial data have some uncertainty concerning their positional accuracy. However, even when information on the level of accuracy exists and is potentially important in terms of fitness for use, it is widely recognised that it is often ignored by users of the data. Several studies have investigated the reasons for this situation and how it can be addressed (Agumya and Hunter, 2002; van Oort and Bregt, 2005). One aspect is the technical challenge of representing uncertainty alongside other attributes of the data, but there is also a need for research to evaluate how the outputs of different techniques are understood by users (Hunter and Goodchild, 1996; MacEachren et al., 2005). This case study seeks to contribute to the literature on these two topics by presenting a sonic technique to represent positional accuracy in Ordnance Survey's MasterMap® Address Layer 2 data and assessing how well it can be used by experienced spatial data users.

The representation of uncertainty in spatial data has been widely discussed in the research literature and visual methods such as colour, blurring or multiple maps are most common (Ehlschlaeger et al., 1997; Appleton et al., 2004; MacEachren et al., 2005). However, these methods can obscure underlying data or limit the amount of information shown. There has been extensive work on tools for visualising spatial data dynamically, resulting in many novel methods of exploratory analysis including interactive visualisation (Dykes et al., 2005) and more recently development of the newer field of geovisual analytics (Andrienko et al., 2007). The vast majority of these displays, however, only employ visual stimuli and displays are reaching higher levels of complexity, sometimes pushing the limits of visual comprehension (Tukey, 1990).

The use of other senses has been suggested as a sensible complement to visual displays. Haptic (touch) maps have advantages for comprehension and can be combined with visual displays relatively easily, but require significant investment of time and money to implement (Golledge et al., 2005). The hardware to utilise sound (sound card and headphones/speakers) is cheaper and more readily available than the hardware required for haptic and consequently has been more widely utilised in research. However, much still remains to be understood in order to make the most effective use of sound in these situations (Pauletto and Hunt, 2009).

4.2 Potential Data Sets

When considering which data to use in this study, one of the first requirements of a potential data set was that it contained uncertainty information. A number of data sets from Ordnance Survey were considered, but only a limited selection have explicitly stated uncertainty data included within them.

Even though uncertainty is inherent in geospatial data because of the collection process (see Section 2.4.1) many geospatial data sets do not have uncertainty included explicitly in them. Of the ones that do, many of these have data-set level uncertainty, such as RMSE (root mean square error) or other uncertainty about the dataset as a whole rather than individual objects within the data set. This level of information is useful, but does not provide any information about how the uncertainty varies within the data set. It may also be slightly misleading if the user is only provided with a data-set level measure of uncertainty and is not aware of the possibility that uncertainty may vary within the data set.

Object level uncertainty can come in two forms, implicit or explicit. Explicit uncertainty data is data where there is a specific field (or fields) in the data set for each object that records the measure of uncertainty. There may be more than one field if more than one type of uncertainty is represented. Implicit uncertainty data is data that does not have a specific field for the uncertainty information. In this case the measure of uncertainty would need to be worked out on a case by case basis. Three different object based data sets will be discussed, change detection, generalisation and address matching.

4.2.1 Change detection

Change is constant in geospatial data and Ordnance Survey has a programme for capturing this change. This is where current geospatial data are updated with new information collected from the field, either in person by a surveyor, or from remotely sensed imagery. This process is starting to become automated and there are a number of decision algorithms which can be used to decide whether a particular object has changed. When the changes were updated manually, a user would decide whether they were ‘certain enough’ that the change had occurred, but with the automation there is a need to measure the uncertainty and flag particularly ‘uncertain’ changes for human review. The frequency that an object changes could be utilised to predict how likely an object is to change in the future, and therefore how likely the current information is to be out of date. This would require input from a number of different elements of the data set to generate a measure of the uncertainty associated with a specific object, including the date of acquisition of the new data, the date of acquisition of the original data and a comparison to see if the gap between the two is great enough for a change to be considered ‘likely’. The threshold for a likely change would be the most complex aspect to calculate but theoretically could be done. The object type and dates of previous changes to an object are already included in the data set as shown in Figure 27.

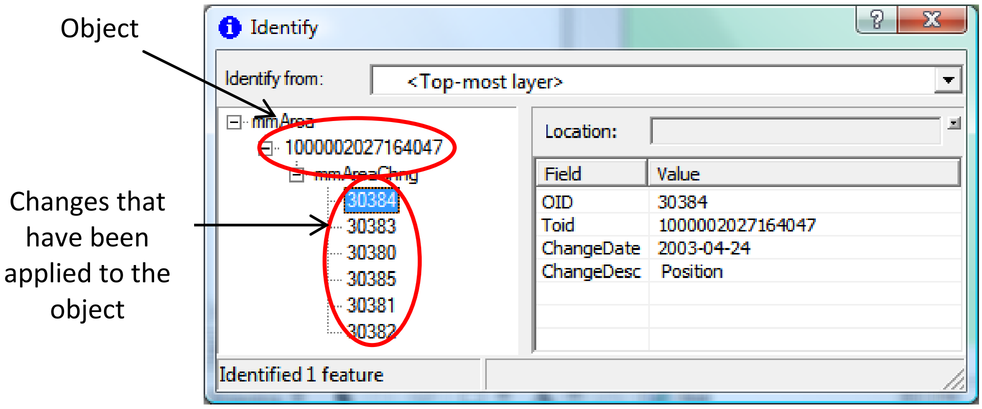


Figure 27. An object from MasterMap Topology layer, showing the object and its change entries.

The uncertainty in this data is implicitly stated, in terms of the object type and dates of previous changes. There is no explicit uncertainty information, which means that the measure of uncertainty would have to be calculated from the existing data before it could be sonified. It would also require significant work to calculate the thresholds for expected change for a variety of types of objects. While there is great potential with this to display the uncertainty in a easily accessible form, it is still very much at the theoretical stage and a vast amount of work would be required to add ‘uncertainty’ information to the data before it could be represented. Therefore it was decided not to use such data for this case study because of the amount of work involved.

4.2.2 Generalisation

When cartographic data are generalised to allow display at different scales, particularly very small scales (e.g. 1:1,000,000) features in the data set may be moved from their true location to avoid conflict with other features or enlarged to ensure the more important figures remain clearly visible. This displacement creates a form of uncertainty within the map that could be represented using sonification. The amount of displacement will vary between different spatial objects and will also vary across the spatial data as a whole (see Figure 28 for an example).

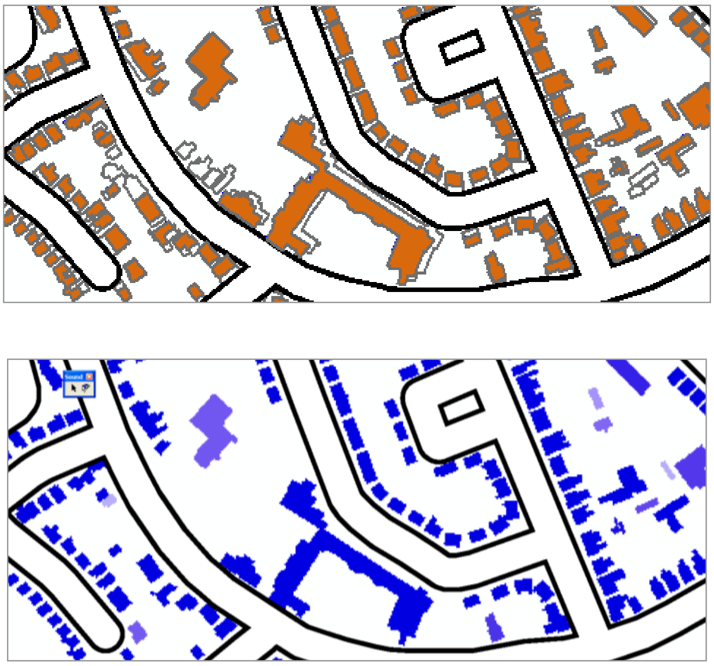


Figure 28. Generalisation example with the original location of the buildings shown in a grey outline, the new location shown in orange (both above), and the displacement shown in blue (below) with a darker blue representing a greater displacement (after Bearman and Fisher, 2012, fig. 6). The data shows Shirley Warren, Southampton; the road data are from the ITN Layer and buildings from the topography layer of Ordnance Survey MasterMap® (Ordnance Survey, 2008). The building displacement was calculated using Radius Clarity (1Spatial, 2008). *Ordnance Survey. © Crown Copyright. All rights reserved.*

When a data set is generalised a measure of the displacement that has taken place is recorded within the dataset, which can be used as an explicit measure of the uncertainty. This would then be easy to sonify, using sound to represent the displacement amount and therefore the uncertainty of the objects spatial location. However the need for a solution like this is unclear, as the aim of the generalisation process is to make changes to the data that will not have an impact on the use of the data in the final product. Equally, the displacement itself will have a degree of spatial autocorrelation which may not exist in other object based data sets. Case Study 2 will consider continuous data sets and how spatial autocorrelation may impact the sonification method used. Therefore it was decided not to use generalisation as an example for this case study.

4.2.3 Address matching

Address matching is the process by which a text postal address (e.g. 86 Amderley Drive, Norwich, NR4 6JH) is matched to a geographical location (e.g. TG 20981 05977). This is of particular use to companies who have large databases of addresses of their customers and want to work out the geospatial location of their customers for a range of reasons including analysing purchasing trends, considering where to focus a marketing mail campaign or to work out where to locate a new store. They may also wish to relate information about their customers to other spatial data, such as demographic or census data in order to understand more about their existing customers and potential future markets. Ordnance Survey provides a product called Address Layer 2 (AL2) which provides a list of all the postal addresses in Great Britain and their spatial location. This then allows the companies who purchase it to match their customers' postal address to a spatial location.

The process of matching customers' addresses in the company's database to the entry in AL2 is a relatively simple text string matching process. However, the most common support question received by Ordnance Survey relating to the use of this data set is 'why isn't this customers' address in the correct location?'. Often the users of AL2 have not looked at the underlying metadata available in AL2, which specifies the method used to capture the information on the addresses location and the accuracy of that point. For example, a point may be derived from accurately located points within the Postcode Unit of surrounding properties, which can be 10 meters or more away from the true location of the property.

The AL2 data set includes measures of accuracy, in the form of Positional Quality Accuracy (PQA) flags which record how accurately the address is geocoded. There is limited awareness of the impact and use of PQA flags in the AL2 dataset within the data user group. While the lack of awareness of the data set can be addressed by training, sonification offers a way of representing this information in a new and different way that may be able to address the issue of representation as well as providing a good example of why being aware of these flags is important. Due to the easily available uncertainty data (the PQA flags) and the need for a method of representation of these data, the Address Layer 2 dataset was used as the spatial data example for this case study.

While all of the data sets considered above have some form of uncertainty within them, it is in a range of different forms. This thesis is concerned with the representation of existing uncertainty data rather than the creation of usable uncertainty data. Due to time constraints, it is only feasible to use geospatial data with explicit uncertainty information in this thesis.

In September 2011 a new product called AddressBase was released as an alternative product for Address Layer 2 (Ordnance Survey, 2012a). AddressBase contains broadly similar information, and additionally includes information from local government's National Land and Property Gazetteer. This product aims to create a single definitive spatial address database for Great Britain moving from the three databases that used to be involved in the process. When the evaluation reported here was designed and run, this product was not available, but it does include a similar measure of accuracy (the Representative Point Code), which could be used in a similar way to the PQA Flags (Ordnance Survey, 2012b).

4.3 Methodology

This section builds on the Methodology chapter that is used for all three case studies as well as the discussion of possible data sets for this case study in Section 4.2. It discusses the details of the AL2 data set used, the sounds used to represent it and the details of the evaluation design and programming.

4.3.1 Address Layer 2 status flags

The AL2 data set provides a list of postal addresses with associated spatial co-ordinates and allows users to geocode their own address data through text matching procedures. It also includes a set of status flags for positional accuracy of addresses to indicate how accurate the spatial co-ordinates are. While significant numbers of papers have commented on the lack of uncertainty data within spatial data sets used by wide numbers of people, AL2 is one data set that does include this uncertainty information. However, it would appear that sometimes these data are not used even though they are present. Telephone interviews with Ordnance Survey Account Managers and Pre and Post Sales staff suggested that often these status flags are not properly considered when the data are used by external organisations. The reasons given for this situation varied, but included the users not being aware that the status flags existed or of their relevance, as well as difficulties in representing the accuracy information in a meaningful way. The reasons may also include a lack of awareness of the impact of uncertainty. A required, structured form for the uncertainty data may raise awareness of the issues surrounding it, and the EU INSPIRE project has the potential to deliver this.

Within the AL2 data set, there are a range of five different fields providing information on the accuracy of each address data point. They are OS Match Status, OS Physical Status, OS Position Accuracy, OS Position Status and OS Structure Type and they cover a wide variety of information, including links with the Royal Mail PAF dataset, the physical state of the building and the method of position recording (see the technical specification, Ordnance Survey, 2009, for more details). When the spatial accuracy of the address is verified as sufficient for the intended use of the product, the five fields can contain a large amount of repetitive information. However when the spatial accuracy is not sufficient, there is a wealth of information on why and how it may be inaccurate.

The main status flag of note is that of OS Position Accuracy, also referred to as Positional Quality Accuracy (PQA). This has five different possible values, as described in Table 4.

Value	Description
Surveyed (most accurate)	Always within the building that defines the addressed premises or close to the addressed structure where the structure is not a permanent building.
Approximate	Approximate position, usually within 50 m, but may exceed 100 m where the addresses relate to temporary or other structure types (for example, caravan sites or under construction) or within an industrial estate.
Postcode Unit Mean	This indicates that the address has been allocated a mean position within the postcode unit. This mean position is derived from those addresses within the unit that have been accurately located.
Estimate	Where no addresses have been accurately located within a postcode unit, an estimated position, if available, is allocated. The accuracy of this reference will vary. In England and Wales the reference given has a resolution of 100 m; in Scotland the reference given is to 10 m.
Postcode Sector Mean (least accurate)	Where no addresses have been accurately located within a postcode unit and no approximate reference is available, the address is allocated a mean position within the postcode sector. This is derived from those addresses within the sector that have been accurately positioned. NOTE: Postcode sector means coordinates can be several kilometres away from the true address location in rural areas.

Table 4. The five different values for OS position accuracy field in Address Layer 2 data set (after Ordnance Survey, 2009).

Each postal address has an entry in the AL2 dataset, as well as various objects without a postal address (such as street lights). Each entry is given a unique TOID (TOpographic IDentifier) by the Ordnance Survey, which uniquely identifies it within the OS MasterMap suite of products. Each address is also linked to the Topography layer of MasterMap which relates an address to a physical building. The Topography layer was used to provide the visual contextual data used in the evaluation.

4.3.2 Sounds and pilot study

The literature review (Chapter 2) discusses the different types of sound that could be used for sonification and their advantages and disadvantages. Given the categorical nature of the information it was decided that variations in pitch (i.e. different notes) would be an appropriate sonic variable to employ. The specific instrument chosen was a piano because it is a common instrument, which the vast majority of people recognise and the notes have a very clear order to them (low to high). Therefore it should be easy for the users to relate the scale of notes to the different accuracy values.

A pilot study with a small number of users was conducted to evaluate the effectiveness of the number of status flag categories, the structure of the questions and the nature of the overall evaluation itself. One of the main pieces of feedback was that five different notes (for the five status flag categories) were too difficult to discriminate between and then relate to the status flag categories. It was decided to reduce the five categories to three by combining Postcode Sector Mean with Estimate and Approximate with Postcode Unit Mean (see Table 5).

Initial Five Classes for Status Flag	Reduced Three Classes for Status Flag
Surveyed (most accurate)	Surveyed (most accurate)
Approximate	Postcode Unit Mean
Postcode Unit Mean	Postcode Unit Mean
Estimate	Estimate (least accurate)
Postcode Sector Mean (least accurate)	Estimate (least accurate)

Table 5. Positional accuracy status flags 5 original classes (left) and how they were reduced to 3 classes (right).

These three categories were represented using the piano notes C_4, G_4 and E_5, based on the CEG triad split over two octaves (see Figure 29). A triad was chosen because these are sets of notes which sound harmonious together (Burrus, 2009) and CEG was the favoured option in the pilot study. Previous research (Bearman, 2008) has also shown that these types of scales work much more effectively than a scale of all the possible piano notes within a range. The highest note (E_5) represented the most accurate level and the lowest note (C_4) represented the least accurate.

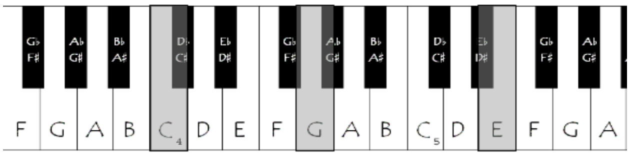


Figure 29. Piano keys with C_4, G_4 and E_5 highlighted.

4.3.3 Evaluation implementation

An ArcScript was created to allow visual and sonic representation of the positional accuracy within the AL2 data as well as controlling the evaluation procedure. ArcScripts allow custom extensions to the ArcGIS interface, programmed in Visual Basic for Applications (VBA). They are mainly used for custom analysis of data; a number are available to download from the support site (http://resources.arcgis.com/) and some have been packaged into extension tools for specific subject areas, such as 'Hawth's Tools' for ecologists[8]. The author had constructed a similar tool for his MSc dissertation, using VBA in the ArcGIS interface, but the tool for this case study was completely rewritten, without reusing any of the previous code. This tool was developed in ArcGIS 9.2, but will work in ArcGIS 10 with an appropriate VBA license from ESRI. This approach was chosen because it allows the addition of sound within an existing GIS environment which participants use frequently. The author was also already familiar with the programming language which reduced the development time required.

The ArcScript tool utilised Ordnance Survey MasterMap® Topography and AL2 data for two terraced streets in the city of Norwich, UK. The AL2 positional accuracy data were linked with the

[8]Available from http://www.spatialecology.com/htools/tooldesc.php

Topography data using the associated TOIDs. Figure 30 shows an example screenshot of the display, with the shading of the building polygons providing a visual representation of address positional accuracy. A sound legend (in the top right-hand corner) was also provided to link specific notes to particular categories. These notes were played as the participant moved the mouse cursor over each building (without needing to click on polygons), allowing them to either query a specific building or scan an area of data to get an overall impression. See Appendix B.4 for details of the files included in the DVD (\Chapter 2 – CS1\), and a similar video is available online at http://vimeo.com/14502978.

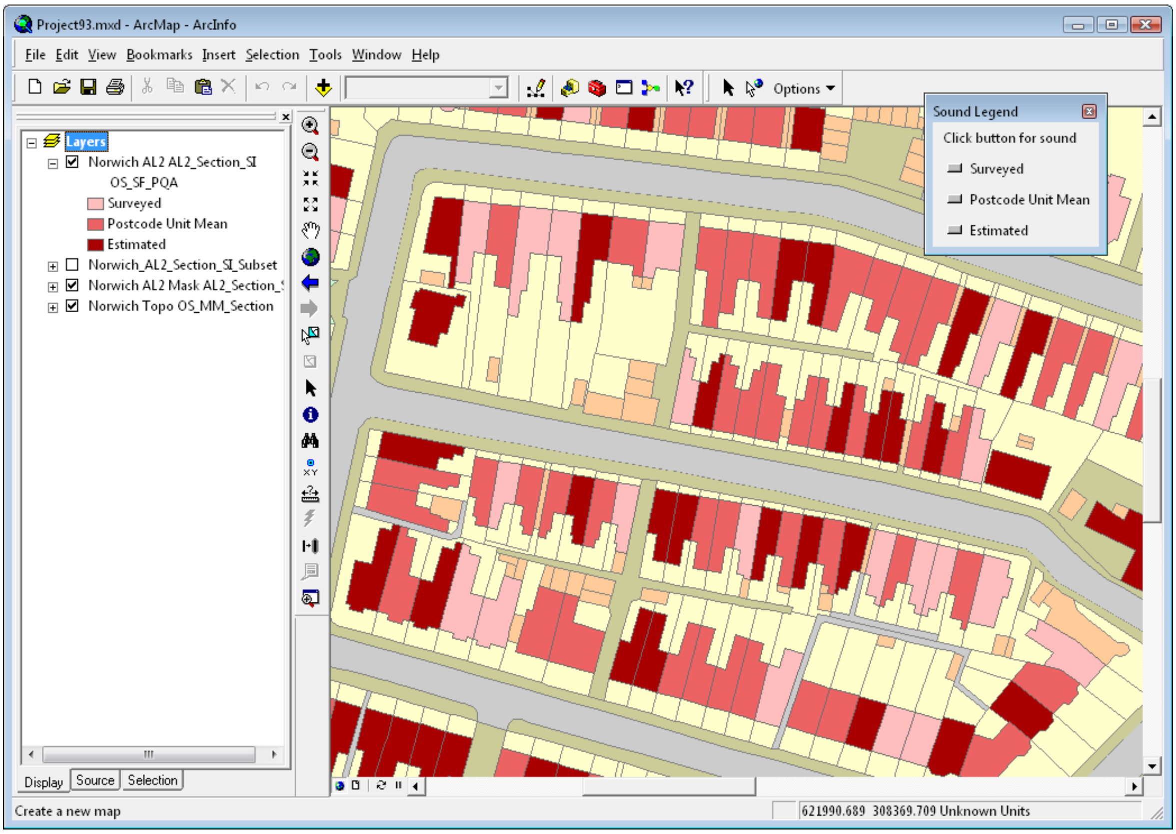


Figure 30. Screen shot of ArcMap showing the positional accuracy classes represented visually and with the sound legend in the top right-hand corner. See video referenced in Appendix B.4 for example. *ArcMap Interface © ESRI 2010, Ordnance Survey. © Crown Copyright. All rights reserved.*

During the survey of participants the data were presented using four different methods in the order shown in Table 6. The four different methods were analogous to the four different representation methods outlined in Section 3.6.3. In all four methods the Topography layer was always depicted in order to provide baseline spatial reference. In the sonic only method (hereafter Sonic) the Topography polygons were shaded according to feature class categories and the AL2 Positional Accuracy was represented using sound (see Figure 31). Visual only (Visual) showed the building polygons with different shadings to represent their AL2 Positional Accuracy. The third method used both sound and visual methods (VS Same) to show positional accuracy in a complementary manner, while the fourth (VS Different) used sound to represent AL2 Positional

Accuracy and a visual shading to depict fabricated "Council Tax band" (local property value) information. Council Tax was used as an additional variable to see whether multiple pieces of information could be displayed at once, with Council Tax band shown visually and AL2 Positional Accuracy shown sonically.

Presentation Method	Abbreviation	Visual Data	Sonic Data
Sonic only (see Figure 31)	Sonic	Topography feature classes	AL2 Positional Accuracy
Visual only	Visual	AL2 Positional Accuracy	None
Visual and Sonic representing the same variable (see Figure 30)	VS Same	AL2 Positional Accuracy	AL2 Positional Accuracy
Visual and Sonic representing different variables	VS Different	Council Tax bands	AL2 Positional Accuracy

Table 6. The four different presentation methods (in the order they were shown to survey participants) and which data were represented visually or sonically.

For each presentation method, the survey participants were asked to identify the proportion of Surveyed values (i.e. the most accurate), from options of 25%, 50% or 75%. A stratified random method was used to assign the status flag values to the building polygons, so each building had a different value for each of the four presentation methods. However, there were always either 25%, 50% or 75% of properties with the value Surveyed. This meant that for each presentation method approximately a third of participants had each proportion of Surveyed values represented. The task of identifying the proportion of Surveyed values was chosen because it combined a question that was not too easy or hard with the need to interpret sound in a way that visual representations are often employed. For the final presentation method (VS Different) participants were asked about both the proportion of houses with the AL2 Positional Accuracy value of Surveyed and the proportion of houses in Council Tax band category 'A&B' (from options of 25%, 50% or 75%).

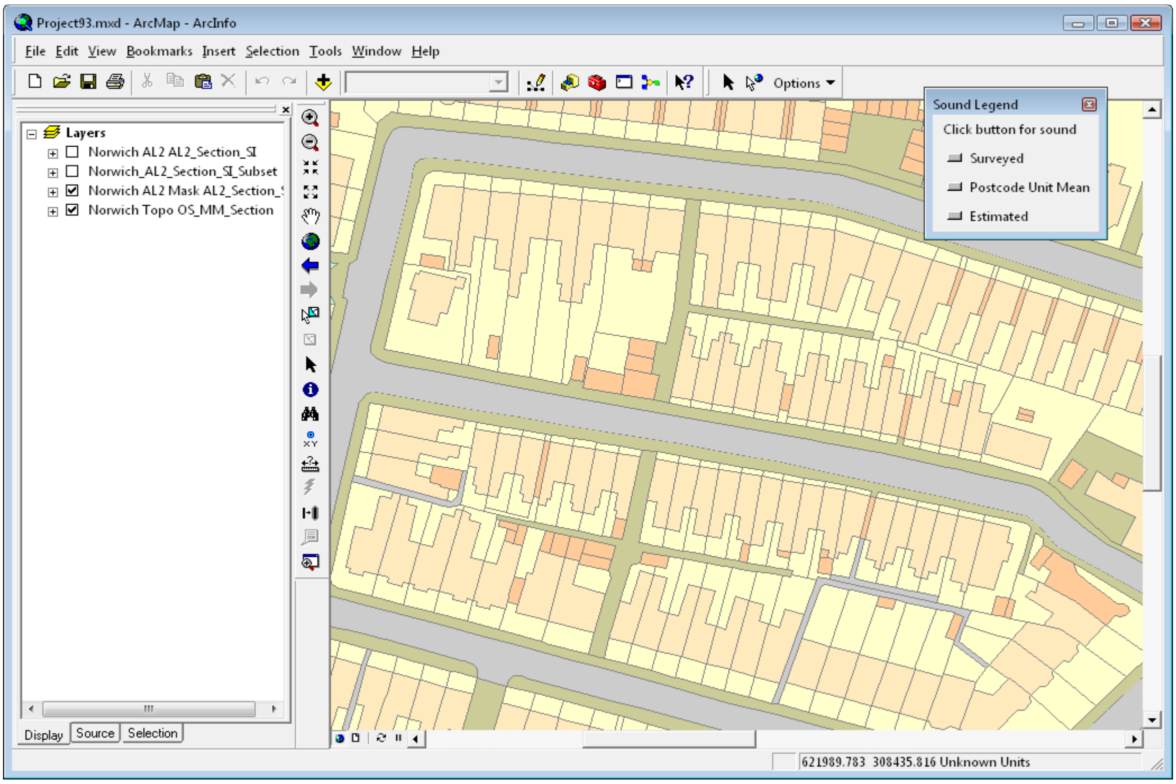


Figure 31. Screen shot of ArcMap showing the sonic only representation method. See video referenced in Appendix B.4 for example. *ArcMap Interface © ESRI 2010, Ordnance Survey. © Crown Copyright. All rights reserved.*

As well as the main computer-based task, during the survey other background questions were asked concerning musical skills, learning preference and experience with AL2 data. The questions covered a range of different areas including people's preferred learning style. However, the learning style question was a relatively basic implementation compared to the way it was implemented in subsequent case studies. For a full outline of the questions, please see Appendix B.1 (page 247).

The evaluations were performed in small groups (~4 – 6 people) and were followed by a discussion session (20 – 30 minutes). The discussion session was in the form of a semi-structured focus group, where the participants were asked their opinions on the experiment, highlighting aspects that did or did not work well and suggesting any possible improvements. The participants were allowed to steer the discussion themselves, with prompting from the evaluator when required. The discussion was recorded for subsequent analysis.

Participants were recruited from Ordnance Survey and local authorities who were thought to have experience of using AL2 data as well as general experience of using GIS. Participants were also recruited from the University of East Anglia as participants who would have less experience of AL2 than those from other organisations. The sessions at UEA were held in the same experimental lab in the ZICER building which can hold up to 20 people, but group sizes were limited to ensure everyone could take part in the discussion session (see Section 3.7.1 for details).

Sessions at the other locations were held in similar rooms where possible. Headphones with adjustable volume were used to provide the auditory stimuli (for more details see Section 3.7.1). Participants were given a copy of the Participant Briefing (Appendix B.1) before they attended the session and a copy of the Participant Information Sheet and Consent Form (Appendix B.2) on arrival.

This section has explained the methodology used for this case study, building on the methodology outlined in Chapter 3. The next section will consider the results from these evaluations.

4.4 Results and Discussion

4.4.1 Participant characteristics

A total of 49 participants completed the survey, consisting of 19 staff from Ordnance Survey, 23 from the University of East Anglia and seven from local authorities around Norfolk. All the participants had at least a basic knowledge of GIS, spatial data and ArcGIS and used these on a regular basis. Table 7 shows an overview of the respondents' characteristics. Three quarters of respondents said they could play a musical instrument (Q2a) and nearly 8 out of 10 listened to music for more than 7 hours per week (Q2b). The learning style questions had a poor level of internal consistency for visual and auditory learning styles, with only 19 (38.8%) selecting both visual and only 9 (18.4%) participants selecting both auditory choices.

Organisation	Ordnance Survey	UEA	Local Authority
	40.8% (20)	44.9% (22)	14.3% (7)

Question	Yes	No
Q2a Musical Experience (Play)	75.5% (37)	24.5% (12)
Q2b Musical Exposure (> 7 hrs/wk)	79.6% (39)	20.4% (10)
Q2c Do you suffer from any sort of colour blindness?	2.0% (1)	98.0% (48)

Question	Visual	Auditory
Q3a Learning Style – Lecture	85.7% (42)	14.3% (7)
Q3b Learning Style – Photos	46.9% (23)	53.1% (26)

Question	High	Medium	Low
*Q4a How do you rate your knowledge of OS Address Data products?	20.4% (10)	38.8% (19)	40.8% (20)

Table 7. Overview of participants characteristics.

*Q4a originally asked participants to choose from five options, but these were collapsed into three groups because some groups had small numbers. *High = Know it very well; Medium = Know a reasonable amount of information about it & Know some information about it; Low = Heard of it but know little / nothing about it & Never heard of it.*

4.4.2 Basic observations

Nearly all participants (46 out of 49) identified the correct proportion of Surveyed values in at least two of the four presentation methods. Table 8 and Figure 32 shows how the mean answer for participants (based on correct = 1, not correct = 0) varied by the proportion included and presentation method, which were the two main influencing factors. VS Same and Visual methods (see Table 6 for abbreviations) performed best, while Sonic and VS Different had lower correct frequencies.

		Proportion of Values with 'Surveyed'			Total
		25%	50%	75%	
Stage	Visual	0.75	0.71	0.95	0.82
	Sonic	0.56	0.57	0.88	0.67
	VS Same	0.95	0.72	0.92	0.86
	VS Different	0.7	0.42	0.76	0.68
Total		0.74	0.62	0.88	

Table 8. The mean answers (where 0 = incorrect and 1 = correct) for each of the different stages, split by proportion of values.

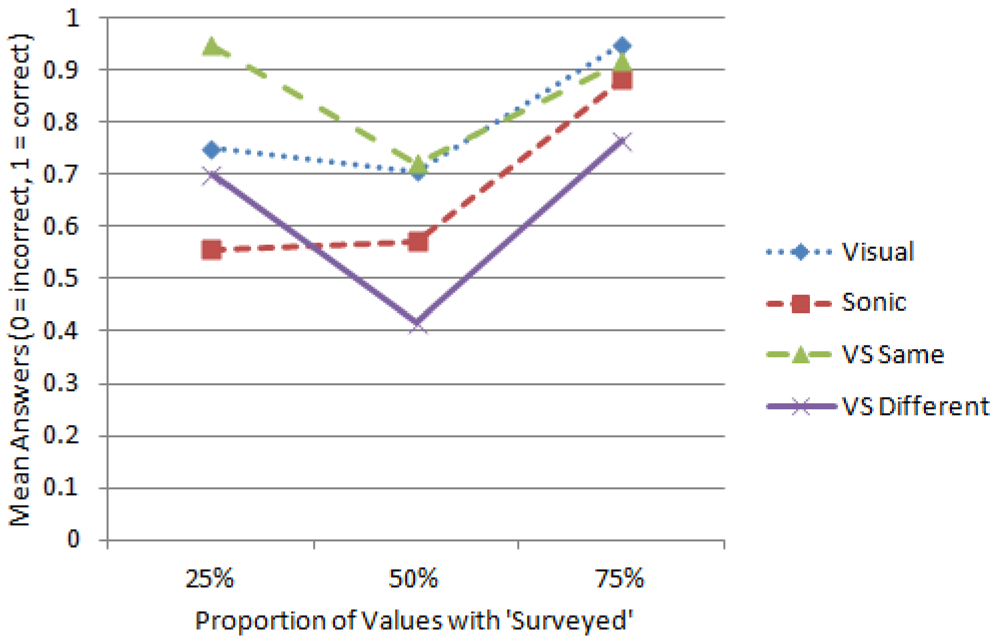


Figure 32. Variations in correct answers by proportion of Surveyed values and presentation method.

The second broad trend was for more correct answers with 25% and 75% proportions of Surveyed values and fewer with a 50% proportion. The exception to this was the Sonic only presentation method, which performed as badly with 25% as it did with 50%. This is likely to have occurred because the nature of the user interface, which made it straightforward to find particular sounds when they were common, but more difficult when they were sparse, leading participants to overestimate the proportion. Both proportion ($p < 0.005$) and presentation method ($p < 0.05$) had a significant influence on whether a participant identified the correct proportion of Surveyed values. Participants were also asked to select the easiest and hardest presentation methods to use, with 59% choosing Visual only as the easiest and 51% specifying Sonic only as the hardest, broadly consistent with the trends in Figure 32.

4.4.3 Objective measures and subjective measures

In addition to the objective measure of success (the mean answers, discussed above) participants were also asked for their subjective views on the sonification, specifically "How easy did you find it to answer the questions with this example?" For all the data a Spearman's Rank Order Correlation was run to evaluate the relationship between the two variables. Participants who found the sonification easier to use performed better ($r_S = 0.213$, $p = 0.003$). The correlation is not particularly strong and this relationship may be dependent on stage or proportion, but the small number of results limits the analysis that can be completed. To show the relationship visually, mean values for both variables was calculated for each stage and proportion combination, and are shown in Figure 33. The outlier in the graph (mean answer = 0.42) was for the VS Different stage using 50% proportion. The 50% proportion was the hardest proportion, and the VS Different stage was more complex than any of the other stages, which meant that participants could have underestimated the number of errors they made.

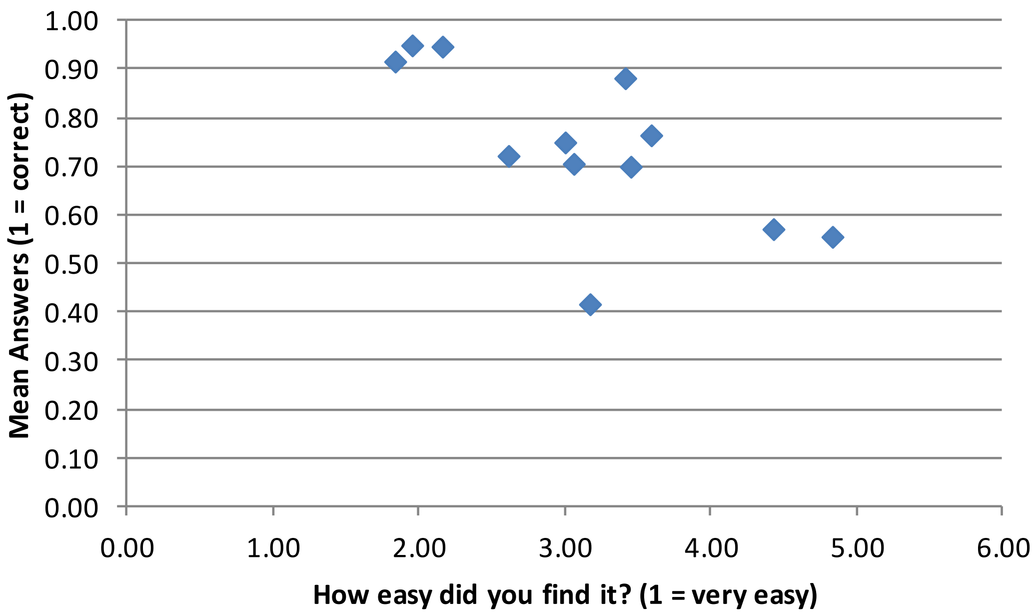


Figure 33. Objective measure (mean answers for each stage and proportion combination) against subjective measure (How easy did you find it?). A Spearman's Rank Order Correlation ($r_S = 0.213$, $p = 0.003$) shows a significant relationship between the two measures.

4.4.4 Logistic regression

A logistic regression analysis was undertaken in SPSS 16 to examine how combinations of different variables influenced the probability of a correct answer. Data for the three respondents who chose incorrect proportions for all or all but one of the methods were excluded from this analysis as they were outliers. In a forward stepwise analysis only the proportion of surveyed values and presentation method made a significant improvement to model fit at the 0.05 level (see Table 9) and the coefficients for these variables showed similar trends to those already discussed.

Factors added to Model	-2 Log Likelihood	Cox & Snell R^2
Proportion of Surveyed	182.01	0.043
Presentation Method	168.579	0.11
Address Knowledge*	167.319	0.116

Table 9. Factors added to the logistic regression model and their impact on model fit. *Address knowledge did not make a significant improvement to the model.

Within each of the survey locations (Ordnance Survey, UEA and local authorities) there appeared to be a trend that indicted higher knowledge of the AL2 data set resulted in an increased likelihood of a correct answer. However, as an overall effect in the regression model this did not provide a significant improvement (see Table 9). In part this is because including this variable resulted in some combinations of attributes with low counts (<3) so a larger sample is likely to alter this interpretation. Similarly, it should be recognised that participant specific attributes (e.g. AL2 knowledge, survey location etc.) were duplicated when the data were analysed by categories such as presentation method and it would require the use of statistical techniques such as multi-level modelling to properly separate out these aspects of the data set structure. However, this refinement is most unlikely to alter the general interpretation above. Learning style and musical experience/exposure did not have any significant influence.

4.4.5 Qualitative results

The recordings from the discussion sessions were coded and summarised. Participants said they were comfortable with the Visual method as it was familiar to them and they were also comfortable with the VS Same method because the sounds confirmed the information acquired visually which made them feel more confident about the proportion they chose. Many respondents said that the Sonic only presentation was hard to use, mainly because it was difficult to get an overall sense of the data when scanning the cursor over the map display. VS Different was also described as challenging, primarily because of the need to separate what was seen from what was heard (i.e. the two sources of information could be in conflict with each other).

The free text answers from the survey indicated that some participants found the sonification very useful and that it added considerably to their interpretation of the data, while others said the sound was very difficult to understand and when combined with vision, distracted them from the visual interpretation. There was a trend that users familiar with the AL2 data found the sonification easier to user and more effective than those who were not familiar with the data, mirroring the results found in the logistic regression. However, this trend was not significant. Sonic and VS Different methods were considered harder to use than Visual and VS Same, consistent with previous results. Sonic only had a relatively low success rate and does not appear to have much applicability with this type of interface.

The discussions after each evaluation session provided further qualitative information and gave participants a chance to suggest changes and improvements to the technique. Preferences for the types of sounds used were subjective and are likely to vary depending on the data set and the analysis taking place. A wider range of audio clips coupled with user choice could allow easier differentiation of sounds and potential for representing a larger number of variables. Possibilities include different piano notes and instruments or completely different environmental or animal sounds. If a user could use a more appropriate sound for the data set or a sound they personally preferred their performance may increase. Colour-blind users were highlighted as a group who might find the sonification useful; however a larger sample size is required to effectively evaluate this possibility as from the sample of 49, only one person was colour blind.

4.5 Conclusion

Knowledge of the data set being sonified (AL2 in this case) appears to be an important factor for effective use of the sonification, with participants with a higher knowledge performing more effectively. However, the information collected on the participants' knowledge of AL2 was limited, so more detailed information should be collected in the future. Additionally, it will be useful to collect information on their knowledge and/or experience of uncertainty in spatial data as this is likely to be an important factor when using a sonification of uncertainty data. Using sound to reinforce data shown visually does improve performance for participants in this evaluation.

Estimating the proportion of a data set may seem unusual to many GIS users, as a more common way of calculating these values would be via statistics. However, the method provided a means of comparing visual and auditory stimuli methods and is common in other fields, such as histopathology. This is in the context of looking at stained cells when testing for cancer and when using a reference card (to compare proportions to) humans perform very well (accurate to within 5%) (Cross, 1998; 2001). However, estimating proportions without a reference card (such as in this study) has a wide range of success rates, and there has been very limited research in this area, despite its relevant significance in both the cartography/GIS and histopathology fields (Cross, 2010, pers. comm.).

The task chosen may limit the wider applicability of the results but there are few existing evaluations of this type so there is little comparative data and very few examples of previous successful evaluations to base the method on. It is obviously possible for the AL2 positional accuracy to be represented visually, using some type of icon or symbol on the map, negating the need for sonification. However, the study sought to determine if sonification could work in this setting and to add to understanding of the use of sonification. The task needed to be easy enough to ensure that most participants could answer at least some questions correctly and avoid a situation where there were widespread incorrect answers. Overall, the research demonstrates how a sonification tool can be added to an industry standard GIS and the results suggest that at least some spatial data users find such a technique a useful complement to visual representation. Upon reflection, the majority of participants answered the questions correctly, so perhaps the difficultly level could have been increased. Possible future options include more complex tasks (such as clustering exercises) and other comparisons of presentation methods, utilising both sound and vision. These could include different types of sounds, such as other musical instruments, non-musical abstract sounds or potentially even real sounds depending on the data being sonified. Another possibility would be to generate sounds representing the overall data quality of the map, potentially using a 3-note chord with the volume of each note proportional to the amount of data with that quality. Such an approach would help negate the issue of the users finding it difficult to get an 'overview' of the map from sound alone, but such a sonic signal might be complex for some people to interpret. Greater repetitions of the different evaluation methods would increase the statistical validity of the results and would highlight any potential learning effect.

The decision to develop this tool as an add-in to ArcGIS was explained at the beginning of this chapter and provides an easy to use sonification tool in an environment users are already familiar with. This is appropriate for GIS users who use ArcGIS on a regular basis. However, it does limit the testing audience to those who are familiar with ArcGIS and have it installed on their computers. This was not an issue for users of the AL2 data set, but users of UKCP09 (used in CS2) cover a much wider range of GIS knowledge and ability and not all of them will have ArcGIS available. Therefore, it was decided to build a tool that was more widely available for the next case study.

When completing the evaluation, one participant who was an advanced GIS user attempted to use the zoom tools provided within ArcGIS to zoom into a specific area of the evaluation data. Unfortunately this disabled the sonification and meant the evaluation had to be restarted for this participant. While this action was a perfectly legitimate tool selection in ArcGIS, it was not anticipated that the participants would attempt this. This is one of the disadvantages of building an evaluation tool within an existing application, rather than a standalone tool just for the evaluation. However, only one participant (out of 49) had this issue.

This research has also highlighted specific characteristics that influenced the ability of users to interpret sound to make proportion judgements. The proportion of the data the user is interested in and the presentation method are the two factors that appear to have the most impact on whether a person will be able to understand the information correctly. Knowledge of the data set being sonified also appears to have some impact, but it is not so clearly apparent with the available data. These issues will be explored further in the subsequent case studies which evaluate the use of sound to represent uncertainty in UK climate scenario data and landscape visualisations.

4.6 Findings in Relation to the Research Hypotheses

This section reviews the results from this case study in terms of the original research aims and hypotheses. They will be compared with results from the other case studies in Chapter 7.

4.6.1 Specific research aims

The specific research aims are:

1. What form does uncertainty information need to be in for it to be represented using sound?
2. How effective are different means of representing uncertainty in spatial data?
3. Which types of sound variation best represent spatial data uncertainty to users?

This case study addresses specific aspects of these questions, but only sets the ground work for aims 1 and 3. The issue of ensuring uncertainty data are available and in a usable form is crucial and unless significant work is undertaken to create explicitly stated uncertainty data for a wider range of data sets, uncertainty will often be ignored because the process of creating the uncertainty data is too complex. Only one type of sound variation was used in this case study (i.e. piano notes) so the main comparison between the different types of sound and uncertainty information will be in the later case studies and between all three case studies.

This case study does enable a good comparison of the representation of uncertainty information visually and sonically, including using visual and sonic methods in combination to improve performance. The different methods were found to have different advantages and disadvantages, depending on the data used and the setting of the sonification. Using sound to reinforce uncertainty information shown visually resulted in better performance and also increased the participants' confidence with their choice. Using sound alone as a method of representing spatial data was worse than any other method of representing only the uncertainty information ($p < 0.069$), but performance did vary with the proportion of the data.

4.6.2 Research hypotheses

This section outlines the research hypotheses in reference to the results from this case study, both in terms of the sonification design and evaluation.

H1. The way uncertainty is conceptualised within the spatial data set will impact how it can be represented in a sonification.

The study indicates that both the conceptualisation of the uncertainty by the data producers and its inclusion (or not) within the data set have a major impact on how the data can be sonified. Change detection and generalisation were not possible options for sonification of uncertainty data because of the work required to get the uncertainty information into a usable form. The AL2 data used in this case study specifically investigated the positional accuracy element of the uncertainty because this was the element that was used most frequently in the real world and was already encoded in the data. From a commercial address matching point of view, some of the uncertainty information (such as whether the address is a commercial or residential property) was not clear and other information (such as whether the address is occupied) was not included.

H2. The technology available to design, implement and evaluate the sonification of uncertainty in spatial data will impact the findings.

The technology used to implement the evaluation (ArcGIS and VBA in this case study) had an impact on the sonification that could be implemented, what information could be recorded in the evaluation, how it could be recorded and where the evaluation could take place. For the local authority participants, it was not known how the evaluation would run on a computer in their office (even if it had the same version of ArcGIS installed) so tested laptops were taken to their office for their evaluations. There was a major programming commitment required for this case study and had the author not already been familiar with the programming language and environment, development would have taken much longer.

H3. The ability of the user to make effective use of the sonification will be effected by a variety of factors including their knowledge and previous experience.

Experience of AL2 data was a factor in the ability to use the sonification, but it was not a significantly strong influence to be statistically significant. This is likely to be because the familiarity with the status flag element of the AL2 data was more relevant to the evaluation task than general familiarity with the AL2 data. Previous experience of GIS and computer use in general was not recorded for this evaluation, but should be included in future case studies. General experience of products used for address matching may also be important. Overall is was recognised that participant experience (of computer use, GIS, the data set and uncertainty within that data set) needed to be recorded in a more detailed manner in the subsequent case studies. Learning style was not a factor in this case study, probably due to the way the question was asked. A more in-depth question was thought likely to provide more information and therefore included in the later case studies. Musical ability and musical exposure did not have any influence on the results.

H4. Providing redundancy (dual-coding) in a task by showing the same information both visually and sonically will improve performance.

Showing uncertainty information redundantly, using both visually and sonically (VS Same) did indeed result in better performance, but there was not a significant difference for all the proportions of Surveyed values. This is likely to be because of the limited complexity of the task, and a more complex task might have generated a larger difference.

H5. Transferring the representation of uncertainty information from the visual to the auditory domain to address visual saturation will reduce cognitive load and therefore improve performance.

There was not an issue of visual saturation with this data, so no direct comparison was completed between a visual data & visual uncertainty representation and a visual data & sonic uncertainty representation. Using sound to show uncertainty and vision to show an additional variable (VS Different) was tested but performance was less good, due to the increased complexity of the exercise compared to previous participant tasks.

H6. Preference for one sound from an option of two will result in better performance with the preferred sound than performance with the other sound.

There was no choice in the sound used for this evaluation, but preference for an option of different sounds was noted in the discussion sessions. The provision of such a choice was therefore noted as an option for future case studies. Participant perception did appear to have an impact, with participants who saw the sonification as easy to use performing more effectively.

This case study fully addresses hypothesis 1, 2 and 4, partially addresses hypothesis 3 and 6 and does not address hypothesis 5. Several aspects of data collection and issues requiring further investigation were identified for further consideration for improvement in subsequent case studies. The results from this case study will be compared with those from the other case studies in Chapter 7.

Chapter 5. Case Study 2: Using Sound to Represent Uncertainty in UKCP09

5.1 Introduction

The main aim of this case study is to evaluate whether sound can effectively be used to represent uncertainty in future climate data. Specifically, this research uses uncertainty information available in the UKCP09 (UK Climate Projections 2009) data set to compare different visual and sonic methods of representing both the uncertainty within the data and the data itself. This evaluation incorporates findings from the previous case study and develops several of the issues highlighted.

This case study evaluates the research aims and hypotheses identified in the literature review (Chapter 2), and summarises the findings obtained at the end of the chapter. In addition it will also consider these specific research questions:

- Can sound be used to represent uncertainty in the UKCP09 data set?
- Is using sound to do this more effective than using existing visual methods?
- How does the participants' learning style and/or previous knowledge of the data set impact on their ability to use the sonification?
- Is a web based evaluation more or less useful than an evaluation based in ArcGIS (as used in CS1)?
- Do the Ordnance Survey participants perform better or worse with UKCP09 data compared to AL2 data (CS1)?

Participants from UEA, Ordnance Survey (OS) and the UK Climate Impacts Programme (UKCIP) were recruited to take part in the evaluation. They completed a quantitative and qualitative evaluation of the different representation methodologies, using the UKCP09 data as the example data set. UKCP09 is the latest in the series of projected future climate variables for the UK, covering a number of different climate variables. This data set was chosen because it is the first set of future climate data that explicitly contains information on the uncertainty of the projections and the users of these data are often not familiar with this level of information so an effective way of displaying it is required. There is some discussion in the literature about how effective and/or useful the UKCP09 data set is (see Section 5.2.1.6, page 140), but this was not explicitly discussed in the evaluation and did not impact the use of the data in this case study. The case study also uses a web based questionnaire and the Google Maps API to capture the users responses, rather than ArcGIS as in the previous case study. The Google Maps API is a free to use

Application Programming Interface for the popular web based mapping service provided by Google. There are various issues surrounding the use of technology such as the Google Map API and reproducible research, which will be discussed later in the chapter.

The chapter is split up into five sections, starting with this introduction, followed by the literature review. This discusses the relevant literature to this case study, focusing mainly on the UKCP09 data set and the Google Maps API methodology. Details on the sonification aspects can be found in Chapter 2. The next section explains the methodology used, including the different methods for presenting the data to the user, the Google Maps API, the questionnaire and the data presented to the participants. The results section discusses the method for calculating the Phi comparison coefficient and the other results, as well as their implications for this study and the wider thesis. The final section concludes the chapter and sets the results in context with the rest of the thesis.

5.2 Literature Review

The section reviews the literature and current thinking on the two main topics relevant to this chapter. Firstly, the UKCP09 data set will be considered, with details on its history, current use and some of the issues created by the inclusion of uncertainty data in the latest version of the projections. Secondly, the use of the web based survey and the Google Maps API for conducting the evaluation will be justified. The discussion will touch upon the sonification aspects relevant to this case study, but the main review of this literature is in Chapter 2.

5.2.1 UK Climate Projections 2009 (UKCP09)

UKCP09 is the latest in the series of future climate projections for the UK (Jenkins et al., 2009), combining the latest in scientific knowledge about the possible future changes in climate and user training on adaptation and mitigation. It consists of temperature, precipitation, air pressure and a number of other climatic variables for seven 30 year overlapping time periods from 2010 to 2099. It is the 5th in the series of climate projections, with each developing and becoming more complex than the last. UKCP09 data were chosen for this case study because they contrast with the AL2 data set used in the first case study. The AL2 data set consisted of discrete objects and UKCP09 is a continuous surface data set.

There are a number of reviews explaining UKCP09 (Murphy et al., 2009) and the science behind it (Sexton et al., 2011; Sexton and Murphy, 2011), as well as the previous four iterations of future climate data for the UK (UKCIP, 2010d; Dessai and Hulme, 2008). The spatial resolution, number of emissions scenarios included and temporal resolution have generally increased with each iteration along with the amount of user engagement in their development and delivery (see Table

10 for a more detailed summary of the changes). UKCP09 continues to push these boundaries, but additionally it includes a range of potential values (and associated probabilities for each of these values) for each parameter, location, emissions scenario and time period combination rather than the previous data sets, which only provided a single value. This is a significant change, both in terms of the amount of data that are available to the users as well as a change in how the users can use the data.

	CCIRG91	**CCIRG96**	**UKCIP98**	**UKCIP02**	**UKCP09**
Factors	1.5% p.a. growth in GHG concentration	IPCC IS92a emissions scenario	0.5 and 1.0% p.a. growth in GHG concentration	IPCC SRES scenarios: B1, B2, A2, A1FI	IPCC SRES scenarios: B1, A1B and A1FI
Spatial Resolution	5° by 5° grid (ca 500 km)	2.5° by 3.75° grid (ca 300 km)	2.5° by 3.75° grid (ca 300 km)	50km by 50km grid (ca 0.44°)	25km by 25km grid (ca 0.22°)[9]
Temporal Resolution	Seasonal averages	Seasonal averages	Monthly/ seasonal averages, plus some daily weather variables	Monthly/ seasonal averages and inter-annual variability, plus some daily weather variables	Annual/ seasonal/ monthly/ averages, plus daily data available from the Weather Generator

Table 10. Summary of the changes in the UK climate projections leading up to UKCP09 (from Jenkins et al., 2009; Dessai and Hulme, 2008).

The set of projections is widely used by many different people across the UK. Many users of the projections (such as local authorities and area specialists including water companies and structural engineers) use the data to provide inputs into their own models about how their facilities or infrastructure will be affected by a changing climate. It is also used by other governmental and non-governmental organisations (e.g. DEFRA, English Nature, Forestry Commission) to understand how climate change will impact their areas of interest.

Some of these users have been using the climate projection data through a number of its previous iterations and all of these have consisted of a specific figure for the projected climate variable, for a particular emissions scenario and time period for a specific location. For example, using UKCIP02, the high emissions scenario projected summer mean daily temperature increase for Dorset (50km x 50km grid square 434) for the 2080s is 5.6° C (UKCIP 2010c). The underlying climate models used to calculate these values had uncertainties within them and scientific judgement was used to reduce the uncertainties as much as possible and to provide the user with one value.

[9] The spatial resolution is for the probabilistic projections element of UKCP09, the main element. Other elements (e.g. Weather Generator or marine data) have different spatial resolutions.

5.2.1.1 Innovations in UKCP09

One of the innovations introduced with the UKCP09 data set was the inclusion of uncertainty in the climate projections as a result of specific requests from the user community (Steynor et al., 2012). More specifically, a series of projections has been provided for each climate variable, emission scenario, time and location combination, with information on the likelihood of each possible outcome according to current science. This is very different to the single value provided by the UKCIP02 data set and the UKCP09 data has to be used in a very different way. This has also required a different scientific way of generating the data. Traditionally, scenarios are used to represent different possible futures depending on the way the world develops as a whole (IPCC, 2000). UKCIP uses an approach known as perturbed physics ensemble (PPE), which takes uncertainty into account in a different way (Gosling et al., 2012). PPE is seen as a more effective way of representing uncertainty and allows for the presentation of more information than using scenarios (Collins et al., 2006). There are numerous advantages to providing a series of projections for each location, as the additional information available is useful to many users. It allows the users to choose their own risk level and therefore how much money they may need to spend on defences to mitigate a particular risk. For example, a 90th Percentile event is less likely to happen than a 66th Percentile event, but preparing for a 90th Percentile event would cost more than preparing for a 66th Percentile event. It also allows the users to make statements like 'it is very likely that the temperature increase for Dorset under a high emissions scenario in 2080s will be between 2.8°C and 8.4°C', which they would not have been able to make using the UKCIP02 data (UKCIP 2010c). The terms 'very likely' and 'likely' have been adopted in UKCP09 in a similar manner to use in the IPCC reports (see Table 11 for details). This gives the users of the data significantly more information about the possible range of temperatures as well as an indication of the uncertainty surrounding them.

Probability Level	Qualitative Term
10%	Very likely to be greater than/very unlikely to be less than
33%	Likely to be greater than/unlikely to be less than
50%	Central Estimate
66%	Likely to be less than/unlikely to be greater than
90%	Very likely to be less than/very unlikely to be greater than

Table 11. Probability levels and associated qualitative terms used by UKCP09 (UKCIP 2011c). These are broadly similar to the terms used by the IPCC.

However, the significantly complex format of the data requires training to enable the users to make most effective use of the data. A number of training sessions were conducted and a range of training materials made available by UKCIP to help the implementation of this approach these were generally received successfully (see Section 5.2.1.3 for more details). In addition, the data have been made available at a number of different levels, ranging from general headlines, through region specific pre-made maps to data available for download for the relevant variables and areas the user is interested in, which allows the user to pick the level of complexity of the data they wish to use.

5.2.1.2 Data available as a part of UKCP09

The UKCP09 data set contains a vast amount of information and a user interface has been created to allow users of the data to access it in a simple and effective way, with contextual help to assist the users throughout the process. A description of the whole data set is well beyond this thesis, but there are two main aspects of the data worth exploring – the climatic variables available (and various sub-choices possible) and the output options available to the user. The most widely used set of data is the land based probabilistic projections, which this research uses, but there are also marine related projections, which contain information relevant to the projected changes over the seas surrounding the UK (Lowe et al., 2009). The probabilistic projections have an associated Weather Generator which allows example 'weather' to be generated, which would be representative of a particular time period in the future (Jones et al., 2009; UKCIP, 2011a). This research uses the probabilistic land based projections, so the marine projections and Weather Generator data will not be discussed further.

The main climate variables available are temperature (mean, daily max and min and temperature of the warmest/coldest day and night), precipitation (mean and wettest day), air pressure, cloud cover and humidity (relative and specific). This study used the mean daily summer temperature (mean of maximum daily temperature of June, July and August) and temperature of the warmest day. The temperature of the warmest day was only available as a change (from 1961-90 baseline) so the absolute value was calculated using the baseline data available from the Met Office (see Section 5.3.5.2). The temperature data were used (in preference to precipitation data) because the uncertainty surrounding precipitation is much higher than temperature. Only one season was used to reduce the number of combinations of data in the evaluation, and therefore the time required to complete the evaluation.

For each of these variables, the time period, temporal averages, emissions scenario and spatial location can be selected. The time periods are a series of 30 year periods, centred around seven decades of the 21st century (starting with 2020s which is 2010-2039, followed by 2030s, which is 2020-2049 and so on all the way up to 2080s, which is 2070-2099). The temporal averages relate to either an annual average, a seasonal average (winter, spring, summer, autumn) or monthly averages (Jan, Feb, etc.). There are three emissions scenarios to choose from (low, medium and high), which relate to the SRES emissions scenarios (B1, A1B and A1FI respectively) (IPCC, 2000). The medium emissions scenario was used for this research to act as an example, and further discussion of the differences between the emissions scenarios and reasons for these is beyond the scope of this thesis. The final option was the spatial location, which could be one (or more) of the 25km x 25km grid cells covering the UK, one (or more) of the Administrative Regions of the UK (UKCIP 2010b) or one (or more) of the River Basin areas broadly based on the Water Framework Directive River Basin Districts (UKCIP 2010e).

The data can be exported in a number of different formats, depending on its intended use and the options chosen in previous sections. The most useful formats are maps (for presentations) or raw data either at given probability levels or above or below a particular threshold. The raw data are often used as an input to further models, for example rainfall models or crop models. Two probability plots are available, the cumulative distribution function (CDF) and the probability density function (PDF) plots, which provide the full range of probability information. They provide the same information, but represented in different ways (as shown in Figure 34). The CDF data were accessed using the raw data option rather than the sampled data option, which does provide more comprehensive information, but was not available at the time the data were acquired.

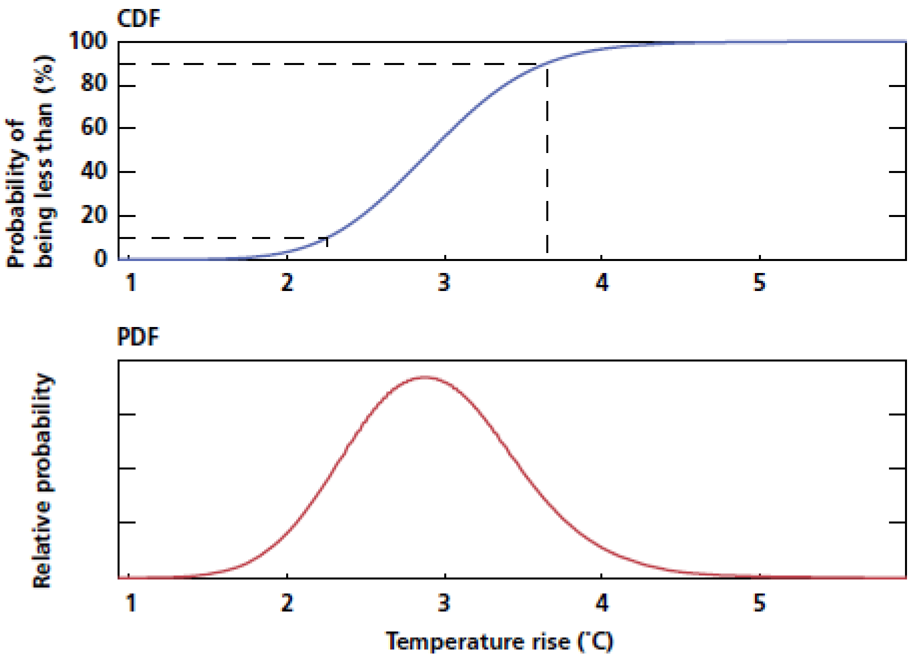


Figure 34. PDFs and CDFs show the same information but in a different form. A PDF can be integrated to form a CDF, and a CDF is a cumulative PDF. A hypothetical example is shown above to show how they relate to each other (Jenkins et al., 2009, fig.8).

The UKCP09 data are available as shape files for use within a GIS, but they are provided in a rotated pole projection, which caused some issues when using the data and analysing the results. The use of a rotated pole grid is the norm in climate models, because it is a much more efficient way of storing world wide data than a standard lat/long grid, as it has fewer grid cells in the polar areas and more in the lower latitudes, where the data are more relevant. The UKCP09 data are provided in this projection to end users and can be transformed to a different projection if required, but this could introduce errors to the data and the rotated pole projection is not an issue for most users.

5.2.1.3 Training on UKCP09

UKCIP (UK Climate Impacts Programme) was established in 1997 to help coordinate scientific research into climate change and its potential impact on the UK (UKCIP, 2010a). Other countries in the world are beginning to start similar programmes, but the UK's was the first. Part of their remit is to help all users make best use of the future climate data currently available. As previously discussed, UKCP09 marked a significant change in the type and complexity of the data being given to the user (Steynor et al., 2012) and UKCIP ran a series of pre-launch training workshops to promote awareness of the data set and help users understand the data that were available to them (UKCIP, 2011b). In the year following the launch of UKCP09, UKCIP ran a series of 51 post-launch free training sessions called PiP (Projections in Practice) around the country for users to receive information and training on using the UKCP09 data. This included training on the probabilistic nature of the data and how to use it effectively, as well as information on how to migrate from the UKCIP02 data set. The training also consisted of a practical element, which was a series of exercises showing how to use the data appropriately. The training was very well received by the users; however the funding for this intense series of training sessions was only for one year. Since then UKCIP have run occasional, charged for (to cover costs) sessions and they provide free online resources and webinars. There is a large amount of user guidance on the website, as well as a helpdesk service, answering users' queries if the information on the website is not sufficient. Small (2010) conducted a review of the UKCP09 data set, and commented that the training material was very well received, but some of it was too technical for the end users, and simpler information is required. There is also a gap between what the users would ideally like (a personalised report, taking into account their attitude to risk for a particular application) and what UKCP09 can provide (the ability to create such a report, but requiring time and input from users in order to do so). UKCIP has released a new tool called AdaptME[10] which helps users assess their own attitude to risk using a questionnaire approach. It is designed to work as a practical support to help users think about their current adaptation processes and attitude to risk and help use the UKCP09 data to modify these.

Provision of climate change adaptation information is undergoing a series of alterations as a result of funding cuts and changes to central government (as of July 2011). Part of the UKCIP function has been subsumed into the Environment Agency (which has subsequently sub-contracted a

[10] Available at http://www.ukcip.org.uk/adaptme-toolkit

significant part of the work), but it is planned that part of the research element will stay with the Environmental Change Institute at University of Oxford, where UKCIP is currently based (UKCIP, 2011e). Under the Climate Change Act 2008, provision was made for a National Climate Change Risk Assessment every 5 years, of which the basis was formed by UKCP09. While the implementation of this work may be moved around different departments, it will always be fulfilled to comply with the legal requirements.

5.2.1.4 Displaying the Uncertainty in the Data

While there is a reasonable amount of support for understanding and using the ranges of the probabilistic data that are provided, it is considerably more difficult to display the latest set of data compared to the previous ones, particularly on a national scale. While planners for particular buildings or over small areas can show and talk about the ranges reasonably easily, it is more difficult to consider a national area and show the ranges effectively. One suggested way is to show three maps, with the 10th and 90th Percentile data to show the very likely range of the climate variable, and the 50th Percentile to show the central estimate (the median) (UKCIP, 2010f). While this does communicate the information, it is quite difficult to get a good idea of the range, and may be difficult for viewers of the map to combine the three into something more easily understood. An example of this is shown in Figure 35, taken from the UK Climate Projections Science Report (Murphy et al., 2009).

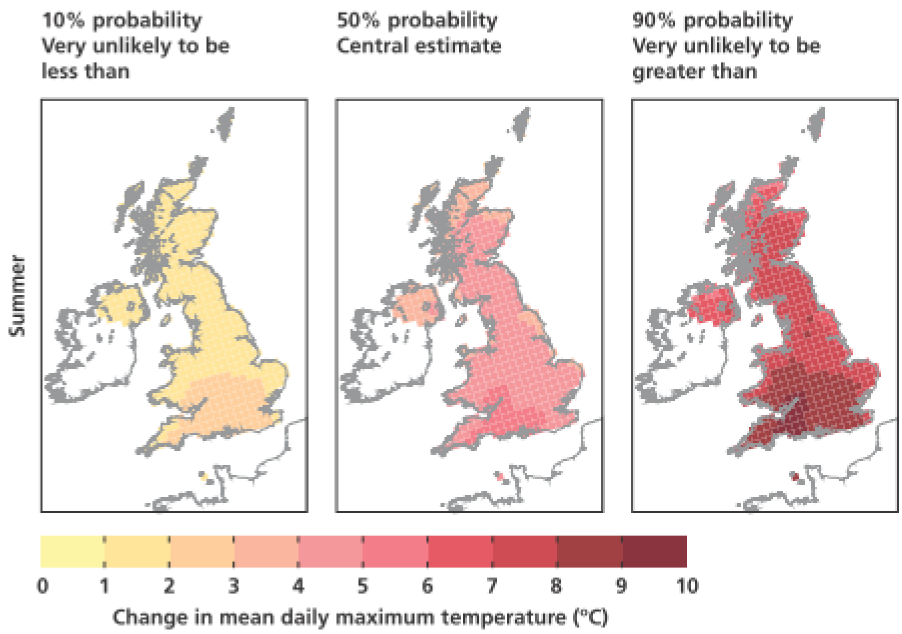


Figure 35. Three maps showing the range of projections within the UKCP09 data for summer 2080s (Murphy et al., 2009, fig.4.4).

Figure 35 shows the projected change in summer (JJA) mean daily maximum temperature for the 2080s under the medium emissions scenario over the UK. The increase in the summer mean daily maximum temperature in the central south west is very unlikely to be below 2°C or above 10°C. The increase in the south east is very unlikely to be below 2°C or above 9°C. The increase in Scotland is very unlikely to be below 1°C or above 7°C.

While the three maps method as described above shows the data, it does have a number of disadvantages, including the issue of the user having to bring the three data sets together to make sense of all of the information. Instead of the three map method, it is possible to show the 50th Percentile data visually, as you would on a normal map, but have the range shown using sound, with different sounds reflecting different ranges. This case study will compare different methods of using sound and vision to show the range data to see whether sound could work in this context.

Another visualisation tool (GRaBS, 2011) is available that uses UKCP09 data as a part of its data source using a Google Maps API interface, but no attempt is made to display the uncertainty of the data as well.

5.2.1.5 Use of UKCP09 Data in this Case Study

Of the various climate variables available for use within UKCP09, it was decided to use some of the temperature variables for this case study because the uncertainty surrounding temperature is much lower than the uncertainty surrounding precipitation. The reason for this is because the historical recordings of temperature have a lower variation than those of precipitation, and consequently a lower coefficient of variation. The lower coefficient of variation means that temperature is easier to predict than precipitation; i.e. there is less uncertainty in predictions of temperature than precipitation. Uncertainty is a new piece of information that is being communicated to the data users, so it was decided to keep it relatively simple to avoid overwhelming the users with lots of information. Both summer mean temperature and temperature of the warmest day variables were used to provide different, but related data sets to allow some randomisation in the evaluation. The evaluation used three different time periods of data; the baseline (1961-1990), 2020s (2010-2039) and 2050s (2040-2069), which gave a set of data that varied in magnitude, but had broadly similar spatial patterns over the time periods.

The UKCP09 data are presented as a series of projections with probability information on the likelihood of each of the projections. The 50th Percentile value of this data was used to represent the projected temperature increase. It is not strictly speaking the mean (it is the median value), but acts as a marker for the potential variation in temperature to be described. The values are provided at all other Percentile levels (1st - 99th) and the values for the 90th and 10th Percentiles are used to generate the 'very likely range'. For example, the temperature for a specific location in a specific time period under a specific emission scenario is very likely to change by between 2.8°C and 8.4°C. The 10th and 90th Percentile values were subtracted to generate a range value, which represented the uncertainty for this evaluation (see Figure 36 for an example of the Range data). Projections with a large range are relatively more uncertain than projections with a small range. The range is used as a measure of the uncertainty of the data, with a larger range representing a larger uncertainty and a smaller range a smaller uncertainty.

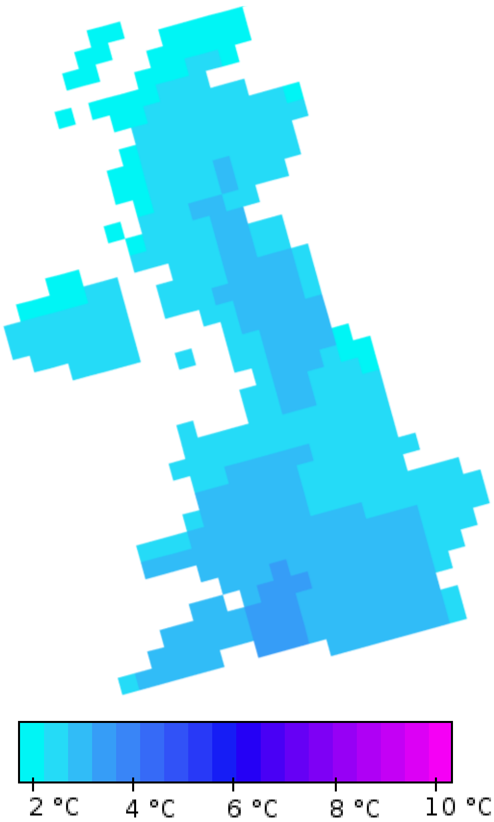


Figure 36. Example of the Range data for Summer Mean 2020s data.

5.2.1.6 Criticisms of UKCP09

While the UKCP09 climate projections do represent a significant change in the type of data provided in this area, they have been criticised by some people. The papers mentioned below are, no doubt, only an example of some of the critical writing on the UKCP09 data, but they are representative of the issues being discussed. The author’s background and speciality is not in the area of predictive climate data, so these issues will not be discussed in depth, and they did not have any impact on the evaluations conducted.

The methodologies for generating the UKCP09 data sets had not been published in peer reviewed journals when the dataset was launched (June 2009), as is the requirement for IPCC and similar reports. This was because the peer review process would have taken a significant amount of time, and by the time the data could have been released, there would be further developments in climate projections, which would need to be included. To address the concern highlighted by some that UKCP09 was not independently peer-reviewed, an independent science review of UKCP09 was conducted (Hoskins et al., 2009). A number of suggestions were made to the UKCP09 project management group, most of which were implemented (UKCIP, 2011). Two journal articles, covering the development and use of the UKCP09 methodology have been published subsequently in October 2011 (Sexton and Murphy, 2011; Sexton et al., 2011).

Hall (2007) comments on the fact that within our understanding of the climate, there are significant gaps in the knowledge of how a number of climate processes function. These are assumed to take the form of a normal distribution, even though there is no evidence that this is the case. This is partly due to a desire from the scientists involved to have some approximation for all the variables and partly due to the users of the data being uncomfortable with the concept of having incomplete information for whichever application they are applying the data to. This paper was written before the UKCP09 data was produced, which has tried to address some of these issues.

Beven (2011) wrote after the UKCP09 data have been published, and he was generally complementary about it, saying that it is taking the correct approach by including the uncertainty information in the data provided to the end users. What he was critical of is the underlying GCM (general circulation model) and how representative the GCM is of what may happen to the climate. He thinks the nonlinear aspect of climate is not adequately represented, given that previous significant shifts in climate have taken place, resulting in dramatic changes in the conditions on earth over relatively short time periods.

These pieces are an interesting comment on the UKCP09 data, and give some of the context in which it is placed in the climate science discipline. These issues were not explicitly mentioned in this evaluation, and the UKCP09 data set was simply used as an example to test the representation methodology, rather than to start discussion of climate change or adaptation strategies *per se*. It is relatively early for deep, in-depth reviews of UKCP09, as it has only been available for public use since 18 June 2009 (25 months at the time of writing), but no doubt more of these will be published in time.

5.2.2 Sonification and climate data

The UKCP09 data are a major step forward in the provision of uncertainty within future climate data. This thesis uses future climate data as one of three examples used with sonification, which does limit the theoretical exploration that can be completed on the use of future climate data and sonification. Equally, sonification as an area is one that has developed gradually within other disciplines, with long term development, particularly theoretical development, being rather patchy. A new 3 year project called SYSSON has the explicit aim of developing sonification techniques and tools for and with climate scientists (Vogt et al., 2012). They do aim to generate a number of sonification tools towards the end of the project, but the majority of the first half of the project focuses on the climate scientists themselves and works with them to discuss and explain how sonification may be able to benefit their area. This approach aims to create a foundation for sonification applications that means they will be useful and relevant for the target audience, something that this area often lacks. The project has only recently started, and it will be interesting to see what findings and outputs are achieved.

5.2.3 Collecting spatial data responses

To measure the participants understanding of the different data representation methods, an exercise was developed where the participant was asked to highlight a specific area of the data shown to them. For example, participants could be asked to highlight the area where the temperature value exceeds 4.5° C. There are a number of different ways this spatial area could be collected, but all methods would have involved selecting the area somehow and this selection method needs to be easy to use. As the target audience included non-GIS users, the interface needed to be reasonably simple and straight forward to use.

The previous case study used ArcGIS to show the spatial data to the participants, and asked them to judge the proportion of data that had a specific value. Allowing the participant to highlight a specific area in ArcGIS would not be simple to do and because of the nature of the program it would be a technical process and potentially difficult for non-GIS users to understand. Also, the ArcGIS software was required to complete the evaluation for CS1, and this would not be available for all of the target audience. For these reasons, it was decided to use a web based interface to develop the evaluation software.

5.2.3.1 Online surveys

Internet surveys are well established as a way of gathering data for academic research in many different fields (such as environmental (Wheaton et al., 2006), geographic (Haklay, 2002) and climate change behaviour (Attari et al., 2010)). Significant attention has also been given to their advantages and disadvantages as a method of collecting data (e.g. Madge and O'Connor, 2002; Berry et al., 2011; Roth, 2006). Primary benefits include wide (potentially global reach) flexibility of design and ease of data processing (Evans and Mathur, 2005). However, representativeness of respondents can be a significant weakness (Peng, 2001) depending on the target audience desired, technical barriers may impede significant sections of the target audience (Evans and Mathur, 2005), different computer configurations may render the survey differently (Wherrett, 1999; Wright, 2005; Rivara et al., 2011) and many of the issues associated with traditional research methods still apply in the virtual arena (Madge and O'Connor, 2002). This thesis will not discuss the strengths and weaknesses of online surveys in-depth unless they are of particular relevance to this setting. One advantage to using online surveys is to reach difficult to access audiences; in this case this includes users of UKCP09 data who are very geographically dispersed.

Online surveys of the nature described above are often limited in the complexity of the questions that can be included, usually to multiple choice and free text answers. When spatial data input is required, these types of questions are often used to choose the nearest town/city, or to type in the user's postcode. More complex, complete or precise spatial information often of interest to geographers can be very difficult to elicit from users. Questions such as 'which route did you use to explore the beach?' (Coombes and Jones, 2010) or 'which area would you define as your neighbourhood?' (Minnery et al., 2009) are very pertinent examples. Asking respondents to draw on paper maps is one solution (Coombes and Jones, 2010), but data processing may be time-consuming and error-prone. Therefore an online method of collecting spatial data is of use in this situation.

5.2.3.2 Online mapping APIs

There are many different online mapping services that could be used to capture spatial data in a questionnaire. The vast majority of online mapping services provide an API (Application Programming Interface), which allows third party websites to include their mapping services and, to varying degrees, customise the interface the user is shown. This started with the concept of 'Google Maps Hacks' or 'mashups' and was then officially rolled out. See Miller (2006) and Gibson and Erle (2006) for an overview of Google Maps and the development of hacks and mashups. A number of different APIs were considered for this project, including Google Maps, Bing Maps and Yahoo! Maps. OpenLayers is an open source equivalent of these commercial mapping APIs, but it is not well known outside the core "open geo" community. It also is reasonably technical to use and would have taken much longer to learn and understand than the commercial APIs, so it was not considered for this case study. Another option that could be used is ArcGIS Server to provide the mapping back-end to the system, but this is not available at UEA because of expense and it would have been more technical to implement.

Out of the commercial mapping services, Google Maps is very much the dominant offering, with 17% of the whole travel market share (including things like Expedia and different airlines as well as other mapping sites) (Hitwise, 2010). This is probably partly due to Google's seemingly unique balance between minimalism and usability, which they first applied to the search industry (Google, 2011b). When this was transferred to the online mapping area, it took off and is now one of the (if not the) clearest exclusive online mapping service (O'Berine, 2010). This does not include mapping such as that from the Ordnance Survey, which is much more detailed, but possibly less appropriate for this application. Due to this dominance it was thought that the majority of users would be familiar with the Google Maps interface and Google Maps API (GMAPI) was chosen to implement this evaluation. They also had very good tutorials and help files showing how to implement the API and a very good user support group to provide information that was not in the tutorials. Additionally, neither commercial tools such as ArcGIS Server (ESRI, 2011), nor the technical resources to run an open-source solution such as OpenLayers (2011) were available for this study.

5.2.3.1 Use of Google Maps API

The Google Maps API (GMAPI) has been used for a number of different mapping tasks by a wide range of users, including local government (Brent Council, 2007; Westminster City Council, 2010) and user generated mashups (Clarke, 2010; MapTube, 2010). It has been said previously that this is often used by computer experts rather than cartography experts (Miller, 2006), but it would appear that this is beginning to change as the GMAPI becomes more accessible and less technical to use.

From the point of view of what the GMAPI is used for, the uses can be split into three categories. Firstly, a relatively basic map showing the user where certain objects are. For example, many local councils use some mapping interface to show their residents where local services are, such as libraries. Secondly some more advanced maps ask users for input into the map and then to represent this data to other users through the existing interface. Thirdly, some maps in the academic area are beginning to use GMAPI as a method for collecting spatial data for later analysis in a GIS. This third category is new and there has been little work in this area, possibly because of some of the difficulties of getting spatial data into and out of GMAPI.

The use of GMAPI (and other mapping APIs) is growing, and will probably continue to grow, both within the academic community, and the wider community. It is important that the methods used with these emerging technologies are well documented, which the next section aims to fulfil.

The GMAPI was used in this case study to show both the UKCP09 data set to the users (both visually and sonically) and to capture their responses. The UKCP09 data was overlaid on top of the Google Maps interface using the data in KML (Keyhole Markup Language) format (OGC, 2010) to show the data visually. The sound element was added by joining the API to a sound add-on and hard coding the data in to the evaluation itself. This is obviously not ideal, but was the best compromise available, because of limitations in the KML class of the Google Maps API (for more details, see Appendix C.2). The participants' responses were extracted from the MySQL database and processed using a GIS at the analysis stage, as discussed in Section 5.4.

5.2.4 Summary

The literature has been reviewed for the UKCP09 data set and the Google Maps API. UKCP09 is the data set used as the example in CS2 and complements the other data sets chosen for the thesis as it is a continuous data set, where as CS1 used a non-continuous data set. The data also have a particular target audience who were involved with the evaluation, as well as more general UEA researchers and staff. The probabilistic nature of this data is new to the user audience and for them to understand and make best use of this an effective representation method is required. The UKCP09 data set has some critics, but this has not impacted its use in this evaluation.

A number of different visual and sonic methods were used to show the data to the users, and the effectiveness of these was evaluated by asking participants to highlight specific areas of the data. Recording spatial areas in a survey is generally limited to a text based response (e.g. postcode) but the use of online mapping APIs allows maps to be included in a questionnaire and a spatially based response to be given. The GMAPI was used to provide this interface to show the UKCP09 data and to collect responses.

The next section will explain the methodology used, covering the choice of sound, the pilot study, the questionnaire, the development of the evaluation and the running of the sessions.

5.3 Methodology

This section discusses the development and implementation of the evaluation, which compared visual and sonic methods of representing uncertainty, building on the themes and areas discussed in the methodology chapter of this thesis. It describes the participants who took part in the sessions, the evaluation itself, some of the background to its development and how results from the first case study were used to help design the evaluation. The evaluation itself consisted of three parts which will be explained in turn. The first part was a computer based questionnaire to gather basic data using multiple choice questions, followed by the different representation methodologies, where the participants were shown six different maps with uncertainty shown using vision and/or sound. The final part was a group discussion where the participants talked about and discussed the sonified maps, explained what worked and did not work for them and suggested future improvements. A number of pilot studies were run at UEA before the main evaluation, to test the procedure and interface. This generated a significant amount of useful information, and resulted in several changes to the evaluation, as discussed in the following section.

5.3.1 Use of sound

A detailed literature review on the use of sound with spatial data in situations such as this can be found in Chapter 2. The selection of sound for this case study was based on information gathered during the first case study, as well as previous research (Bearman, 2006, 2008). One of the main pieces of feedback from CS1 was that people would like the sound to sustain (i.e. be a continuous note that varied as the mouse moved over the different values) rather than resounding each time a new value was encountered. The ideal aim was to have a sound that changed seamlessly from one note to another, but that was not possible in the interface used. Instead, the note was only resounded if a different one was required to the one currently playing.

The piano does not provide a continuous note in its standard form, so it was decided to use a musical instrument that could provide a continuous note, such as a wind instrument. The UKCP09 data have positive spatial autocorrelation, with the value at a particular location being similar to values near that location. The data varies smoothly over different areas, with few sharp changes. The sounds used to represent this information should also vary continuously and this is why a continuous note was used. While the implementation meant that it was not possible to vary the note on a continuous basis, it would sound more natural with a wind instrument than a piano. In addition, one of the aims was to compare different instruments to see how they compared. To this end, it was decided to use a set of trumpet notes. The trumpet notes were continuous and were recorded from the same MIDI application as previously used (Oliveira, 2008). The notes used were from a type of pentatonic scale (C, D, E, G and A) repeated over each octave to give a total of 18 notes. The pentatonic scale was used for this case study because it was one of the preferred scales in previous research (Bearman, 2008).

5.3.2 Session participants

The evaluation required a larger number of participants than previously, because in CS1 the numbers of participants in each logistic regression group became very small ($n < 4$) which restricted the analysis that could be conducted. A larger number of participants would prevent these groups from becoming too small again. Three different organisations were approached to take part in this study. These were UK Climate Impacts Programme (UKCIP), University of East Anglia (UEA) and Ordnance Survey (OS).

UKCIP are the organisation in charge of distributing the UKCP09 data set to interested parties through the UK and they are also responsible for running the training sessions for users of the data. In the evaluation they formed two sub-groups; firstly the staff working at the UKCIP office in Oxford who knew the data set reasonably well and had a wide range of backgrounds; and secondly, the users of UKCP09, people who used the data set in their jobs were asked to participate, to see whether this method of representation would help their understanding of the uncertainty within the data. A significant proportion of the respondents would comprise of PhD students, research staff and faculty from UEA. They were mainly from the School of Environmental Sciences, but did consist of a wide range of people who were interested in the UKCP09 dataset, including GIS users, climate scientists and climate impact researchers, as well as MSc Climate Change students. Members of staff from Ordnance Survey (OS) were asked to participate to see how their results compared to the first case study, and to act as a control group who were unlikely to be using the UKCP09 data set, or be looking directly at climate change within their work.

The participants at UEA were recruited through informal networks and internal department advertising (n = 62). The OS staff (n = 8) were recruited using the same method as in CS1, through a gate keeper and the UKCIP staff (n = 8) and UKCP09 users (n = 3) were recruited with the help of staff in the UKCIP office. In total, 81 respondents completed the evaluation. A prize draw was included as an incentive, with two prizes of £10 shopping vouchers. The UKCP09 users were asked to complete an online version of the evaluation as they were unable to be evaluated together in a group in one physical location. There were not enough participants in the web survey to analyse the results (n = 3) and they could not easily be added to the UKCIP staff group, as the collection method was relatively different, so these results were excluded from the analysis.

5.3.3 Session overview

The evaluations were held in small groups, usually between 3 and 6 people. The participants were sent a Briefing Document one or two days before their session (see Appendix C.3) and they were asked to read and sign a Participant Information Sheet & Consent Form) on their arrival at the session (see Appendix C.4).

The evaluation sessions consisted of three main sections – the questionnaire, the maps and the discussion session. The questionnaire and maps section were both computerised, with the participants working on their own and following instructions on screen. The discussion session was completed as a group. Each session was scheduled to take around one hour, but the time varied depending on the speed of the participants in the computer based section and the levels of discussion at the end of the session.

The computer based section was run as an online evaluation, using a series of PHP pages to present the questions and data to the participants. The data from the participants (questionnaire answers and highlighted areas from the maps) were stored in a MySQL database for subsequent processing. The questionnaire and maps were developed in HTML and JavaScript in Firefox 3.5. For a technical overview and more technical details, see Appendix C.1.

5.3.4 Questionnaire

The computer based questionnaire consisted of questions covering a number of different areas (including the participants knowledge and use of the UKCP09 data set, learning style and academic background), which it was thought could have an impact on users ability to understand the sonification (see Appendix C.5 for a full list of the questions).

The questionnaire was a development of the one used in CS1, with a greater number and depth of questions to provide more background information on the participant. The amount of time that participants were willing to give was limited, so it was decided to have exclusively closed questions in the evaluation instead of the mix of open and closed questions used in CS1 as closed questions were easier and quicker to answer. The information previously captured in the open questions after the evaluation was captured in the discussion sessions instead.

Two particular questions from CS1 appeared to be important to using the sonification effectively, so they were developed for this case study. Firstly, learning style was restricted to a 'visual' or 'aural' choice in CS1, with two questions asking the participant to choose their preferred learning method. These were of limited use, so a greater range of questions was asked and a total learning score was calculated. These questions were based on Advanogy learning styles (Advanogy, 2004) and the participants were asked to choose whether statements were very much like them, a bit like them or not at all like them. Some examples of these are (the full list is available in Appendix C.5):

- Aural - 'Jingles, themes or parts of songs pop into your head at random'
- Visual - 'You can easily visualise objects, buildings, situations etc from plans or descriptions'
- Visual - 'In school you preferred art, technical drawing, geometry'
- Aural - 'Music was your favourite subject at school'

The statements were either visual or aural, and were scored with not at all (= 0), a bit (= 1) and very much (= 2). These were totalled for the visual and aural sections, then aural was subtracted from visual. If the value was positive, the participant had a visual learning style, and if it was negative they had an aural learning style.

The second question from CS1 that was expanded was the participant's knowledge of the data set. This was expanded to include knowledge and use of the UKCP09 data set itself, attendance at the PiP training sessions run by UKCIP, knowledge and use of probabilistic data, general climate change and GIS knowledge, and qualification area. Education level may be a factor that influences effective use of web map interfaces (de Mendonça and Delazari, 2012), so this was also included in the participant questionnaire. The full list can be seen in Appendix C.5 and the idea was to capture a wide a range of information about the participant's knowledge of the data set to see which factor(s) had the most influence on their ability to use the sonification successfully.

5.3.5 Development of the map interface

One of the main premises of the evaluation was to compare different visual and sonic methods to see which was most effective at representing the uncertainty within the data. The participants were shown six different maps using the Google Maps interface. The interface itself is discussed, followed by an overview of the data shown in the map and information on the representation method employed.

5.3.5.1 Google Maps API

As discussed in Section 5.2.3.1 the GMAPI was used to provide the interface for this evaluation. There are two versions of the GMAPI available for use: a JavaScript version and a Flash version. The Flash version was explored since opinion among developers suggests that Flash handles graphics and sound more effectively than JavaScript (Google Groups thread, 2008). However, it was found to be less developed than the JavaScript version in terms of general capabilities and displaying and interacting with existing spatial data; the survey was therefore constructed with the JavaScript version, adding a single Flash element to handle the sound (Ribeiro Amigo, 2010).

For visual representation, the data were added to the Google Maps interface via a KML (Keyhole Markup Language) file. The KML was created by including the data in ArcGIS and adding the colour scheme before using the KML export option. A number of ways of including the data for the sound with Google Maps were explored, but in the end the values were stored in an array (written in the JavaScript code) with their location (latitude and longitude). For the sonification, the nearest value was looked up in real time and the relevant sound was played using a Flash add-on. When the participant moved the mouse to a different grid square, the sound continued playing if the same sound was required, or if a different sound was required, the original one was stopped and the new one started.

5.3.5.2 Data sets used

The data sets used came from the UKCP09 data set, which provides a series of projections for the future climate of the UK over a series of time periods up to the 2080s, for temperature, rainfall and a number of other climatic variables. The two data sets used in this evaluation were summer mean daily temperature and temperature of the warmest day. The time periods were Baseline (1961-1990), 2020s (2010-2039) and 2050s (2040-2069). The 50th Percentile (also known as the central estimate) was used directly and the Range was calculated (90th Percentile - 10th Percentile). Participants were shown six different maps. The order that the data were shown to the participants was varied, as shown in Table 12. Varying the order was aimed at reducing the learning effect as the participants worked their way through the evaluation and this variation did not impact the participants' results. While the order of the data was varied, the order of the representation methodology was not varied (see Section 5.3.5.3). The colours chosen for the maps were based on the official UKCP09 colours, used in the online interface.

The baseline data was not available directly from UKCIP, but were downloaded from the Met Office website (Met Office, 2010). The data was provided in 5km grid squares and this was aggregated up into 25km grid squares to match the UKCP09 data set. The warmest day in summer was not provided, so this was calculated by selecting the warmest day in June, July and August for each year and then averaging the values across the years before the data were added into the evaluation.

<table>
<tr><th>Map No</th><th>Data Randomly Selected From:</th></tr>
<tr><td>1</td><td rowspan="2">EITHER Summer Mean Daily Temperature (Summer Mean) for Baseline
OR Temperature of the Warmest Day (Warmest Day) for Baseline</td></tr>
<tr><td>2</td></tr>
<tr><td>3</td><td rowspan="4">Summer Mean for 2020s OR
Summer Mean for 2050s OR
Warmest Day for 2020s OR
Warmest Day for 2050s</td></tr>
<tr><td>4</td></tr>
<tr><td>5</td></tr>
<tr><td>6</td></tr>
</table>

Table 12. The order that the data were shown to the user. The data for Maps 1 and 2 were randomly chosen from summer mean or warmest day, and the data for Maps 3-6 were randomly chosen from the four remaining data sets.

5.3.5.3 Representation methodologies

Five different representation methodologies were used for this evaluation, and participants were always shown these in the same order (see Table 13, Figure 37 and Figure 38, over). The methodologies comprised of either one or two maps, showing the 50th Percentile value of the dataset and the Range of the dataset using either visual or sonic methods.

Map No	Code	No. of Maps	Method & Data on First Map	Method & Data on Second Map
1	V	One	Visual (50th Percentile)	*(none)*
2	S	One	Visual & Sound (50th Percentile)	*(none)*
3	VV	Two	Visual (50th Percentile)	Visual (Range)
4	VSVS	Two	Visual & Sound (50th Percentile)	Visual & Sound (Range)
5	VS	One	Visual (50th Percentile) Sound (Range)	*(none)*
6	VS	One	Visual (50th Percentile) Sound (Range)	*(none)*

Table 13. Data shown for each map and information on how the data were shown to the user

These combinations of data and time period were chosen for the evaluation as they were different, but related data sets, which were required for the randomisation aspect of the evaluation. The summer mean temperature and warmest day data sets (both 50th Percentile and Range data) had different spatial distributions and the data over the different time periods generally had similar distributions but varied in magnitude.

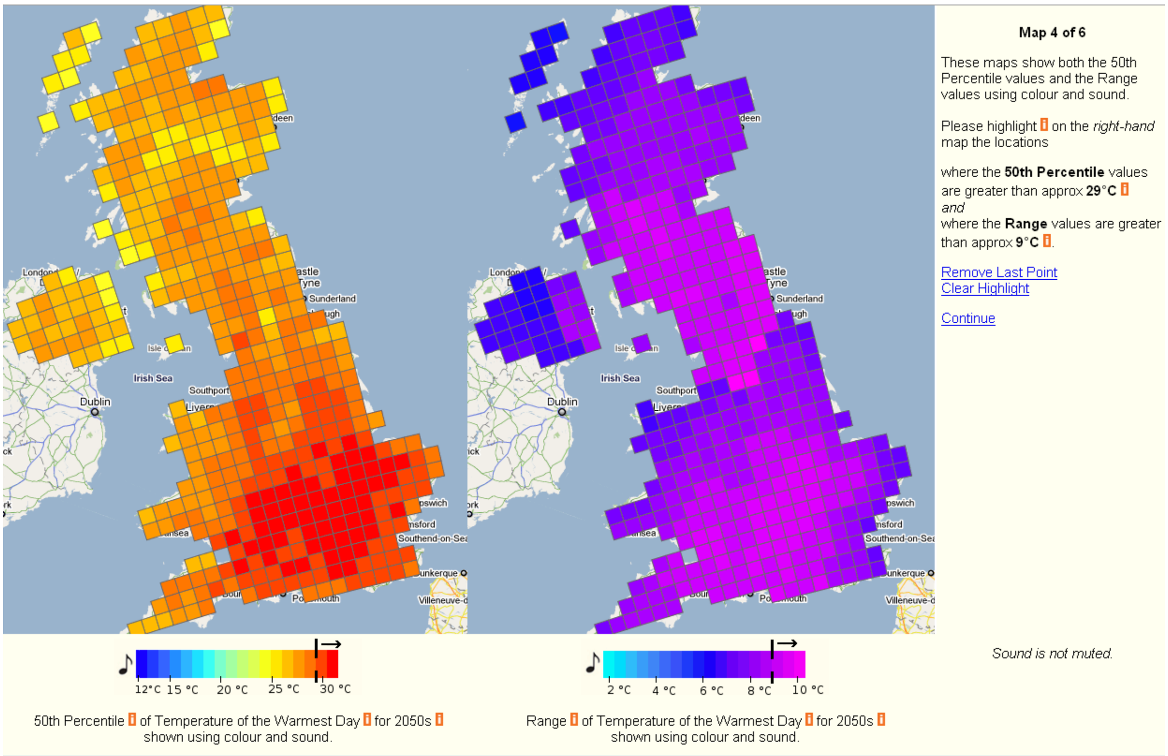


Figure 37. Screenshot of Map 4, an example of VSVS where vision and sound are used to represent both the 50th Percentile temperature data and the range data. See video clip referenced at Appendix C.6 for an example of this method. A similar clip is also available online at http://vimeo.com/17029341.

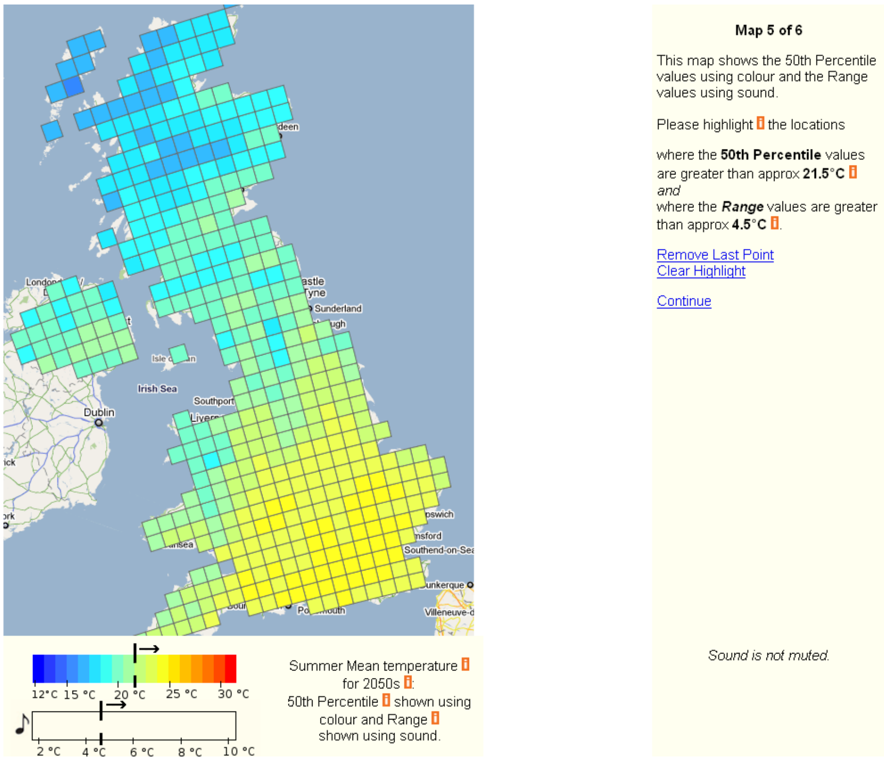


Figure 38. Screenshot of Map 5, an example of VS where the 50th Percentile temperature data are represented visually and the range data are shown using sound. See video clip referenced at Appendix C.6 for an example of this method. A similar clip is also available online at http://vimeo.com/17029358.

Figure 37 shows the 50th Percentile and Range values for the warmest day for 2050s. The data are shown both visually and sonically (VSVS method) with the visual method using colour ranges as shown in the legend at the bottom, and the sound using different notes (as explained in Section 5.3.1). See video referenced at Appendix C.6 for a clip of the sound interface being used. The participant is asked to highlight the area where the 50th Percentile exceeds 29°C and where the Range exceeds 9°C and then to highlight this combined area (i.e. where 50th Percentile > 29°C AND where Range > 9°C) on the right hand map. Figure 38 shows the 50th Percentile and Range values for the Summer Mean temperature for the 2050s. The 50th Percentile data are shown visually (using colours, as before) and the Range data are shown sonically, using different notes, which is the VS method (see video clip referenced at Appendix C.6 for an example). However, the data are both shown on the same map, which is different to previous methods. The participant is asked to highlight area where the 50th Percentile exceeds 21.5°C and where the Range exceeds 4.5°C on the map.

For Maps 1 and 2, participants were asked to highlight the area where the variable shown (summer mean for baseline or warmest for baseline) exceeded the stated value. For Maps 3 to 6, participants were asked to highlight the areas where the 50th Percentile and the Range exceeded the specified values. Maps 1 and 2 worked as a training section to allow users to familiarise themselves with the interface and the highlighting procedure. The main focus of the evaluation was Maps 3 to 6, but the first two maps added information to the results.

Initially the main task was going to use just one variable, however the results from the pilot studies showed that this was too easy as the results were consistently good. The mental load was increased to ensure a wider variety of scores by asking the participant to highlight the areas where both variables exceed the specified thresholds. Figure 39 shows some examples of areas that users highlighted, using the VS methodology.

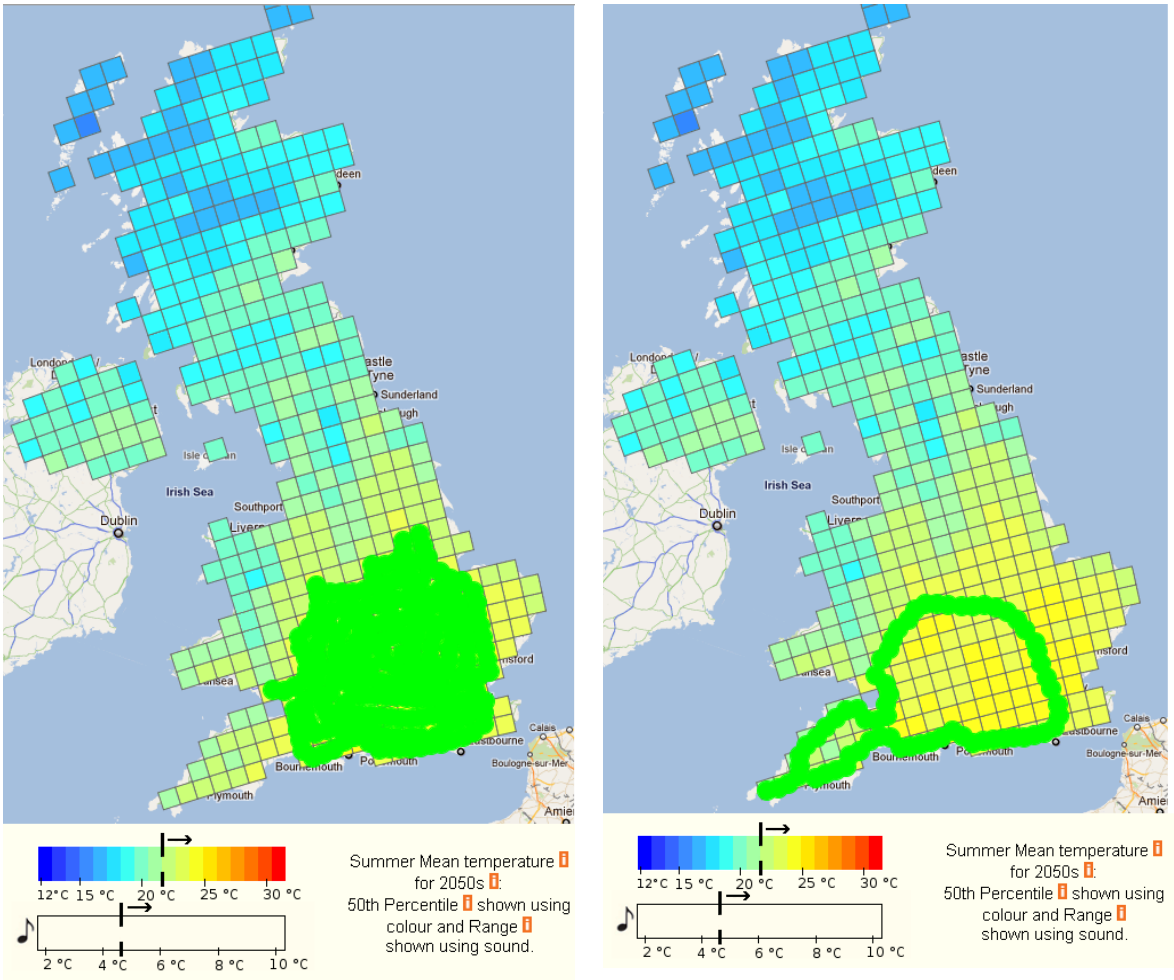


Figure 39. Two examples of highlighted areas provided by participants. If an outline was provided (like the example on the right) this was filled in manually before the analysis.

In a standard evaluation design it is standard practice to randomise the order of the different maps. This was attempted in the pilot study, but the VS stages were too complex for the participants to understand. They said they needed the simpler maps (such as VV or VSVS) beforehand to introduce the concept before moving on to the more complex VS map.

5.3.6 Discussion sessions

Discussion sessions were held at the end of each evaluation session, taking around 20-30 minutes, although this varied between groups. A semi-structured method was followed, but the discussion was allowed to take its own course if the participants were suitably enthused. The discussions were recorded (with consent) and subsequently summarised.

5.3.7 Session setting

The UKCIP staff and OS sessions were run at their respective offices, in groups of between 6 and 8 participants. The OS had a computer teaching room where the exercise was run, and the UKCIP session was run in a meeting room, using the training laptops available from UKCIP. The UEA sessions were run in the same computer lab as CS1 (see Section 3.7.1 for details). The lab could hold 20 but the sessions were run in groups of 4-8 to make group management easier and to ensure everyone could contribute to the discussion sessions.

5.3.8 Web only evaluation for UKCP09 users

One group of users (UKCP09 users) were very geographically dispersed and did not come together on a regular basis. As a result it was decided to create a version of the evaluation that could be completed online without supervision. This was very similar to the supervised version of the evaluation, but had a greater detail of instruction and contact information if the participants required assistance. De Mendonça and Delazari (2012) completed a similar online evaluation on web map use without supervision, but had significant issues with survey completion. Of the participants that took the survey, only data from 55% of the participants could be used in the evaluation because the rest did not complete a sufficient amount of the evaluation. Reasons for the lack of completion were not apparent, but stand alone web surveys do require specific design and support considerations that are not necessary in supervised surveys (Bearman and Appleton, 2011).

5.4 Results and Discussion

This section will explain how the results were processed, discuss their validity and relate them to the original research questions.

Participants' responses to the questionnaire were recorded in a MySQL database and the highlighted areas for each of the 486 maps (81 participants, each with 6 maps) were also recorded as a series of lat/long points. A Python script (using ArcGIS 9.3 libraries) was written to read these into a point shape file, which was processed using a point-in-polygon analysis to calculate which UKCP09 grid cells were highlighted by the participant for each map. Manual checking highlighted two issues. Firstly, that the relatively large size of the paintbrush tool may have led users to believe a particular square had been selected when the precise point location actually lay in an adjacent square. This resulted in gaps in the highlighted area after processing, which were filled in manually. Secondly, some users drew an outline around the area of interest rather than highlighting the whole area (see Figure 40, over for examples). These outlines were filled in manually. Those participants who highlighted the area using an outline (rather than filling in the whole area) were quicker to complete the map tasks, but there were no other differences in the results. These issues could in the future be addressed at the design stage by allowing users to toggle grid cells' selection status, rather than draw on the map, but this was not possible in the available time for this case study. The highlighted areas for one participant could not be discerned, as it consisted of a series of horizontal lines and it was not clear which side of the line was the highlighted area. This participant’s responses were excluded as well as one participant who had a very poor result (see Section 5.4.1 for details).

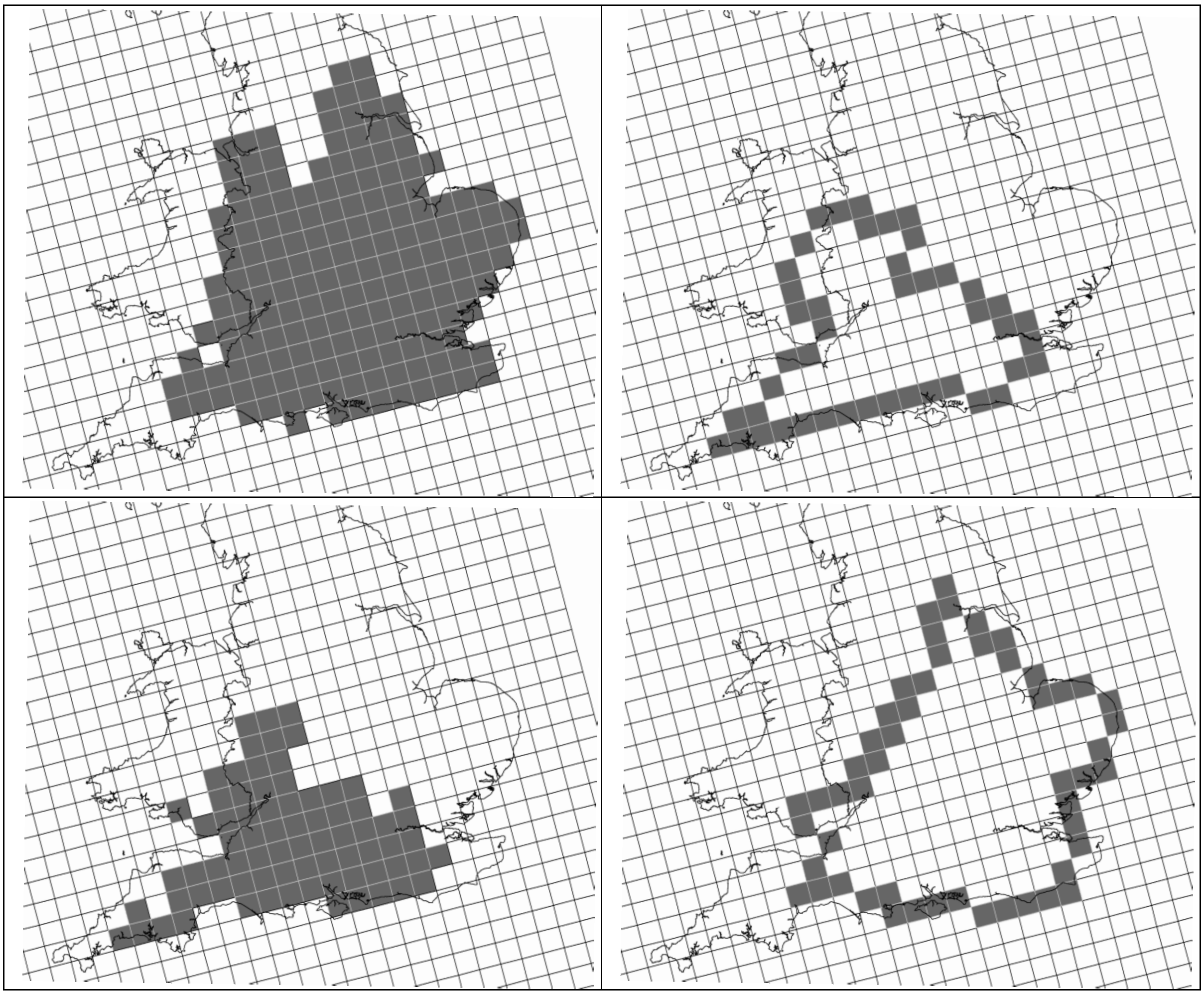


Figure 40. Examples of different user’s responses. The responses where the user just provided an outline (right) were filled in manually before further processing.

5.4.1 Phi calculation

For each participant and map combination, the highlighted area was compared with the correct area (for the relevant combination of data and time period) using Pearson's Phi measure of agreement for binary data (Field, 2000a, p.695). This coefficient is a key part of the results and was the main method of summarising how effectively people understood the method used to represent the data. The Phi values were calculated for each map the participants created, by comparing the highlighted areas with the 'correct' areas – i.e. the area the participant should have highlighted. The Phi score was calculated by creating a 2x2 matrix (see Table 14) of the counts of the cells that were and were not highlighted (user highlighted) against the cells that should and should not have been highlighted for that data combination (correct answer). Only the land based cells were used; any highlighted cells that were sea (and therefore not part of the UKCP09 data set) were excluded. The Phi value was calculated using the formula specified in Equation 1.

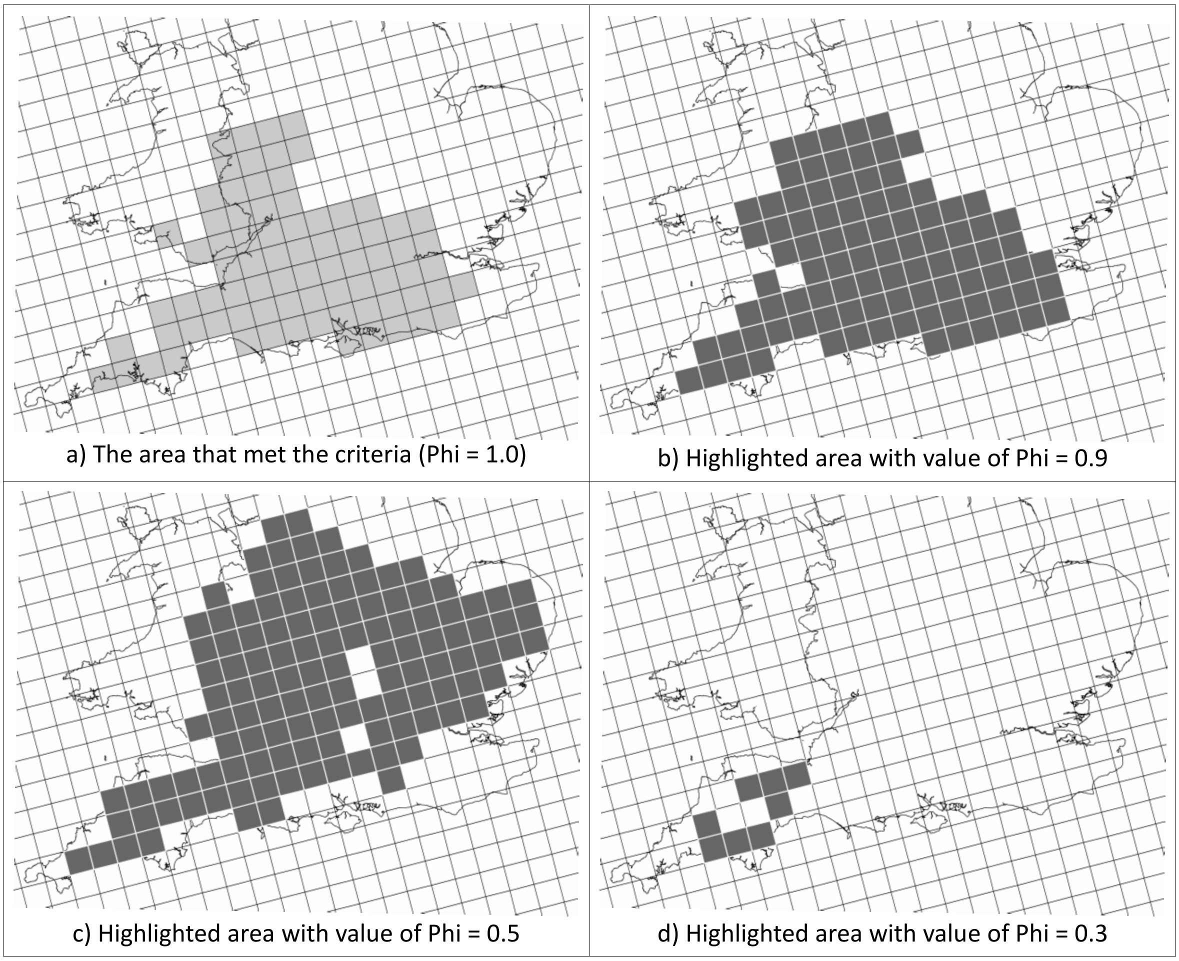
		User Highlighted		
		0	**1**	**Total**
Correct Answer	**0**	a	b	e
	1	c	d	f
	Total	g	h	n

Table 14. Example matrix of the values used for the Phi equation

$$\varphi = \frac{ad - bc}{\sqrt{efgh}}$$

Equation 1. Formula used to calculate the Phi value for each map. Values *a, b, c* and *d* relate to the table above and ϕ is the Phi value.

The possible values of Phi ranged from +1 to -1. A value of +1 would indicate that the participant highlighted exactly the cells required, where no incorrect cells were highlighted and no correct cells were missed. A value of -1 would mean that the participant highlighted the exact opposite of what was required, where all the cells that should have been highlighted were blank, and all cells that should not have been highlighted were highlighted. A value of 0 would mean the cells had been highlighted completely randomly. The values from this evaluation ranged between 1.0 and 0.2, see Figure 41 over for a range of examples. One Phi value from one participant was much lower than the rest (-0.3) so this participant’s results were excluded from the analysis.



a) The area that met the criteria (Phi = 1.0)

b) Highlighted area with value of Phi = 0.9

c) Highlighted area with value of Phi = 0.5

d) Highlighted area with value of Phi = 0.3

Figure 41. The 'correct' answer (a) and some examples of user responses (b-d), shown with their calculated Phi values. The Phi value would be reduced by either incorrectly selecting additional areas (e.g. c) or not selecting required areas (e.g. d) or a combination of the two.

5.4.2 Basic observed participant characteristics

The series of tables over shows an overview of the participant‘s characteristics. The majority of the participants were from UEA, so it was not possible to compare the OS or UKCIP groups to each other or to UEA in any meaningful way. A number of the questions had very skewed answers (e.g. Q3, Q19), which meant that the results could not be analysed using these factors. Some of the questions had more balanced responses (e.g. Q1, Q2) and analysis using these factors was carried out. Some of the participants were removed from the analysis, or groups combined because of a small number of issues; see questions marked * for details.

The participant’s knowledge of climate change and GIS / spatial data (Q1 and Q2) was distributed much as expected, with the majority of participants having a medium level of knowledge. A larger proportion (45.1%) of the participants said that uncertainty had no relevance to their work (Q10) but their results were no different to the rest of the participants. Participant’s confidence in using the UKCP09 data (Q11) was toward the lower end of the scale, but this skew is not surprising, considering that some of them had not used the data before.

Organisation*	Ordnance Survey	UEA	UKCIP
	11.3% (8)	81.7% (58)	7.0% (5)

Question	1 (A lot)	2	3	4	5 (Nothing)
Q1 Knowledge of Climate Change	9.9% (7)	25.4% (18)	50.7% (36)	14.1% (10)	0% (0)
Q2 Knowledge of GIS / Spatial Data	15.5% (11)	11.3% (8)	39.4% (28)	21.1% (15)	12.7% (9)
Q10 How relevant is uncertainty to your work?	15.5% (11)	15.5% (11)	9.9% (7)	14.1% (4)	45.1% (32)
Q11 How confident are you that you are using (or could use) the UKCP09 data correctly?	7.0% (5)	12.7% (9)	26.8% (19)	28.2% (20)	25.4% (18)
Q12 Knowledge of Probabilistic Data	8.5% (6)	19.7% (14)	38.0% (27)	29.6% (21)	4.2% (3)

Question	Yes	No
Q3 Are you a member of the Royal Metrological Society?	4.2% (3)	95.8% (68)
Q4* Were you aware of the UKCP09 data set before today?	54.9% (39)	45.1% (32)
Q5 Have you used the Projections website?	25.4% (18)	74.6% (53)
Q6 Have you attended the PiP training sessions?	7.0% (5)	93.0% (66)
Q7 Have you used the Moodle online training?	11.3% (8)	88.7% (63)
Q9 Are you aware that the UKCP09 data contains probability information?	52.1% (37)	47.9% (34)
Q18 Have you been involved in any of the author's previous studies?	33.8% (24)	66.2% (47)
Q19 Are you colour blind?	0% (0)	100% (71)

Question	High	Low
Q13 Do you have a high or low knowledge of betting?	47.9% (34)	52.1% (37)

Question				
Q8* Have you downloaded data from the user interface?	Yes – For a Specific Project - 15.5% (11)	Yes – Just to see how to do it*	No – 84.5% (60)	
Q20a Qualification Level	PhD - 74.6% (53)	MSc - 18.3% (13)	BSc - 1.4% (1)	High School – 5.6% (4)
Q20b* Qualification Area	UKCP09 – 2.8% (2)	Climate Change – 12.7% (9)	Environmental Sciences – 59.2% (42)	Other – 25.4% (18)

Table 15. Overview of responses to the questionnaire. Starred questions have information below:

*The number of UKCP09 Users who completed the online version of the evaluation was small (n = 3) so they were removed from the data set. The number of participants completing the UEA MSc Climate Change was also small (n = 3) and these were combined with the main UEA group.
*Q4 originally asked for the duration the participant was aware of UKCP09, but this was simplified to those who know about UKCP09 and those who didn't, because the duration of awareness had no impact.
*Q8 originally had a third response of 'Yes – Just to see how to use the interface' (n = 8) but these were removed from the data set because they were not consistently different from the other results and the group was too small to have any significant difference.
*Q20b data were collapsed into two groups – Climate Change (UKCP09 & Climate Change) and Not Climate Change (Environmental Sciences & Other) because of the small numbers.

Knowledge of probabilistic data (Q12) has a reasonably normal distribution and this did have an impact on the participants results; those with knowledge of probabilistic data did perform more effectively. Awareness that the data contained probabilistic information was very similar to participants who were aware of the data (52.1%), which suggests that those who knew about the data knew how it was different to previous iterations. Awareness of the data before the evaluation was more-or-less split 50-50 between those who knew and those who did not (Q4), but only 25% had accessed the projections website (Q5), 15.5% had used the data in a specific project (Q8) and even fewer attended the PiP training (Q6 - 7%). A similar trend is shown with the online training (Q7). Only a third of the participants had been involved in the previous case study (Q18) and the qualification level and area were biased to PhD level in Environmental Science, due to the high dominance of UEA based (and School of Environmental Sciences based) participants. The learning style results have not been included in these tables because they are difficult to represent (as raw scores between +14 and -14) and do not have any relationship with any other variable (such as organisation). These do not appear to have any impact on the results, as discussed later in the results section.

Results from some questions were combined, to allow sufficient sample size for comparisons (see starred text in Table 15). Data could be split into sufficiently sized subgroups, particularly for the knowledge questions, awareness and use of the data. Organisation, qualification, RMetSoc membership and training attendance were not evenly distributed enough to allow comparison. A small number (n = 3) of participants were completing the UEA's MSc in Climate Change degree. The whole year group was approached to take part, but only three completed the evaluation. Their results were not different to those for the result of the UEA group, so the MSc students were combined into the larger UEA group.

This sample was reasonably representative of some of the target audience, who were academic experts in climate change using the data and academic experts who use the data but do not specialise in climate change data. However, non-academic expert participants of the data (i.e. UKCP09 users) were not very well represented at all. It was planned that these users would complete the online version of the evaluation, but because of the very small number of respondents these results could not be used. The sample is not representative of the general public, but this was not the intention. However, the general public is a potential target audience for this technique as a novel and interesting way of representing this data set, which could be evaluated in future research.

The basic duration (seconds) and Phi values were initially considered. Duration showed a reasonably straight forward pattern, with Maps 1 and 2 being much faster than Maps 3 – 6 (Table 16). This is discussed later in Section 5.4.7.

Map \ Duration	Mean	SD	Min	Max	Range
Map 1 (V)	143.920	79.628	29.281	436.388	407.107
Map 2 (S)	152.894	103.696	55.529	700.947	645.418
Map 3 (VV)	246.523	145.582	40.374	819.100	778.726
Map 4 (VSVS)	239.332	139.416	38.837	703.573	664.736
Map 5 (VS)	261.493	130.422	47.062	596.605	549.543
Map 6 (VS)	215.443	124.730	37.750	858.777	821.027

Table 16. Mean duration of the responses, split by Map, with SD, min, max and range.

The mean values of Phi were considered initially, but it soon became apparent that there was a very wide range of values, which the mean did not accurately reflect (see Table 17).

Map \ Phi	Mean	SD	Min	Max	Range
Map 1 (V)	0.861	0.156	0.313	0.989	0.676
Map 2 (S)	0.913	0.100	0.517	0.994	0.477
Map 3 (VV)	0.680	0.213	0.204	1.000	0.796
Map 4 (VSVS)	0.786	0.202	0.293	1.000	0.707
Map 5 (VS)	0.783	0.180	0.227	0.986	0.758
Map 6 (VS)	0.821	0.147	0.479	1.000	0.521

Table 17. The values of the Phi variables for all respondents for each map. Map 1 (V) shows one data set visually, Map 2 (S) shows one data set visually & sonically, Map 3 (VV) shows two data sets visually, Map 4 (VSVS) shows two data sets visually and sonically, Maps 5 & 6 (VS) show one data set visually and one data set sonically. For a more detailed description, see Section 5.3.5 (page 149).

Further analysis was undertaken, as the range was quite high for all of the Maps. However, these data show that when sound is added to an existing visual method (e.g. Map 1 → Map 2, and Map 3 → Map 4) the mean Phi score increases and the range decreases. This is also seen when the same method is repeated (e.g. Map 5 → Map 6), so this is likely to be the addition of sound increasing the score; a learning effect as participants become more familiar with the interface and specific representation methodology, or a combination of the two. There were no significant differences between the results for the two different data sets used, summer mean or warmest day data.

5.4.3 Individuals graph

To get a better handling of the results, the individual Phi values were plotted for each participant for each of the six maps (see Figure 42, over).

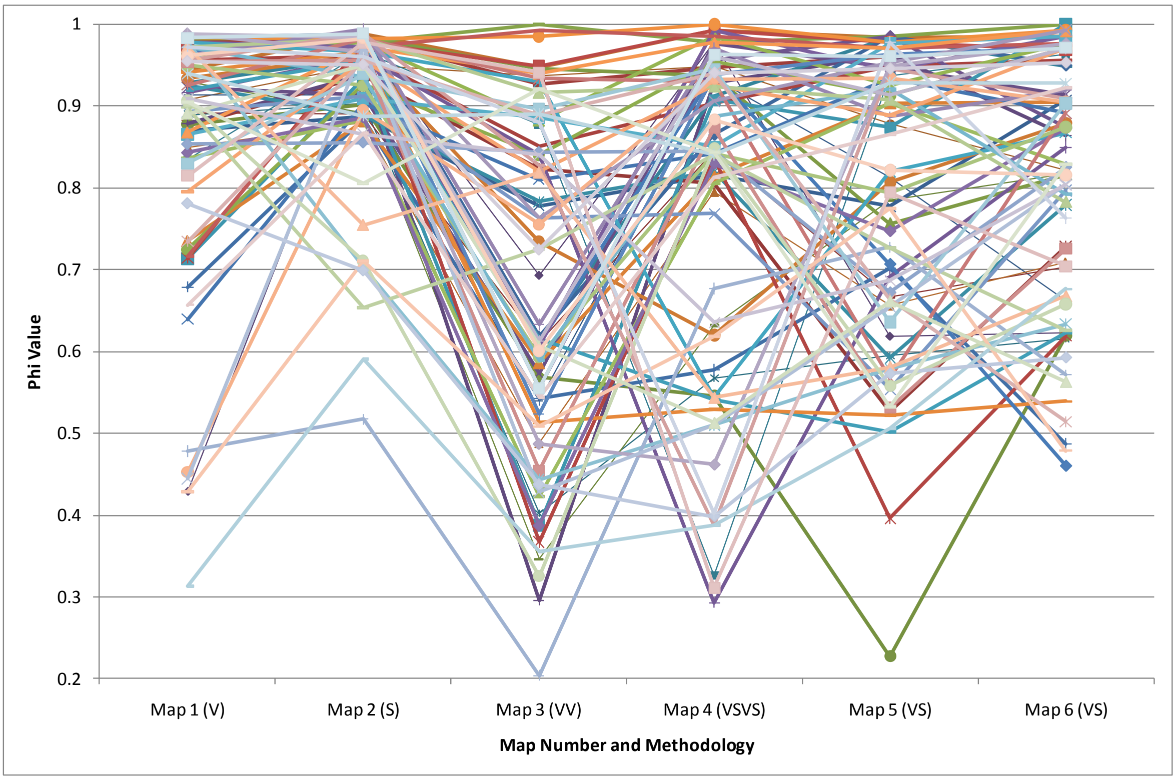


Figure 42. Each individuals result for each of the six maps.

Figure 42, showing each individual's results for the six maps looks complex, but there are some clear trends. Some users perform well (0.85 < Phi < 1.0) for all maps and most users perform well for Maps 1 and 2. Many users score fairly low on Map 3 (VV), but then improve for Map 4 (VSVS). There is an inverse group, who perform well on Map 3 but badly on Map 4. Phi values for Maps 5 and 6 vary widely.

In general, adding sound to a visual method (V -> S and VV -> VSVS) improves the Phi scores for most participants, but there are a few exceptions to this. There is also likely to be a learning effect, as shown between Maps 5 and 6 for most of the participants. A large number of participants found VV difficult, and this is because they were asked to select the area on the map where two criteria were met, rather than just one criterion as they were asked previously. This is a significantly more complex spatial operation and during the discussion session people commented on the increased mental effort required to complete this task.

A number of independent samples t-tests were used to compare the mean Phi values for Maps 3 to 6 (see Table 18). Map 3 scored significantly lower than the other maps (p = 0.005). This shows that adding sound in some manner makes a significant improvement to the score (Map 3 compared to the others), but which method works best is unclear. There is also a learning effect (comparing Maps 5 and 6), but this difference is not significant.

	Map 3 (VV)	Map 4 (VSVS)	Map 5(VS)	Map 6 (VS)
Map 3 (VV)	-	-	-	-
Map 4 (VSVS)	0.005*	-	-	-
Map 5 (VS)	0.004*	0.968	-	-
Map 6 (VS)	0.000*	0.198	0.159	-

Table 18. Independent samples t-test, comparing the means of Phi for the four maps. Significant p values are starred*.

5.4.4 Learning effect

From the results there was a learning effect where the later maps have better Phi scores than the earlier maps. This is also shown in the time taken to complete the maps. Unfortunately, it was not possible to separate the learning effect from the different representation methodologies because of the design of the study, as the different representation methodologies were always shown to the participants in the same order and the only one to be repeated was VS (Maps 5 and 6). A future experiment would ideally randomise the order of the maps to negate this issue. Randomisation was considered for this experiment in the early stages, but during the pilot testing it was found to be too complex for the participants. In the future, a different approach to the randomisation could be used, which might enable it to be used effectively. The main issue with the randomisation in the pilot study was that it had the potential to start with a relatively complex interface with very little introduction. A more explicit training session and/or demonstration by the session leader could reduce this problem, but it would make the evaluation longer.

5.4.5 Phi elements

As explained in the Phi formula (Equation 1, page 156) there are four different elements that contribute to the calculation to the Phi value; a, b, c & d. These could also be thought of as cells that were correctly highlighted (d, 11), cells that were correctly not highlighted (a, 00), cells that were incorrectly highlighted (b, 01) and cells that were incorrectly not highlighted (c, 10) (see Table 19).

Should the cell have been highlighted by the user?	Was the cell highlighted by the user?	Formula letter	Text description	Was this correct or incorrect?	
0	0	a	The cell should not have been highlighted and was not highlighted.	Correct	
0	1	b	The cell should not have been highlighted, but was highlighted.	Incorrect	Highlighted more cells than necessary
1	0	c	The cell should have been highlighted, but was not highlighted.	Incorrect	Highlighted fewer cells than necessary
1	1	d	The cell should have been highlighted, and was highlighted	Correct	

Table 19. Explanation of the different scores that were entered into the Phi equation. 0 = No, 1 = Yes

Category	Map 1 (V)	Map 2 (S)	Map 3 (VV)	Map 4 (VSVS)	Map 5 (VS)	Map 6 (VS)	Correct Answer
00 a	0.96	0.97	0.88	0.93	0.93	0.95	1
01 b	0.04	0.03	0.12	0.07	0.07	0.05	0
10 c	0.11	0.07	0.18	0.13	0.13	0.12	0
11 d	0.89	0.93	0.82	0.87	0.87	0.88	1

Table 20. Proportion of cell counts for the Phi coefficient components, as explained in the main text. If the user had highlighted exactly the correct area, category 00 and 11 would have a value of 1 and category 01 and 10 a value of 0.

The cell counts were converted into proportions, as the number of cells for each category varied depending on the data used (see Table 20). The scores vary from 0 (no cells) to 1 (all possible cells). The count of 00 cells (a) is relatively high all cases, as there were a large number of cells that did not need to be selected, for example in Northern Ireland, Scotland & northern England.

These four components were compared and showed very similar trends to those seen in the overall Phi result. There is an interesting trend with the 'learning effect', where participants are selecting fewer extra cells (category 01) as they progress through the maps, whereas the missed cells (category 10) stay roughly the same.

5.4.6 Clusters (phi and duration)

5.4.6.1 Grouping the results

While the scores varied between maps, there were some clear trends and four distinct groups could be seen in Figure 42 (page 161). The first stage needed to establish the number of clusters to split the data into. A hierarchical cluster analysis was performed, using Ward's method (the most common) and 6 clusters appeared to be the optimal number for this data (Field, 2000b; Everitt, 1980; Ward, 1963). The hierarchical method does have a number of disadvantages; the main one being that due to the nature of the analysis, early 'bad judgments' (in terms of the first level of clustering) cannot be changed at a later stage. Additionally, the output will be impacted by the order of the data and by the method chosen to create the clusters (Field, 2000b). However, one major advantage is that it does not require a number of clusters to be specified *a priori* (before the analysis starts). This is important to avoid because if a number of clusters is chosen, it would significantly impact the results.

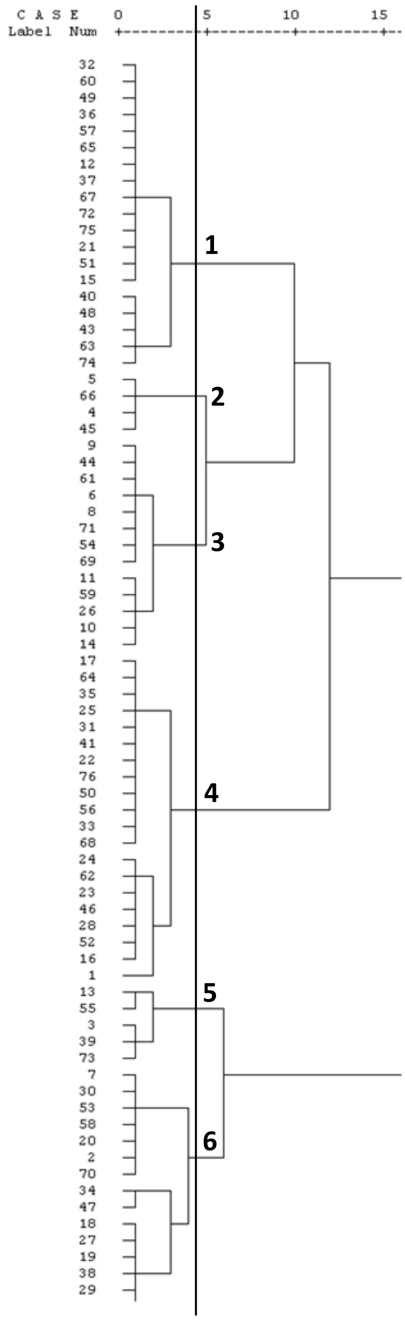


Figure 43. Dendogram from the hierarchical cluster analysis, showing the threshold for the six clusters and the clusters themselves.

The second stage took the number of groups from the first stage and split the data up into these groups using a K-means cluster analysis. This analysis required a number of clusters to be specified, then uses an iterative method to assign the points to the specified number of clusters. Obviously the number of clusters has a significant impact on the output and 6 clusters was chosen as this was the optimal number from the hierarchical clustering (see Figure 43). There was a reasonable level of overlap between the membership groups of the hierarchical cluster analysis and the K-means cluster analysis, but this was nowhere near 100%, so the two stage process of the analysis was required.

The six clusters each had different trends for the 6 maps (see Figure 44), but there were some commonalities between the clusters. However, none matched up with the organisations (see Section 5.4.8.1 for details).

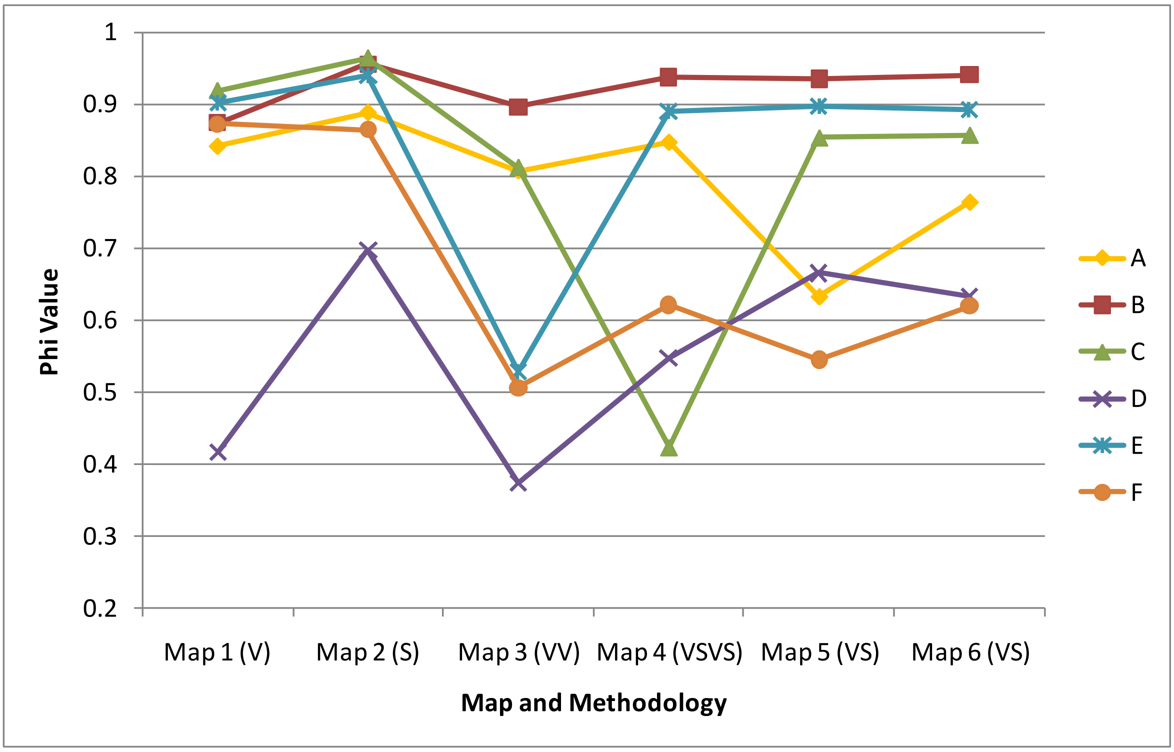


Figure 44. Graph of the six clusters and their Phi values. Each cluster has a unique trend, see main text for full description.

Table 21 shows the responses and Phi values for each of the six clusters and is referred to in the text below. Cluster B performed well for all of the maps and it showed a slight improvement when adding sound (V to S and VV to VSVS). Cluster A also performed reasonably well in Maps 1 to 4, but had difficulty with Maps 5 & 6 (VS). The lower performance in VS is explained by a more visual learning style (1.33) meaning that they had trouble with VS because they could no longer rely on the visual stimuli and had to rely on the sound. Cluster A was also much faster than other clusters (except D, see below).

Cluster E generally performed very well, apart from Map 3 (VV), which took the participants much longer to complete than any of the other maps (263.34 seconds). Their limited experience of the data explains this low score as Map 3 is the first time they have see the Range data and they required the extra time to understand the data. Cluster C also generally performed well, but scored very low on Map 4 (VSVS). The cluster was fairly small (n = 7), which limits the information that can be interpreted from the results, but it may be that for Map 4 they were more interested in the sound than the exercise, which would explain their low score. Clusters C and E have a number of differences, including the fact that cluster C had used the interface and cluster E had not used the interface (Q8, the difference approached significance, p = 0.264). In addition cluster

E had probabilistic data knowledge, cluster C did not, (Q12, a significant difference, p = 0.032) and cluster C performed badly at Map 2 (VSVS) and cluster E performed badly at Map 1 (VV).

Cluster	A	B	C	D	E	F	All
n	12	17	7	4	19	12	71
% from UEA	75%	94.1%	100%	100%	73.7%	66.7%	81.7%
Learning Style (mean)	1.33	0.06	1.43	-1.75	2.63	-1.67	0.70
Learning Style (range)	19	28	8	8	21	22	28
Map 1 (V) Phi	0.841	0.893	0.920	0.416	0.899	0.889	0.861
Map 2 (S)	0.891	0.960	0.965	0.698	0.941	0.863	0.913
Map 3 (VV)	0.805	0.906	0.813	0.374	0.526	0.503	0.680
Map 4 (VSVS)	0.847	0.944	0.423	0.548	0.891	0.625	0.786
Map 5 (VS)	0.629	0.938	0.854	0.666	0.897	0.532	0.783
Map 6 (VS)	0.759	0.942	0.858	0.633	0.898	0.630	0.821
Map 1 (V) Duration	102.522	140.944	193.099	100.609	134.260	190.577	143.920
Map 2 (S)	96.776	138.003	241.096	117.867	145.036	202.771	152.894
Map 3 (VV)	159.672	250.443	302.364	179.220	263.34	291.047	246.523
Map 4 (VSVS)	164.735	244.636	265.019	158.186	221.271	347.079	239.332
Map 5 (VS)	208.222	286.038	309.619	181.266	212.639	356.009	261.493
Map 6 (VS)	175.107	246.195	217.947	119.281	226.125	225.892	215.443
Q1 Knowledge of Climate Change (1 = high, 5 = low)	2.42	2.41	2.71	3.25	2.68	3.17	2.69
Q12ProbKnow (0=No, 1=Yes)	0.25	0.59	0.14	0.00	0.58	0.33	0.41
Q8UserInt (0=No, 1=Yes)	0.25	0.24	0.29	0.00	0.05	0.08	0.15
Q5UsedData (0=No, 1=Yes)	0.25	0.35	0.29	0.00	0.21	0.25	0.25
Q2GI Knowledge (1 = high, 5 = low)	2.67	3.06	3.29	3.75	2.53	3.83	3.04

Table 21. Responses and Phi values for the six clusters identified in the cluster analysis.

Cluster F performed well on Maps 1 and 2 (V and S), but then did not perform as well on Maps 3-6. These maps were the evaluations using two maps (3 and 4) or one map representing two variables (5 and 6), and the participants GIS knowledge was lower (3.83, compared to 2.62 overall), which explains this change. Lower GIS knowledge means that the participants had less experience of GIS which reduced their performance on the more complex maps. The low GIS knowledge also explains the lower times, particularly for Maps 4 and 5 as these were more complex spatial processes. Cluster D was very small (n = 4) and scored relatively badly on all of the maps. They also did not spend very long on the maps, which suggests they did not understand the task or did not find the task suitably engaging. Their knowledge and use of the UKCP09 data was very low, which is another factor that explains the low scores.

There are distinct differences between the clusters and some of these differences are significant. It is not clear what caused these differences as there were no significant patterns in the participants' characteristics that related to the clustering. It is possible that a larger sample would help with understanding this, as some of the clusters are quite small.

5.4.6.2 Clusters and Phi Components

Using the clusters previously devised, the different components of the Phi coefficient were compared, as explained in Section 5.4.1. Most of the clusters showed the same pattern, but there were two that were markedly different.

Cluster F highlighted slightly more cells that necessary on Maps 3-6 (see Figure 45) meaning they did not understand the 'AND' GIS nature of the question, which would be consistent with their lower GIS knowledge. They may have understood 'AND' as select the 50th Percentile cells that meet the criteria and the range cells that meet the criteria rather than the cells that met both of the criteria. Cluster D highlighted a lot more cells on the first map than necessary and this is likely to be because they were working out how to use the interface.

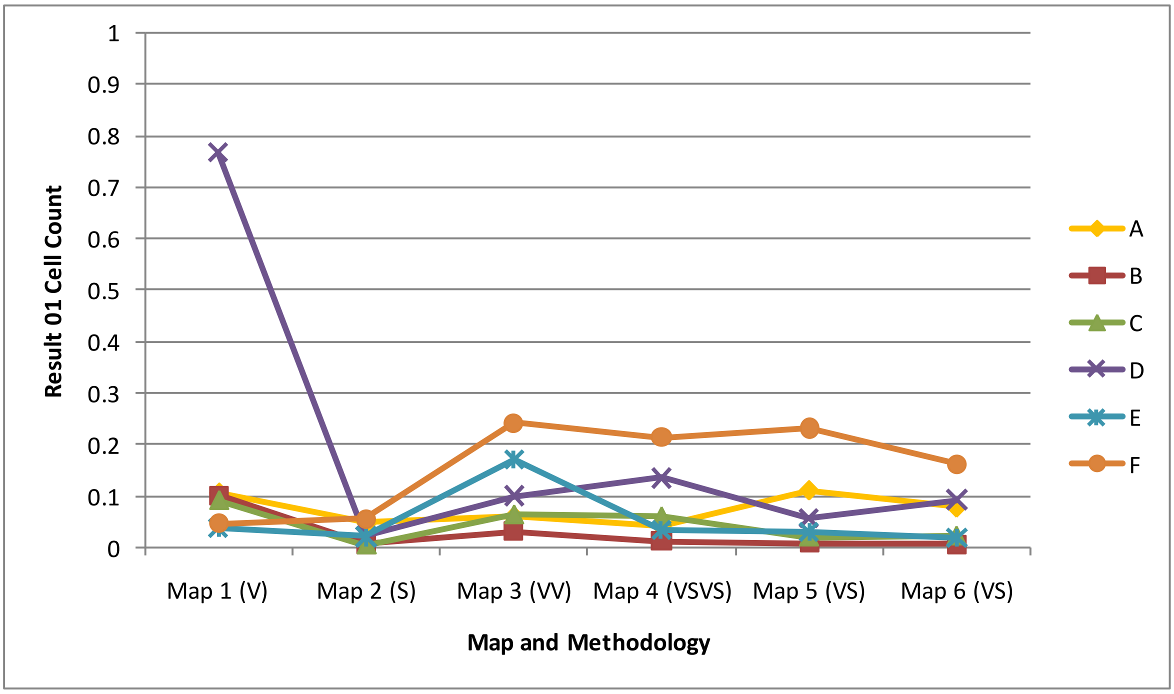


Figure 45. Cells that were highlighted that should not have been highlighted, by cluster. Note how cluster F select more extra cells than any other cluster, and cluster D results for Map 1.

Generally participants highlighted required cells (incorrectly not highlighted, 10) and did not select incorrect ones. Cluster D is the main exception to this, leaving out many of the required cells (see Figure 46). Cluster C missed many cells in Map 4 and this could be because they were not used to the use of sound in this way, which may have confused them.

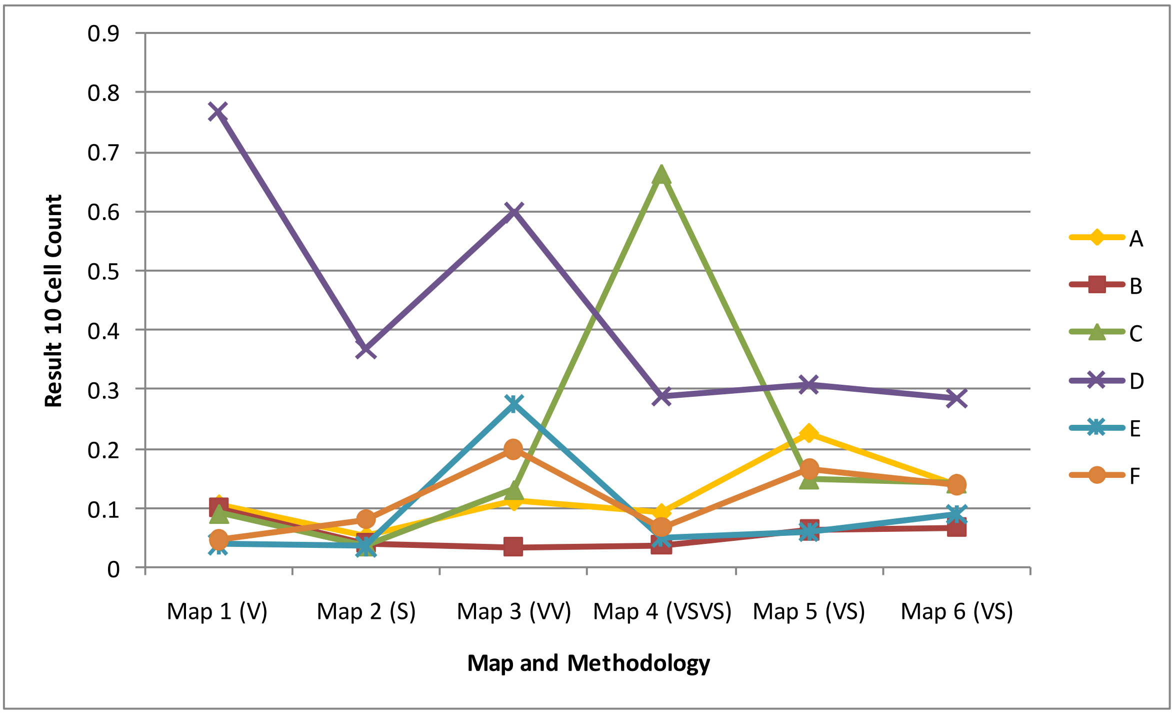


Figure 46. Cells that should have been highlighted, but were not. Note cluster D, who missed out more cells than the other clusters, and cluster C who had trouble with Map 4.

Apart from Cluster D (as discussed above) participants tended to overestimate the cells, rather than underestimate them (i.e. select ones they did not need to, rather than not select ones they should).

5.4.7 Duration

Using the clusters derived from the cluster analysis previously, there were some interesting trends when looking at the duration participants took to complete each map (see Figure 47).

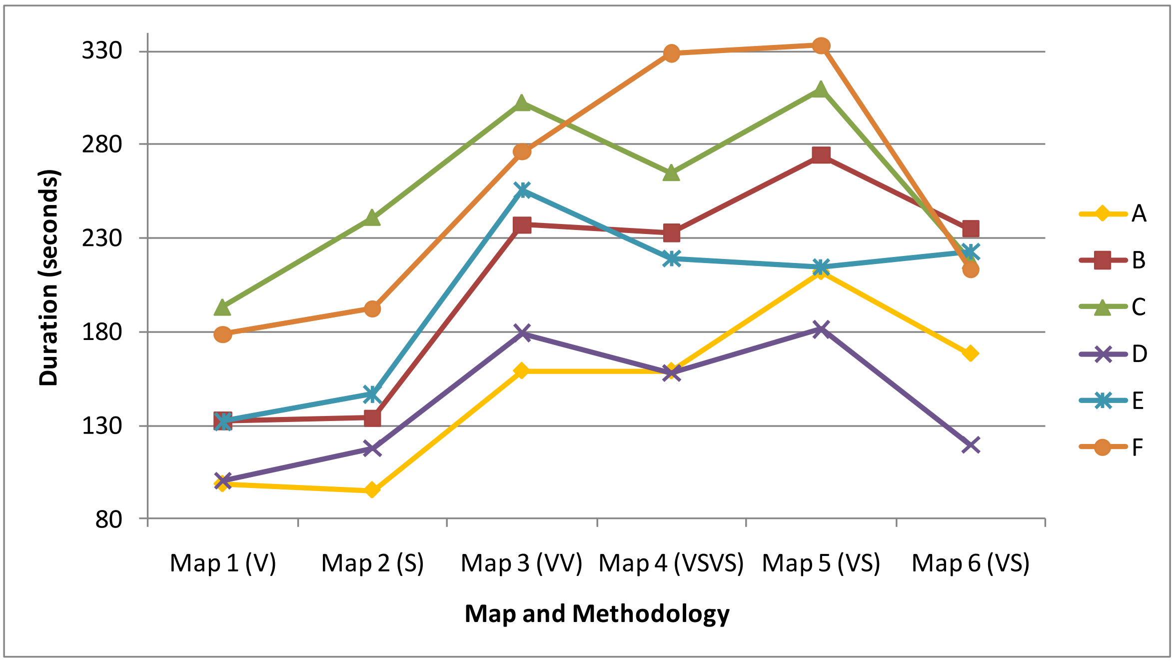


Figure 47. Graph of the results from the six clusters and duration taken, in seconds.

The two training maps were relatively simple and quite quick to complete. The four main maps were harder because they asked the participants to highlight the area on one map that met the criteria on two maps, which requires more mental effort. This was shown by the higher durations taken for these maps (see Figure 47) and from the participant responses in the discussion sessions. The most marked increase was from showing the data on one map to showing it on two maps (Map 2 -> Map 3), but there was also an increase from VSVS to VS (Map 4 -> Map 5), which indicates that the VS methodology was harder than the previous ones. The learning effect can be seen very clearly, with the duration for the second VS map (Map 6) much lower in nearly all of the clusters.

In general, all of the clusters followed the same pattern with the timing, but the magnitudes varied. All of the clusters took longer on VS than VSVS, which is likely to mean it took more mental processing. The exception was cluster E, which had the highest GIS knowledge and this explains this impact.

5.4.8 Differences by respondent characteristics

5.4.8.1 Organisations

There were no significant patterns between the different organisations in terms of the Phi results. The organisations also did not relate strongly to any specific cluster by Phi (Figure 48) or duration (Figure 49).

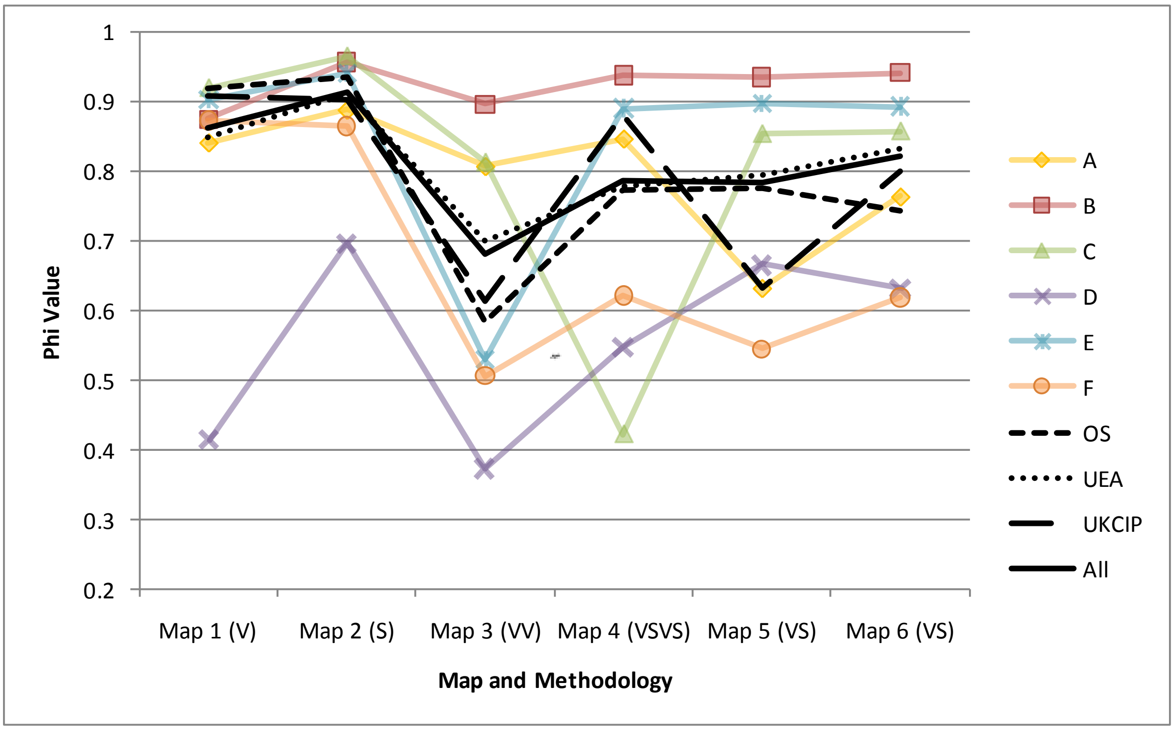


Figure 48. Phi results by cluster with results by organisation and total average (overlaid on Figure 44).

Participants from UKCIP showed a spike in their Phi results in Map 4, which is because of their aural learning style (so they performed more effectively with VSVS than VV) and their lower GI skills (meaning Map 5 (VS) was poorer).

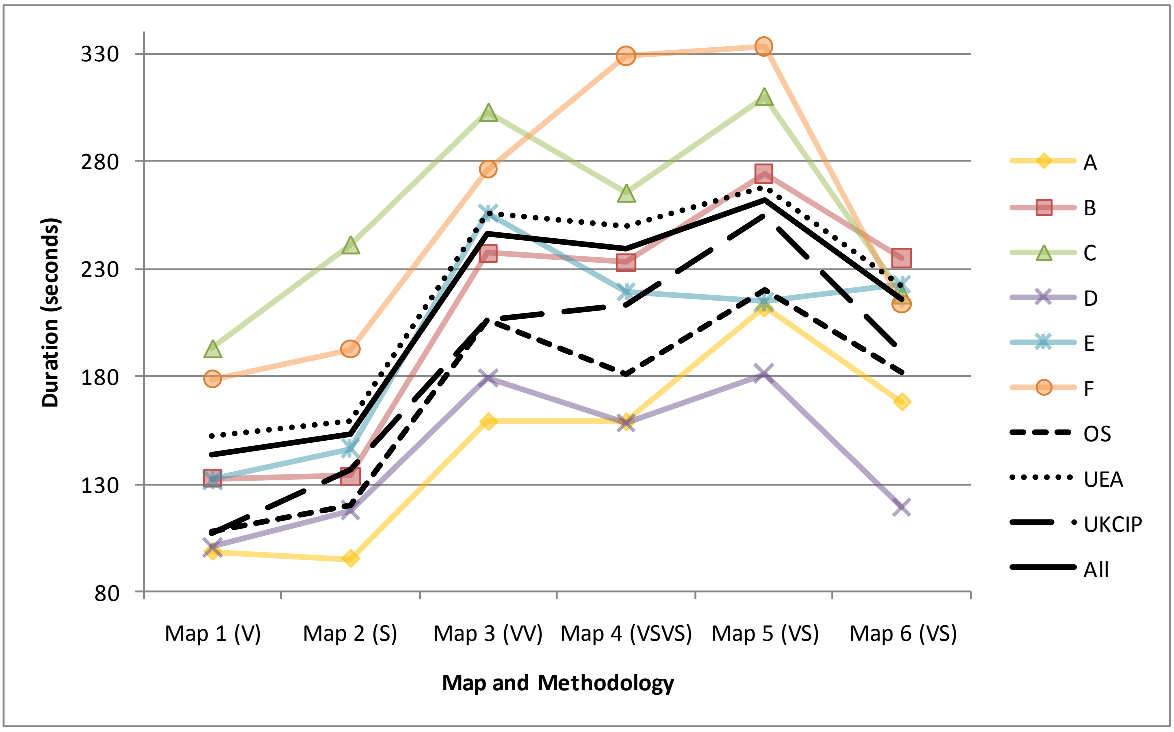


Figure 49. Graph of duration by cluster with organisation and total average (overlaid on Figure 47).

Participants from OS were slightly faster (particularly on Map 4) due to their higher GI experience, but the differences between organisations were not significant.

5.4.8.2 Significant results

The questionnaire element of the evaluation gathered various pieces of information, some of which made a difference to the Phi value. The questions that had an influence on the results are shown in Table 22. See Section 5.3.4 for details and Appendix C.5 for a full copy of the questionnaire.

Question Number	Question	Possible Answers and Notes
Q4 Aware of UKCP09	How long have you been aware of the UKCP09 data set?	The answers from this question were summarised to be either Yes (aware of UKCP09 for any period) or No (not aware of UKCP09 before today)
Q5 Used Projections Website	Have you used the UK Climate Projections website which contains information on what the projections are and how to use them? (http://ukclimateprojections.defra.gov.uk)?	Yes or No
Q6 Attended PIP Training	Have you been to the PiP (Projections in Practice) or any other training organised by UKCIP?	Yes or No
Q7 Used Moodle Training	Have you looked at any of the online training material available at http://moodle.ukcip.org.uk from UKCIP?	Yes or No
Q8 Used User Interface	Have you downloaded any data from the User Interface site (apart from as a part of the training session? http://ukclimateprojections-ui.defra.gov.uk)	Yes to download data for a specific project Yes just to see how to download data No There were a very small number of answers (n = 5) of 'Yes for a specific project' so these were removed from the analysis.
Q9 Aware UKCP09 contains Uncertainty	Before today, were you aware that the UKCP09 data contains information on the probability of the changes of the future climate?	Yes or No
Q12 Probabilistic Knowledge	On a scale of 1-5 (with 1 being high) how would you rate your knowledge of probabilistic data?	1 – 5
Q13 Betting Knowledge	Which statement best describes your understanding of betting?	There were four statements, and the results were grouped together as either high knowledge (1) or low knowledge (0)
Q20B Qualification Area	Which area is this [your current or highest] qualification in?	Climate Change (using UKCP09 data) Climate Change (not using UKCP09 data) Environmental Science Other Area / Non-Area Specific

Table 22. Details of the questions asked in the questionnaire.

A series of t-tests were run on the Phi values, to see whether the answers to these questions had any impact on the results (Table 23). Only Maps 3 to 6 were compared, as these formed the basis for the comparison of the different visual and sonic methods.

		Phi Value				
		Map 3 VV	**Map 4 VSVS**	**Map 5 VS**	**Map 6 VS**	**Mean**
***Q4 Aware of UKCP09**	Yes	.719 (p = .086)	.854 (p = .002)	.842 (p = .002)	.896 (p < .001)	.856 (p < .001)
	No	.632	.703	.710	.730	.747
Q5 Used Projections Website	Yes	.667	.771	.781	.810	.799
	No	.718	.831	.786	.853	.832
Q6 Attended PIP Training	Yes	.685	.779	.790	.824	.808
	No	.618	.877	.688	.776	.801
Q7 Used Moodle Training	Yes	.731	.829	.685	.825	.825
	No	.674	.780	.795	.820	.805
*** Q8 Used User Interface (ignore Yes-See How, n = 5)**	Yes (Specific)	.773	.810	.790	.863	.850
	No	.663	.782	.781	.813	.799
***Q9 Aware UKCP09 contains Uncertainty**	Yes	.703	.863 (p = .001)	.826 (p=.031)	.886 (p < .001)	.849 (p = .001)
	No	.655	.702	.735	.750	.761
Q12 Probabilistic Knowledge	High	.675	.849 (p = .018)	.826 (p = .093)	.861 (p = .055)	.842 (p = .030)
	Low	.683	.742	.753	.793	.783
Q13 Betting Knowledge	High	.664	.840 (p = .027)	.831 (p = .030)	.844	.834 (p = .060)
	Low	.695	.736	.738	.800	.783
Q20B Qualification Area	Climate Change	.770	.809	.720	.826	.830
	Not Climate Change	.664	.782	.794	.820	.803

Table 23. T-test results for highlighted questions.

Awareness of the UKCP09 data set has a significant impact on the Phi score, both for the groups and the results individually ($p < 0.001$) (Q4). Length of awareness (e.g. 'knew about UKCP09 since launch' and 'knew for 6 months') did not appear to influence Phi results, which is why they were combined.

Awareness that the UKCP09 data contained information on the uncertainty of the projections was significant ($p = 0.001$) (Q9 ‘Are you aware that UKCP09 contains uncertainty information?’), but more detailed knowledge about probabilistic data (Q12, Q13) had variable impact, depending on the map stage in question. Use of the UKCP09 (Q5) data and/or attending the PiP training sessions (Q6) appeared to increase the effectiveness of this methodology, but this difference was not significant.

The impact of awareness was significant for the methods that utilised sound, however, the trend was the same for the exclusively visual methods, but it was not significant for all of them. This may be a learning effect, or may mean that participants need to have a reasonable awareness (and understanding) of the data to utilise a sonification.

Higher awareness that UKCP09 contained data on uncertainty and the relevance of this to the participant's work, seemed to make the respondents worse at the exercise, although this could have been because they were thinking about the uncertainty issues rather than the task in hand. Climate change knowledge did improve results and GI knowledge had a much smaller (but still positive) effect.

Learning style seems to have some impact, but the details are unclear. It appears there is an optimal combination between visual and aural preferences, as those who have a high preference for both seem to perform better than those who have a low preference for both. The lower visual learning style is most likely to impact on the ability to utilise spatial data and GIS, which would impact the results (e.g. Cluster D). However, the interaction is probably more complex, as the differences between visual and aural for Clusters A, B, C and F are quite small, but they have very different trends.

The subjective measures (which method did you find easiest / hardest) showed no significant relationship with the objective results. There was a trend for participants who said they found VSVS or VS easier to score more highly on these methods than the other methods, but this was not significant.

5.4.9 Results from group discussion sessions

The group discussion sessions provided some very interesting feedback from the participants including comments on the positives and negatives of the sonification, general use of sound and the design of the experiment.

Peoples' responses to the sonification varied quite significantly; some found it very useful, and some just found it to be a distraction and a negative impact on the evaluation process. Vision provided a good overview of the data, whereas sound was useful for finding information on specific grid cells and some users also used it as a check, to confirm the visual information, if the colours were unclear. Users also commented that using the sound interface to access all the data (as required in VS) was very slow, as they had to move the mouse over each cell individually to hear the sound.

One of the most commented aspects (by around two thirds of the participants) was on the use of two maps to show the data in VV and VSVS. This was when two maps were shown side by side, with the 50th Percentile data on the left and the Range data on the right (see Figure 37 on page 151 for an example of this). People were asked to work out the combined area where the 50th Percentile and Range data exceeded the specified threshold and highlight this on the map on the right hand side. Conceptually, this was a reasonably complex spatial operation and even those with higher GI knowledge said that they had difficulty completing this task. Some participants with low GI knowledge had to ask for more detailed instructions before they could complete the task. The main difficulty was with working out where a specific cell on one map was located on the second map (i.e. co-locating a particular cell). Cells near the coast were relatively easy to locate, but cells further inland were much harder to co-locate, because there were no obvious features nearby. It is not thought that this had a major impact on the study results, because all participants were equally effective, but it did result in the training Maps (1 and 2) performing significantly better than the other maps, as they consisted of a single map rather than two maps. Additionally, the VS methodology (see Figure 38, page 152) showed both of the data sets on one map, so this issue was avoided and may have artificially decreased the other scores, relative to this. However, the VS methodology required the users to understand the 'AND' operation, which probably explains the difference between GI and non-GI users, as the former would be more familiar with this as a spatial operation concept.

The preference for the sounds used varied, with some people saying the sounds were 'ok', but a significant proportion of people said that more 'natural' or 'musical' sounds would have been preferred. The sounds did appear to be much more computerised than intended and this was because of the processing and compression required to play the sounds through the web interface. The majority (around 80%) of people said they did not like the 'high' sounds because they were irritating when used repeatedly and seemed to 'pierce their heads'. A few people did mention that they found the higher sounds easier to distinguish between than the lower sounds, even though they found them annoying. This is likely to be something that varies between users and ideally the user would be able to choose between different sets of sounds. A third of participants (n = 24) had also participated in CS1 previously and some said that they found the sound in CS2 easier to understand than CS1, because it had the sustain element which CS1 lacked. Some of the other participants said that they found the sound in CS1 less annoying than the sound in CS2 because the sound used in CS2 sounded computerised where as the one in CS1 did not.

About half of the participants mentioned that some of the colour combinations were not very easy to tell apart, particularly when the thresholds fell in certain locations on the blue scale. The colours used were based on those used by the UKCP09 online interface, but there were still instances where the differences between the colours were unclear. Using the colours used for UKCIP09 may have avoided some of these issues and would have provided greater consistency with the likely real world usage. If users were interested in a particular threshold and were suitably equipped GI users, they could alter the colours themselves, and just have two colours: those areas that met the criteria and those which did not. The shades of colour on the monitors varied if they were viewed at a different angle. It was not a major problem, but it prevented the users from moving. This could be dealt with by using a different type of monitor which would not suffer from this issue, or using different colours.

A number of participants (generally the more advanced GI users) asked who the target audience was for this sonification, or why sonification was used, when this example could be presented exclusively visually. There are situations where using sound in conjunction with vision to represent a variable (or set of variables) is useful, if the visual sense is already saturated. For example, if the visual sense is taken up with a standard map, something that is readily accepted or understood by a large group of people, it may be better to represent an additional variable with sound, rather than trying to include it in the already predefined visual sense. It could also be useful when visual saturation has occurred, where no more data can be represented visually without obscuring the underlying data. This study is looking at whether the principle of using sound to represent uncertainty data is feasible and this is why the examples used are relatively simple. It is true that both the 50th Percentile and Range values could be represented visually on a map (e.g. colour and hatching or something similar), but adding a sound variable to this would result in a much more complex evaluation task than the one used, which would potentially result in the participants performing very poorly, which would not provide any information about the use of sonification.

In addition, experienced GI users asked why the user could not just 'Select by Attribute' to find the data that met the criteria. Again, this was just an exercise and in a real world situation this would have been a useful approach to take. However, it would not provide information on the spread or distribution of the data, which might be useful to the user.

One participant said that they thought the difference between colours located next to each other on the legend varied between colour pairs (e.g. it is easier to tell the difference between the two lighter blues than the two darker blues in Figure 38), where as the difference between sounds next to each other on the legend was always the same. This participant had more experience of and exposure to music than average, so the fact that the differences between the sounds were the same and the differences between the colours were not, were obvious to him may not have been so clear to the other participants. While an in-depth discussion of colour and music 'spaces' or ‘palettes' is beyond this thesis, it would be an interesting piece of research in itself, particularly relating the palettes of vision and sound together.

While this was not a main focus of the evaluation, participants were asked whether the sonification helped them understand the UKCP09 data set. The sonification helped some people to understand the data, but most of the participants said they abstracted the task very quickly, so they were looking for a change in sound / colour rather than a 'change in uncertainty' as such.

5.4.10 Using the Google Maps API

The GMAPI provided a very usable interface to run the evaluation. This element of the interface was not the main part of the evaluation and participants were not explicitly asked about the interface. However, they generally found it easy to use and none of the participants had issues with using the interface.

5.4.10.1 Technical issues with the Google Maps API

There were a couple of technical issues with using the API, one of which had a minor impact on the analysis. When the mouse was moved over the different areas of the map, a new sound was played if the value underneath the mouse cursor changed. This worked as expected when moving over un-highlighted areas, but when moving over highlighted areas the values beneath the highlight could not be “seen”, so the interface got stuck on the previous sound. Only two participants mentioned this as an issue during the discussion sessions and upon further questioning, they said it did not impact their results.

The other issues were more aesthetic than important and included potential issues with the changeability of the GMAPI code. One change that occurred resulted in the Google Maps Street View 'Pegman' appearing in the interface (see Figure 50, over). This was easily removed, using the posts on the forum asking 'why has the pegman appeared?' but it does highlight the potential problem of relying on a third party service. Completing the survey was reliant on continuous internet connectivity which was disrupted during two of the sessions at UEA. This was a result of a complete network failure at UEA, which could not have been predicted. This did not impact the results (the sessions were rescheduled successfully) but does highlight the potential issue of relying on Internet connectivity or any centralised resource.

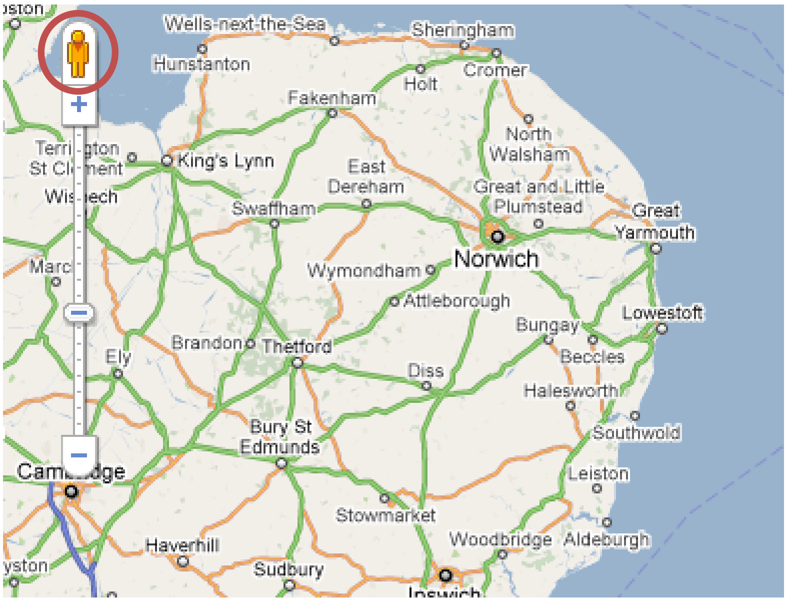


Figure 50. The Google Maps Street View 'Pegman' icon (circled), which appeared suddenly during the evaluation sessions run at UEA. This is the icon used to go from normal Google Maps view into Street View mode.

Additionally, observation suggested that 2-5% of sessions saw map tiles received slowly or not at all from Google's servers; although specific feedback was not solicited, it was felt that this might cause users in unsupervised surveys to abandon sessions. Future applications should allow users to refresh pages without losing responses. Performance is reliant on the remote service and some evaluation sessions had to be rescheduled due to local network disruption, however such problems would affect any online survey.

5.4.10.2 Longevity issues with the Google Maps API

Ideally all academic research should be reproducible, based on the information available in the research publication (Brunsdon, 2011). This is fairly simple process for experiments with standard equipment that will not change over time, but this research used web browsers and the GMAPI, which are changing at a very rapid pace. This is an acknowledged problem given the growing necessity of archiving digital data, and is generally addressed via either emulation of obsolete technical environments or migration of digital objects to a currently accessible format (Brown, 2006). The Internet based nature of this evaluation limits the usefulness of this technique, so to ensure this research is reproducible, screenshots of the interface are included, as well as a video showing the dynamic sound elements. In addition, fully commented computer code and flowcharts of the process involved are available in Appendix C.1.

The GMAPI operates on a series of rolling changes with new minor versions being released every 3 months and each major version only being supported between 9 months and 3 years (Google, 2011a). This timeframe is stated in the documentation, and while RSS feeds and email alerts about new versions are available, they were not widely publicised during the development of this project. Google recommend developing the application with a specific version of the GMAPI and once launched, testing against new versions as they are released every 3 months.

5.5 Conclusion

The participants in the evaluation had a reasonably wide range of knowledge of GIS and climate change, with the majority not having experience of using the UKCP09 data set. Awareness of the data set made a significant difference ($p < 0.001$) to the ability to complete the exercise. Using sound in combination with vision to show the data improved scores for the majority of the participants (71.8%, 56 out of 78).

Results in the literature show that combining sound and vision together or sound, vision and haptic together in interfaces can improve user performance (Jeong and Gluck, 2003; Harding and Souleyrette, 2010) however, the details of the impact of sound specifically are unclear (Constantinescu and Schultz, 2011). It is likely that this variation is due to a non sound factor in the experiment, such as the data being sonified or the background of the participants. Other literature is mentioned below when appropriate in relation to the case study results, but the amount of sonification literature on spatial data is limited.

5.5.1 Original research questions

This research set out to answer these specific research questions:

- Can sound be used to represent the uncertainty in the UKCP09 data set?
- Is using sound to do this more effective than using existing visual methods?
- How does the participants learning style and/or previous knowledge of the data set impact on their ability to use the sonification?
- Is a web based evaluation more or less useful than an evaluation based in ArcGIS (as in CS1)?
- Do the Ordnance Survey participants perform better or worse with UKCP09 data compared to AL2 data (CS1)?

Sound can definitely be used to represent uncertainty in the UKCP09 data set and the method used for this evaluation appears to work effectively. Using sound to reinforce what is shown visually (VSVS) was significantly more effective than the existing visual method tested in the evaluation ($p = 0.006$). Using vision and sound to show two different variables on the same map (VS) was more successful that the traditional visual method, but respondents felt it took much longer to get the information from the map.

However, sound does have its limitations, particularly with the way it was implemented in the VS methodology. Participants said they felt the Range data (represented using sound) took too long to access, as they had to move the mouse over each grid cell individually. In fact the mean and range of durations for VS (Map 5 and 6) were not that different to VSVS (Map 4), and the Phi scores were generally better. There are more effective ways of sonifying this data and looking into the sonification literature showed some simpler examples including using left/right pan to show locations, using a combination of instruments to demonstrate thresholds (Constantinescu and Schultz, 2011) and more complex possibilities including a 'sweeping sound' option to generate an overview (Bujacz et al., 2011). There are also ways to make the sound less tiring on the user, primarily by making use of silence so the sound is not continually playing. A suggestion of using sound to show either an overview of the entire data set, or an overview of a particular area was common and some type of weighted neighbourhood average could be used to provide the input into the sonification element.

Other visual methods could be used to represent this data and they may be more or less effective than the sonic methods used in this experiment. One commonly suggested option is to use colour for the temperature and shading or monochrome adjustment for the uncertainty, but this could be very difficult to understand for a large spatial area (e.g. the UK) or if a continuous colour scale is used. A method such as this was not included in the evaluation, but would make a very interesting piece of comparative research.

As mentioned earlier, previous knowledge of the data set helped participants significantly. Learning style appeared to have a slight impact, where those with a visual learning style performed more effectively, likely due to the mapping element of the exercise, but this was not significant. There have been a number of pieces of research in the sonification community looking at whether musical knowledge / ability to play musical instruments / read music etc. makes a significant impact on peoples abilities to use a sonification (Mauney and Walker, 2007; Constantinescu and Schultz, 2011; Tiuraniemi, 2011). In the vast majority of studies there has been no significant difference and it is the opinion of the community that this does not have an impact (Walker, 2011, pers. comm.). There were definite groups of participants who found the sound useful and groups who did not, but from the data collected in this evaluation it was unclear what factors are involved. More detailed information from the participants may help with this.

There was an element of learning effect within the results (particularly shown in Map 5 and Map 6, using the same method) and no doubt with more experience of the sonification, people's scores would increase. This learning effect could have been removed by randomising the order that the maps were presented to the participant, but this was attempted in the pilot stage and was found to be too confusing for the users.

The first case study used an interface within ArcGIS to sonify the spatial data and collect responses and this case study used the GMAPI in a web based interface to collect the data. The results from this case study generally agree with the results from the first case study, with the VSVS method being the most successful (VS Same in CS1). VS scored reasonably well, but VS Different (CS1) did not, and this difference could be because of the differences in the data used. CS1 showed that knowledge of the data set might be important, and CS2 explored this area and shows that just being aware of the data has a significant impact.

The two case studies used very different methods to collect the data, with CS1 using a custom coded extension to ArcGIS and CS2 using a web based interface with the GMAPI. CS2 took longer and was more complex to code, but was easier to roll out to different users because an installation of ArcGIS was not required. The web based nature also helped access to the geographically dispersed population. The sounds were slightly limited by the choice of sound add in, but this can be addressed by use of a different add-in, as there are a number of other options now available. Additionally, the development of HTML5 adds a significant amount of multimedia capability directly into the web browser (including audio), which could be used for the sound element instead of requiring an add-in (Hickson, 2011).

When running the evaluation, more difficulties were encountered with CS2 than CS1, primarily due to the greater reliance on third parties for the evaluation to work (e.g. Google and Internet connections). However, CS1 evaluation was not problem-free so a balance needs to be struck and what impacts this will depend on how (and where) the survey is being run.

5.5.2 Learning style

The aim of the learning style question was to see whether participants' preferred learning method had a significant impact on their ability to use the sonification interface. There was a range in the participants learning style, but this did not significantly relate to the results. Upon further consideration, it appeared that learning style worked as a measure of GIS and map skill. Those participants who had a visual learning style tended to like diagrams and pictures and it is not a great extension to suggest that they like maps as well. Equally, those who had an aural learning style do not get on well with diagrams and probably do not get on well with maps. There was no significant relationship between learning style and results, but there was a tendency that those with a visual learning style were more likely to have greater GIS knowledge than those with an aural learning style.

5.5.3 Future development

There are a number of areas that could be developed further, particularly with the sonification method and the participant details. The sonification method used is a reasonably basic one compared to some of the others in the literature (Bujacz et al., 2011; Constantinescu and Schultz, 2011; Hermann and Zehe, 2011) that work more in harmony with the users listening experience. This could be developed in future research, along with an increased amount of information being represented sonically (e.g. the whole of the probability distribution function / PDF for each point) if this did not impact user effectiveness. More details about the participants could be gathered to understand if there are any factors that influence the usefulness of sound in this type of situation. As highlighted by one of the participants, there may be some relation between different groups of colours and different groups of sounds (and differences between individual colour pairs and sound pairs), which are intuitive for a wide group of users. If these do exist in this case study it would be purely coincidental, but there is great potential for a future piece of research in this area.

The use of GMAPI is also a relatively new development and has generated significant interest from other academics interested in using the technique. The software available will continually develop, which does raise reproducible research issues, but also will provide new features and innovations (see Section 5.4.10.2 for details).

5.6 Findings in relation to the Research Hypotheses

This section reviews the results from this case study in terms of the original research aims and hypotheses. They will be compared with results from the other case studies in Chapter 7.

5.6.1 Specific research aims

The specific research aims were:

1. What form does uncertainty information need to be in for it to be represented using sound?
2. How effective are different means of representing uncertainty in spatial data?
3. Which types of sound variation best represent spatial data uncertainty to users?

This case study addressed all three of the specific research aims, with a comparison with CS1 important for aims 1 and 3. The data used in this case study had uncertainty in a very different form to that of CS1, which has major impacts on how it can be presented to the user. A different type of musical instrument was used to represent the sound, but there was not time in the evaluation to complete a comprehensive comparison between different types of sound. This case study also compared different methods of representing the uncertainty data (Aim 2) with the use of sound both as an reinforcement and as a way of showing additional information significantly better than vision alone ($p = 0.005$).

5.6.2 Research hypotheses

This section outlines the research hypotheses in reference to the results from this case study, both in terms of the sonification design and evaluation.

H1. The way uncertainty is conceptualised within the spatial data set will impact how it can be represented in a sonification.

The UKCP09 data set contains a very detailed set of data on the uncertainty surrounding the predicted temperature increase for each location in the UK. The level of detail is important when using this data set and forms a crucial part of the data. It also creates a wide range of possibilities for representing the information. The data producers (UKCIP) struggled to show the full detail visually, which makes this data set ideal for sonification. Many different sonification approaches could be used to represent the uncertainty data and it was decided for this study to simplify the uncertainty data before sonifying it. The way the uncertainty is stored within the data had a significant influence on how the data and the uncertainty was represented in the sonification.

H2. The technology available to design, implement and evaluate the sonification of uncertainty in spatial data will impact the findings.

The participants for this case study did not all have access to ArcGIS on their computers and were dispersed around the country, so it was decided to use a more widely available tool to create the sonification and the evaluation. The UKCP09 users were distributed around the country and very rarely came together in the same location so it was decided to create a web-based version of the evaluation that could be completed without an on-site session. The use of a web-based version for the evaluations also allowed it to be completed by the UKCP09 staff who used a series of laptops available in their office with no compatibility issues. The GMAPI was used to present the evaluations and collect the user responses, with a JavaScript add-on providing sound. These choices did influence the choices of sonification available but provided a much more accessible evaluation interface than available using ArcGIS.

H3. The ability of the user to make effective use of the sonification will be effected by a variety of factors including their knowledge and previous experience.

The evaluation recorded a wide variety of pieces of information about the users GIS, UKCP09 and uncertainty experience. While attending the UKCIP training sessions (PiP) and use of the UKCP09 data did not have a significant effect on performance, knowledge of probabilistic data did have a significant impact ($p = 0.03$) with participants with higher levels of knowledge performing more effectively. Awareness of UKCP09 was also significant ($p < 0.001$) as well as awareness that UKCP09 contains uncertainty information ($p = 0.001$). Therefore knowledge of uncertainty and the effects it can were of crucial importance to getting the most out of this sonification. Learning style had no impact.

H4. Providing redundancy (dual-coding) in a task by showing the same information both visually and sonically will improve performance.

H5. Transferring the representation of uncertainty information from the visual to the auditory domain to address visual saturation will reduce cognitive load and therefore improve performance.

Using sound with vision for these data did improve performance significantly ($p = 0.005$). However, there were no significant differences between using sound for redundancy (VSVS) and using sound to show additional information (VS). This is due to the complexity of the task, particularly the sonification method used for VS, which meant that the evaluation took longer to complete. A different type of comparison would be required to see which method of using sound provides better performance.

H6. Preference for one sound from an option of two will result in better performance with the preferred sound than performance with the other sound.

It was not possible to compare a second type of sound in this case study in addition to the other factors tested as this would have extended the evaluation length too much. A proportion of participants did comment that they would have liked different sounds, possibly more ‘natural’ or ‘musical’ than the ones used in this evaluation. The sounds came across as very computerised, which was due to the compression methods used for the web based nature of the evaluation. Within the set of sounds used, participants had a preference for the lower sounds as they found the higher ones piercing and very tiring. This did vary between users and is something where the users would benefit from being able to choose their own sound scheme for the sonification. The participants that took part in CS1 and CS2 were asked to compare the sounds used in each case study. Each had advantages and disadvantages, often occurring because of limitations from the data or evaluation interface used. Participants who preferred the sonification methods performed slightly better than those who did not prefer the sonification methods, but the difference was not significant.

This case study fully addresses hypothesis 1, 2, 3, 4 and 5 and partially addresses hypothesis 6. The user factors element of the research was substantially developed by this case study, establishing that knowledge of the uncertainty element of the data being represented in the sonification (probabilistic data in this case) was an important factor in whether the participant can make effective use of the sonification. This case study is compared in more detail with the other case studies in Chapter 7.

Chapter 6. Case Study 3: Using Sound to Estimate Distance in Landscape Visualisations

6.1 Introduction

This case study develops and builds on the research completed in CS1 and CS2, as well as applying the concept of sonification in a different setting. This case study evaluates how sonification can be used in a virtual reality / landscape visualisation setting. Currently landscape visualisations are essentially visual tools and the option of using sound to either provide greater immersion or provide additional information may make them significantly more useful.

This case study evaluates the research aims and hypothesis identified in the literature review chapter and summarises the findings obtained at the end of the chapter. In addition it will also consider these specific research questions:

- Can sound be used to represent additional data in a virtual reality / landscape visualisation setting?
- Is using sound to do this more effective than using existing visual methods?
- How does the participant's learning style and / or previous knowledge of the data set impact on their ability to use the sonification?
- How does using an audience response system to collect evaluation responses compare to the methods used in CS1 and CS2?
- Do OS and UKCIP respondents perform better or worse with this data set compared to their 'expert' data set? (i.e. AL2 for OS and UKCP09 for UKCIP).

The research questions were developed from the ones originally stated at the beginning of the thesis based on results and subsequent reflection from CS1 and CS2. Specifically, the learning style and data set knowledge elements of the questionnaire were extended, as well as the evaluation method (using an audience response system).

This case study builds on the previous two case studies, using similar questions as a starting point and a similar group of participants. In addition, a number of people involved in local authority planning were approached to participate, with volunteers from Norfolk County Council and the Broads Authority taking part. The case study used a similar combination of quantitative and qualitative questions as before, but used a different evaluation method to collect the data. The questions were shown as a series of PowerPoint slides, with participants selecting their answers using the keypads provided. This chapter finishes by reviewing the results from this case study, before leading on to a discussion and conclusion of the findings for all three case studies.

6.2 Literature Review

The previous two case studies covered the use of sonification with two different data sets in a computer based GIS environment. This case study considers the use of sonification in a different environment: a landscape visualisation. Spatial data are very important in this environment as they form the basis for the visualisation and sonification can assist with the representation of this data, but in different ways to the previous examples.

A landscape visualisation setting was chosen for this case study because it provides a different, but related setting for the sonification. Uncertainty is still relevant to the visualisation, but there are additional factors that could be represented using sound. The example chosen for this exercise was distance because many people have difficulty estimating distance in a visualisation and using sound to show this may help communicate distance. Distance is also a very important aspect in the planning process, for which landscape visualisations are often used to assess the visual impact of the development. Landscape visualisations are often set within a VR (virtual reality) theatre.

6.2.1 Use of sound within a virtual reality setting

A significant amount of research has been done into the use of virtual reality (VR) for different applications, and there is a compelling argument that sound should be included in these VR environments to increase the realism of the environment. The term 'soundscape' partly encapsulates this idea, with a soundscape being the sounds heard at a particular location at a particular time. Soundscapes have had a reasonable amount of research completed on them (Schafer, 1994), but little on using soundscapes (or sound more generally) as an integral part of a VR environment (Caquard et al., 2008). Carles et al. (1999) completed some work evaluating how different sounds impacted the values people placed on different landscapes and he found that for some landscapes, sound is much more important than others, particularly for those landscapes where a congruent sound is played. This is also seen in the application of serious games, where provision of sonic stimuli in combination with visual stimuli increases the feeling of immersion in the game (Morgan et al., 2012). Looking at the effect of sound on the immersion level within a VR landscape visualisation would be a very interesting piece of research, but is beyond the scope of this thesis. Additionally, incorporating realistic sound in a VR environment is very challenging and faces many problems including synchronisation of sound with visual elements and dynamic correction of sound when moving through the environment (Manyoky et al., 2012).

In this case study sound was used to show an additional dataset to users in a VR setting in a similar manner to the previous case studies, using the example of a landscape visualisation. There has been very limited research in this area, particularly using sound to display additional data to the user (see Section 2.4.6 for more details). Therefore this case study builds on the data collected from the first two case studies and translate this from the computer based GIS setting to a landscape visualisation setting.

There are a number of different variables in the environment that could be helpful to users. Many could be displayed as an overlay on the visualisation for the relevant area (e.g. scale, distance, coverage, membership, area boundaries, etc.), but these could obscure the underlying visual information. Sound is an obvious sense to use to attempt to display some of this information, for the reasons mentioned earlier in this thesis.

A landscape visualisation can have many different objects within it and may represent the current situation of a location or some possible future, given a particular set of circumstances. The level of detail included in a particular visualisation will vary, depending on the software used, the developer creating the visualisation, the target area and the target audience, as well as many other variables.

Choosing an appropriate level of realism is very important, as realism is one of the main factors that the general public will observe and comment on when looking at a visualisation (Appleton and Lovett, 2005). It is also important to be very clear about the purpose of a visualisation and to say whether it is showing a proposed new structure on an existing landscape or a potential future landscape that may or may not happen. While no visualisation can be truly 100% complete because of the limits in geospatial information capture and VR information representation, for a visualisation of a current location at the current time there are many pieces of information that could be included in the visualisation, derived from on-site measurements and photographs. Interaction in a landscape visualisation is also important in terms of giving the user a realistic impression of the landscape (Lange, 2005). Future landscape visualisation is very different, with a much more limited amount of information available. For example, the future base data may only provide the proportion of a 5km grid square that has changed from one land cover to another without providing any information on where this change has taken place (Munday et al., 2010). The process of getting from this to a landscape visualisation is very complex. The visualisation designer could either add in more detailed information, based on their skills and experience (but this would not represent how the landscape will look, or the limitations of the data available) or just include the available information (but this could provide a relatively low quality visualisation which may confuse and/or discourage viewers). This ‘lack of information’ is something that is not well communicated in visualisations and can be a very difficult concept for the author to communicate as well as for the audience to understand (Sheppard and Cizek, 2009).

6.2.2 Perception of distance within a landscape visualisation

There are a number of elements within a landscape visualisation that have uncertainty that could be presented using sound, particularly with a future landscape visualisation. In this case, each element of the landscape (e.g. field, river, wind turbine, solar panels, trees, etc.) will have a different measure of the likelihood of it being present in this future landscape. However, the types of details required for such a visualisation are often not included in the data (Munday et al., 2010) and even if they were, it is unlikely that the uncertainty information would be included explicitly. Distance is something that many users have trouble estimating in visualisations and this is often difficult to estimate in the real world. Using sound to provide additional information for distance may help users to estimate the distance of different objects.

Estimating relative distance (i.e. is tree A closer to us or further from us than tree B?) is moderately easy for most users. However, estimating absolute distance (i.e. how many meters away is tree A?) is much more difficult. Estimating distance in the real world is often quite difficult for some people and this is equally the case in a landscape visualisation (Montello, 1991; Plumert et al., 2004). Many studies have reported that participants underestimate ego-centric distances in Head Mounted Display virtual environments (Zhang et al., 2012) and a similar observation has been seen in large screen VR displays, similar to the one used in this case study (Klein et al., 2009; Grechkin et al., 2010). This includes both estimating distances by moving through the landscape and verbal estimations of distance without moving (Alexandrova et al., 2010) as well as using visual and audio stimuli (Rébillat et al., 2012). Methods of assisting users to estimate distance in these environments are therefore likely to be useful.

Landscape visualisations are increasingly used to attempt to assess the amount of impact a new development will have on an existing area (Sheppard, 2001; Paar, 2006). Wind turbine developments require planning permission and around half of applications in 2010 in the UK were refused (Milne, 2011; Renewable UK, 2011), often on the basis of visual impact (Business Green, 2011). The UK Government’s Renewable Energy Strategy sets a target of generating 15% of the UK’s energy from renewable sources by 2020 and the vast majority of this will come from wind turbines (Renewable UK, 2010). This means that the emphasis on planning applications for these types of developments will only increase with time and the perception of distance within these visualisations is a very important aspect in understanding and applying these visualisations.

A significant amount of research has been conducted on the visualisation of wind turbines within a specific landscape setting, looking at how well the turbines can be seen at different distances and how effective different visualisation methods are. Bishop (2003) gives a good overview of how wind turbines are perceived in landscapes and also Bishop and Stock (2010) evaluate a method of using a collaborative virtual environment to assist with the planning process.

Evaluation of the visual impact of turbines at different distances is known to vary significantly (Bishop and Miller, 2007) so people's ability to estimate distances in landscape visualisations is important. It has also been shown that turbines at a distance of > 20km have little or no visual impact and recognition of turbines at a distance of 10km occurs in around 1 in 5 people. The furthest distance at which visual impact is significant is 5-7km (Bishop, 2002; University of Newcastle, 2002; Benson, 2005; Jallouli and Moreau, 2009), so any evaluation involving turbines must have the turbines closer than 5-7km. Additionally, the depth of a visualisation has a significant impact on participants opinion of the landscape (Bishop et al., 2000) and minor changes in the landscape (such as adding or removing a few trees) can have a significant impact on participants perception (Sang et al., 2008).

Wind turbines were chosen for this case study because they are an extensively studied feature in landscape visualisation. The justification for the number of wind turbines and distance to them is presented in the methodology section, as is the amount of information made available to the users about the scale of the visualisation and the precise wording of the question.

6.3 Methodology

This case study looks at how sound can be used to represent data in a landscape visualisation setting. This is different from the two previous case studies and it is important to see whether the methods of sonification evaluated in a computer based GIS setting are also effective with landscape visualisations in a VR setting.

The evaluation followed a similar pattern to the first two case studies, with participants from UEA, OS and UKCIP and a computer based evaluation followed by a group discussion. In addition, a group of planners were invited to take part, from Norfolk County Council, a number of district councils and other related organisations in Norfolk. The participants were shown several landscape visualizations with a number of wind turbines in them and asked to judge how many of the turbines were within a certain distance. Responses were collected using an audience response system, where each participant selected their responses using a keypad.

The evaluation was run in small groups, and usually took around 1 hour to complete. Participants were not paid for their time. The main section of the task was viewing the visualisations and answering the associated questions. Both the visualisations and questions were shown to the participants using Microsoft PowerPoint 2007.

6.3.1 Visualisations

The visualisations used for the task were based on a series of wind turbines placed within a fictitious landscape. A fictitious landscape was used to avoid the potential of participants having different levels of familiarity with the landscape. The arrangement of the turbines was varied to prevent participants from remembering where the turbines were (and therefore the correct answer) from question to question. The same landscape was used for all of the questions, as any changes in the landscape would affect participant's ability to judge distance, and may alter the results (Shang and Bishop, 2000). There are a number of software packages available that could be used to create the required visualisations, but due to limitations in available time, it was decided to use Visual Nature Studio 2.85 (3D Nature, 2011), which is a piece of software used by colleagues within UEA. The use of the software at UEA allowed a rapid learning of the required skills and easily available advice. Visual Nature Studio is regarded as state of the art software for landscape visualisation and has been used by a wide range of organisations for visualisations including academic research (Appleton and Lovett, 2003; Paar, 2006; Wang et al., 2006).

The software used to create the landscape visualisations has the ability to generate the landscape shape (in the form of a DEM) and this was used to create the fictitious landscape. The height was set to range between 0m and 500m and the landscape was intended to look roughly similar to the South Downs, as this would provide a rolling landscape within which to set the turbines (see Figure 51). Flat landscapes can make distance judgment much more difficult due to the limited reference points and very hilly landscapes could potentially obscure wind turbines placed upon them. Very hilly landscapes would also limit the realistic placement of turbines as the turbines would only be placed at the top of hills, where the wind is the greatest.

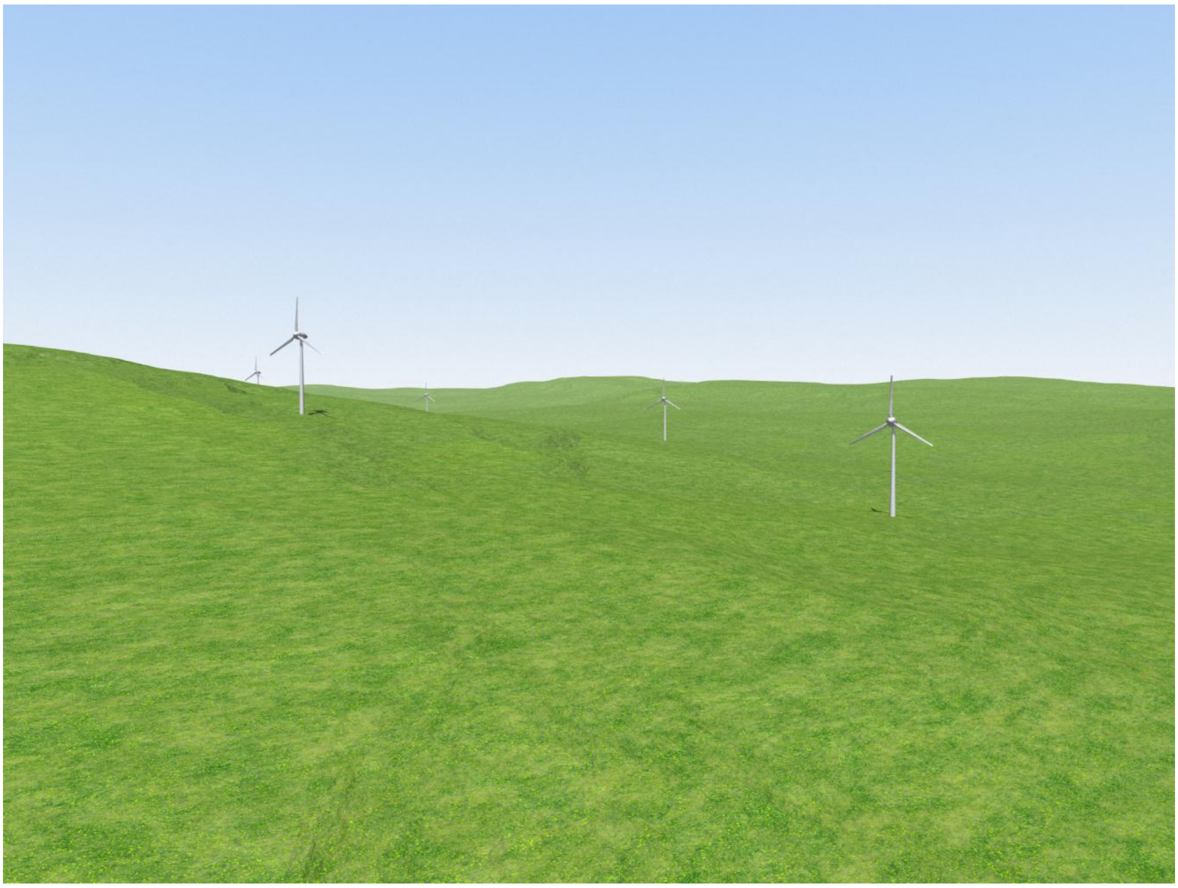


Figure 51. An example of the visualisation show to the participants, with five wind turbines.

When generating the visualisations, it was decided to have a field of view of 60°. This is because this is the same field of view as used in a standard 35mm photo taken using a wide angle lens (Sevenant and Antrop, 2011). Altering the field view has a large impact on the image so this is a very important aspect of the visualisation.

Choosing the position and number of turbines is important for the evaluation to work correctly. The first step was to generate the landscape; for this case study a generic landscape was developed. It was not based on a real world location or designed to be representative of anywhere. This was done so none of the participants would be able to locate it, as it is anecdotally observed that if participants know the location it can affect their judgement of the visualisation (i.e. they often spend time thinking about how the visualisation differs from their experience of the landscape, rather than focusing on the task). The turbine design used was a pre-created model, based on the turbines located at Swaffham Ecotech Centre from previous research at UEA. The turbine had a total height of 100m and is one of the larger land based turbine designs (Ecotricity, 2011).

6.3.2 Pilot testing

It has already been noted that the turbines should be within 5km to be easily seen and a number of different distances and a number of turbine combinations were experimented with. The number of turbines and distances were initially chosen by the author and then pilot tested by three volunteers. The numbers and distances were tweaked based on their responses. The final set of visualisations consisted of five turbines, at distances between 800m and 3000m from the viewer. Twenty different visualisations were created, with turbines at different distances from the viewer and locations within the image.

During the pilot phase, a number of different distance judgment questions were piloted. It was subsequently decided to ask the participants 'how many turbines are within 2000m of your current location?'. The 2km threshold was decided upon from a combination of the literature suggesting 5-7km as the maximum distance for perception of wind turbines, the visual appearance of the turbines within the visualisation and feedback from the pilot studies. It was also apparent from the pilot that the 2km threshold would give a variety of correct and incorrect answers, avoiding floor and ceiling effects discussed previously.

6.3.3 Audience response system

The previous case studies collected responses on an individual basis, with each participant using a computer on a one-to-one basis to provide their responses. For this case study it was decided to run the sessions as a group and collect responses as a group to avoid some of the issues that could occur with individual data collection. For example, some participants were waiting around for up to 15-20 minutes in the first case studies while waiting for the other participants to finish, before the group discussion session could begin. Running the session as a group avoided this.

The visualisations were shown on a projector and the participants responses were collected using an audience response system. These consist of a series of hand held keypads with the numbers 0 to 9 on them, one given to each participant (see Figure 52) and an associated USB dongle to plug into the system running the presentation. The questions appear on screen and the participants select their answer. The system used was TurningPoint (Reivo, 2011) and the software element of the system integrates with Microsoft PowerPoint 2007, allowing questions to be inserted and the results to be extracted. Audience response systems have been well researched as a teaching tool (Cranston and Lock, 2010; Efstathiou and Bailey, 2011) and have been shown to be effective research tools (Ulla, 2005).

Figure 52. The response keypad for the TurningPoint system (image from http://www.bgsu.edu/images/provost/img49376.gif, see also Reivo (2011)).

6.3.4 Group setting

The evaluations were run in group settings, with between 2 and 8 participants. The questions and tasks were shown on an overhead projector, with the participants answering the questions using the audience response system. For the evaluation questions relating to the visualisations (i.e. how many turbines are within 2000m?), the software allowed the speed of response to be recorded as well, which was analysed in conjunction with responses and participants characteristics.

The group evaluations at UEA took place in the VR theatre in the ZICER building using the display technologies available there (see Figure 53). At the other locations a standard projector and set of speakers were used to present the evaluation. A test slide was used at each location to check the display ability of each projector to ensure the visualisations would be presented consistently to all of the participants. The audio was also tested to ensure a constant level and during the practice phase, it was confirmed with participants that they could hear the sound clearly.

Figure 53. The Virtual Reality theatre at UEA where the evaluation for CS3 was held.

6.3.5 Sounds used

The sounds used for the evaluation were chosen based on feedback received from the first two case studies. In addition to piano notes, a more natural sound was chosen for this evaluation as well. This was because in previous case studies participants said they would prefer a more natural sound than the musical notes, as it would be easier to listen to for extended periods.

Initial consideration was given to bird calls of different species of birds, which would range from very beautiful (e.g. nightingale) to very harsh (e.g. gull). While there is a scale of sorts to these calls, upon further investigation it was found that many bird calls would be considered either harsh or beautiful, but there would be very few in between. A ratio scale of sounds was needed to represent distance in this case study, so the scale of different species of bird calls was not appropriate. Instead a series of 'bird calls' that varied in pitch were chosen from the same MIDI source as the piano and trumpet notes used previously (Oliveira, 2008). For both the piano notes and bird calls, a series of 9 sound clips were recorded and used in the evaluation, with lower notes representing further distances. The highest piano note (or bird call) represented a distance of 333m and the lowest a distance of 3000m. The polarity of this mapping was tested during the pilot phase, and found to be the preferred mapping by the majority of participants.

MIDI is a technology that is used to play predefined synthesised sounds on a computer, as well as more extensive communication between different music instruments and technology. The MIDI specification includes a list of instruments (of which #1 Piano and #124 Bird Tweet are used in this case study), but does not specify the sound samples to be used (MIDI Manufacturers Association, 2011). This is decided by the manufacturer of the relevant device. For this research, the sounds were generated by a computer (the authors' university PC). These were then recorded as a WAV file for reproduction in the evaluation. The sound card (which produces the MIDI output) is 'SoundMAX Integrated Digital HD Audio' and the manufacturer, SoundMAX, do not provide any additional information on the MIDI sounds generated (SoundMAX, 2011).

The evaluation was created within PowerPoint 2007, with Turning Point integration used for data collection. The landscapes were shown as a series of slides, with the sounds attached to the objects within PowerPoint. As no programming was required this mechanism allowed for a significantly faster development of the evaluative prototype than in the previous two case studies. A demonstration of the evaluation (including both the sounds and the PowerPoint slides) is included in the DVD (see Appendix D for details).

6.3.6 Visualisation methods

As well as comparing two different sounds (piano notes and bird calls, as discussed above) a number of different representation methods were compared. These were generally based on the ones used in the previous two case studies, providing a baseline visual only, as well as vision & sound, sound only and vision with scale. For each method, participants were asked to judge how many turbines were within 2000m. Examples of the methods used implemented in a PowerPoint file are available on the DVD, see Appendix A for details. The methods used were:

- Visual Only
  The turbines were shown visually on screen and participants were asked to judge how many were within 2000m. This was designed to be a fairly difficult exercise (as there were no other objects in the landscape to help the user judge scale), but the users were told the height of the turbines (100m). The turbines were highlighted in turn in exactly the same manner as the sound method (see below) to ensure the turbines were on screen for the same time.
- Visual & Sound
  Turbines were shown visually on screen and highlighted in turn, but when each was highlighted the relevant sound was also played. The sounds were three seconds long and a gap was left between each turbine. Participants were played two reference sounds (1000m and 3000m) before each slide.
- Sound Only
  The users were just played the sounds for the turbines in turn, without being able to see the visualisation. Instead, a black screen was shown.

- Vision with Scale
  The turbines were shown visually on screen (the same as Visual Only). The turbine at a distance of 1000m was labelled on screen to give the participant a sense of scale and this label stayed visible while the other turbines were highlighted.

The Sound Only method was included to see how well users could understand the sounds without the visual stimuli. In CS1 the Sound Only option performed very poorly, but this data set and task were very different, so there was interest in seeing whether the sound only sonification would work better.

The order of the methods shown to the users was very important and in pilot testing it was found that if the Vision with Scale method was shown before any of the others, then participants used information from that to work out the distances of the turbines in the other methods. The order chosen was one that gradually increased the amount of distance information available to the participants so earlier methods could not impact later ones. Ideally in evaluations the order of methods would be randomised to avoid any potential learning effects. As explained above, this was not possible.

Also during the pilot study the number of repetitions of the reference sounds was varied between different methods. Some repetition for each repeat was required because participants would forget how the sounds related to distances after one or two repetitions. When a particular method or sound was introduced, the reference sounds were repeated three times and for each repeat of the same method, the sounds were repeated once before the visualisation was shown. It was ensured that all participants heard the same number of repetitions. Feedback from the pilot sessions said this was the best compromise and the approach was used for all of the evaluation sessions.

When the visualisations were shown to the users, each turbine was highlighted in turn (working from left to right on screen) and the sound played, if relevant. A two second gap was left between each sound to avoid overloading the user. An initial pilot study established that 18 visualisations could be completed within the one hour session, including time for background multiple choice questions at the beginning and a group discussion session at the end.

6.3.7 Questions

As well as the visualisations, a number of questions were asked of the participants. A full list is available in Appendix D.2 and the questions included some background demographics, learning style and opinions on the sounds and visualisation methods used. The learning style questions were a subset of those asked in CS2, covering the visual and aural learning styles. As well as the question about the number of turbines within 2000m, participants were also asked to rate how confident they were that their answer was correct on a scale of 1 (high) - 7 (low) to see whether this had an impact on the likelihood of getting the correct answer.

6.3.8 Participants

The groups of participants recruited for this evaluation were based on the groups used for the previous two case studies (total n = 45). Ordnance Survey and UK Climate Impacts Programme groups were similar to those used in CS1 and CS2. Gatekeepers were used to access these organisations and the groups primarily consisted of people who had been involved in the previous case studies but there were also some participants who had not been involved in the previous case studies. The UEA group was similar to that in previous case studies, with the bulk of the participants from ENV (School of Environmental Sciences). The group from Norfolk County Council included a number of planners who have been involved with other research projects at the UEA as well as some of their colleagues. The evaluations were carried out in September 2011, with a Briefing Document emailed to participants at least 24 hours before their session (see Appendix D.3) and a Participant Information Sheet and Consent Form given to them on arrival (see Appendix D.4).

6.4 Results and Discussion

This section provides an overview of the results from this case study, starting with the methods used to process the results, the basic observed values and the participant characteristics. Results of a multiple regression analysis are then presented, followed by evaluation of the output from the discussion sessions and how this related to the rest of the results. The findings from this case study will also be set in context with the wider aims and objectives of this thesis.

6.4.1 Participant characteristics

Table 24 (over) shows information about the participants involved in this case study. Over half of the participants were from the UEA (55.6%) with the rest split between OS, UKCIP and planners. The results did not vary between different organisations, which is consistent with the expected outcome. Knowledge about GIS was split relatively evenly, but knowledge about landscape visualisation and planning applications was heavily biased towards to the lower knowledge end. This meant it was quite difficult to conclude anything from this data, but a more even distribution of knowledge would be ideal for future research.

Learning style will be discussed in more detail later but response was generally fairly similar between all the statements (except 'You can play a musical instrument') with most people either saying it is partially or very much like me. The musical instrument question had a more bi-polar distribution, with people either being able to play an instrument (i.e. very much like me) or not (i.e. nothing like me). There was also very little internal consistency between the visual measures and the auditory ones. There was a reasonably even split between those who were or were not involved in previous sonification evaluations, and those who were scored slightly worse than those who were not, but the difference was not significant.

Organisation	Ordnance Survey	UEA	UKCIP	Planners
	17.8% (8)	55.6% (25)	11.1% (5)	15.6% (7)

Question How would you rate your:	1 (A lot)	2	3	4	5 (Nothing)
GIS / Spatial Data Knowledge	24.4% (11)	15.6 (7)	40.0% (18)	13.3% (6)	6.7% (3)
Landscape Visualisation Knowledge	6.7% (3)	6.7% (3)	28.9% (13)	26.7% (12)	31.1% (14)
Experience of Planning Applications involving wind turbines	4.4% (2)	8.9% (4)	26.7% (12)	15.6% (7)	44.4% (20)

Learning Style Statement	1 (Nothing Like Me)	2 (Partially Like Me)	3 (Very Much Like Me)
You can easily visualise objects, buildings, situations etc. from plans or descriptions.	8.9% (4)	35.6% (16)	55.6% (25)
You pay attention to the sounds of various things. You can tell the difference between instruments, cars or aircraft, based on their sound.	13.3% (6)	42.2% (19)	44.4% (20)
You can play a musical instrument or you can sing on (or close to) key.	42.2% (19)	15.6% (7)	42.2% (19)
You use diagrams and scribbles to communicate ideas and concepts. You love whiteboards and colour pens).	4.4% (2)	33.3% (15)	62.2% (28)
You are a tinkerer. You like pulling things apart, and they usually go back together OK. You can easily follow instructions represented in diagrams.	11.1% (5)	44.4% (20)	44.4% (20)
Music evokes strong emotions and images as you listen to it. Music is prominent in your recall of memories.	6.7% (3)	33.3% (15)	60.0% (27)

	Yes	No
Have you been involved in my previous work?	57.8% (26)	42.2% (19)

Table 24. Overview of the participant characteristics, including learning style and organisation.

6.4.2 Processing the results

During the evaluation responses were recorded using the audience participant system, as explained in the methodology (see Section 6.3.3, page 192). The responses recorded were the number of turbines the participant thought were within 2000m, as well as various additional pieces of information. The length of time the participants took to respond was also recorded. The 'Number of Turbines' responses were processed to assess whether the correct response was given for each visualisation by subtracting the correct answer from the participants' responses. Therefore a value of 0 was a correct answer, a negative answer meant the participant underestimated the number of turbines within 2000m and a positive answer meant the participant over estimated the number of turbines within 2000m. These are the answers used in the subsequent results section.

6.4.3 Practice questions

The first visualisation the participants were shown was a practice question, in this they were shown a visualisation similar to the main visualisations and then asked 'How many turbines are easily visible?' This was to give the participants some practice at looking at the visualisations and wind turbines to prepare them for the later questions. There was a range in the responses given (from 2 to 5) with the majority choosing 5 (see Table 25).

Number of Turbines Easily Visible	Mean Correct Answer	Number of Responses
2	2.99	5
3	2.97	4
4	3.08	10
5	3.34	26

Table 25. Responses for the question 'How many turbines are easily visible?'

Over half of the respondents said they could easily see five turbines and the mean correct answer from the rest of the evaluation (as described in Section 6.4.2) for these participants was higher. This might indicate a different level of engagement but the difference was not significant and there was no difference in the time taken or confidence during the evaluation. This indicator was included in the regression, as it could have had a significant impact, but it did not actually relate to the other factors involved. It is likely to be worth having some measure of engagement in future research, to see whether people who are less engaged perform less effectively in the results.

6.4.4 Basic observations

The scores varied between methods (see Figure 54) with the most effective method in the evaluation being Sound Only (Piano), which was significantly better than any of the other methods ($p < 0.001$) (see Table 26). Vision & Sound (Piano) was the next most effective method, at a similar level to Vision & Sound (Bird) and Vision with Scale. Sound Only (Bird Call) was slightly worse, with Visual Only being the worst performer. This is consistent with the output from the discussion session, with piano being the most useful sound, followed by bird calls and Visual Only being the hardest.

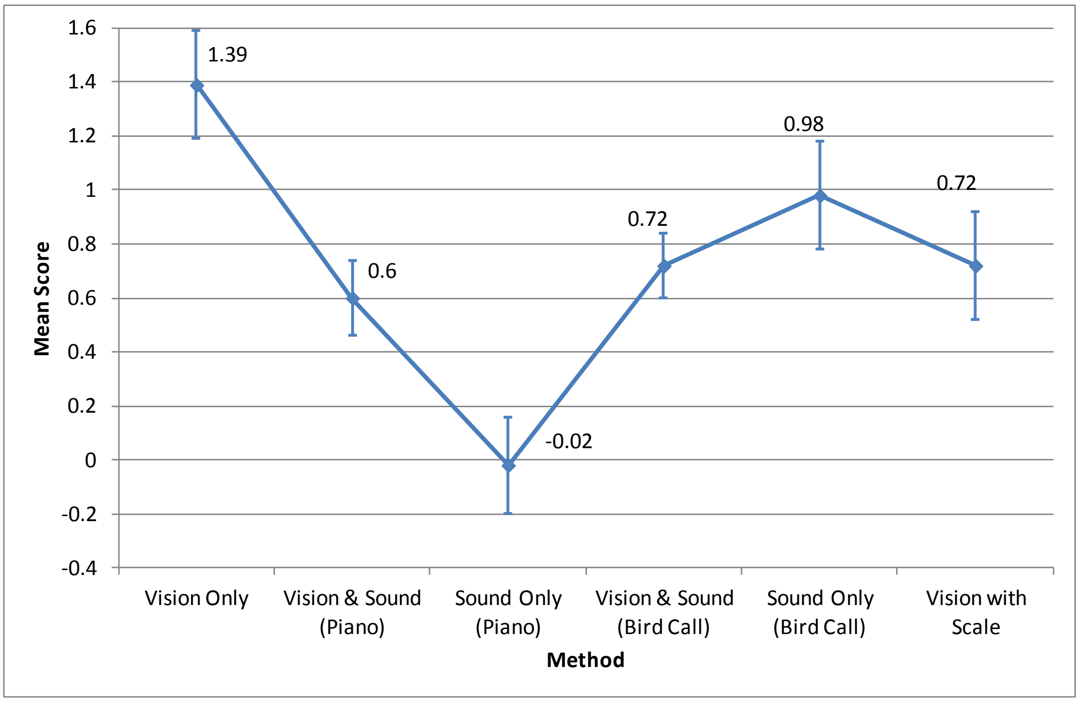


Figure 54. Mean scores of all results for each method, with error bars (2 x standard error). The results are not bi-modal and this graph is an accurate representation of them (see also Figure 55). Sound Only (Piano) was the best result, and Visual Only the worst.

		Vision Only (no scale)	Vision & Sound (Piano)	Sound Only (Piano)	Vision & Sound (Bird)	Sound Only (Bird)	Vision with Scale
1	**Vision Only (no scale)**						
2	**Vision & Sound (Piano)**	p<0.001*					
3	**Sound Only (Piano)**	p<0.001*	p<0.001*				
4	**Vision & Sound (Bird)**	p<0.001*	p=0.188	p<0.001*			
5	**Sound Only (Bird)**	p=0.006*	p=0.002*	p<0.001*	p=0.022*		
6	**Vision with Scale**	p<0.001*	p=0.969	p<0.001*	p=0.278	p=0.006*	

Table 26. T-test results for the 6 different methods. Starred figures are significant*.

Vision Only is significantly worse ($p < 0.001$) than all other methods. Sound (all) and Vision with Scale are broadly similar. Piano performs significantly better than bird calls ($p < 0.001$). Overall, sound with vision performs worse than sound alone ($p = 0.028$), but Sound Only (Piano) is better than Sound & Vision (Piano) ($p < 0.001$) and Sound Only (Bird) is worse than Vision & Sound (Bird) ($p = 0.022$).

It was expected that using vision and sound would be more effective than using sound alone, but this was not the case with piano notes. Participants said that they felt there was a contradiction between the information received from the two different senses, specifically with the distance between the turbines. The visual distance appeared to be greater than the distance between the sounds for the turbines. This is because of the difference in visual perception over different distances, specifically further distances which appeared not be taken into account by the software used for the visualisation.

Figure 55 (over) shows how it was very common for participants to overestimate the number of turbines within 2000m, although this was much more common in the methods using the bird calls than the piano notes. It also shows the distribution of the responses for each method. The difference between piano notes and bird calls is probably because participants found bird calls much less easy to interpret, so they also relied on the vision element to help their answer. Sound Only (Piano) was more successful as participants were more used to listening to tones and some said that the sound contradicted what they saw visually. If the correct sound is used, then it can definitely make the representation more effective, but if the wrong sound is used (or if the right sound is used incorrectly) then it can make performance worse than an alternative visual method.

Figure 56 shows how the individual participants' answers varied over the 18 visualisations. Participants' responses varied over the 18 individual visualisations, with most of the answers between +2 and -1 (94.9%) with 90.6% between +2 and 0. People were much more likely to overestimate the number of turbines than underestimate them, although this did vary with the sonification used (Figure 55). There was a general trend of improvement within each group, particularly Sound & Vision (Bird) (no. 11 to 14 in Figure 56). For Sound Only (Piano), participants were much more likely to underestimate the number of turbines, and for Sound Only (Bird) participants were much more likely to overestimate the number of turbines.

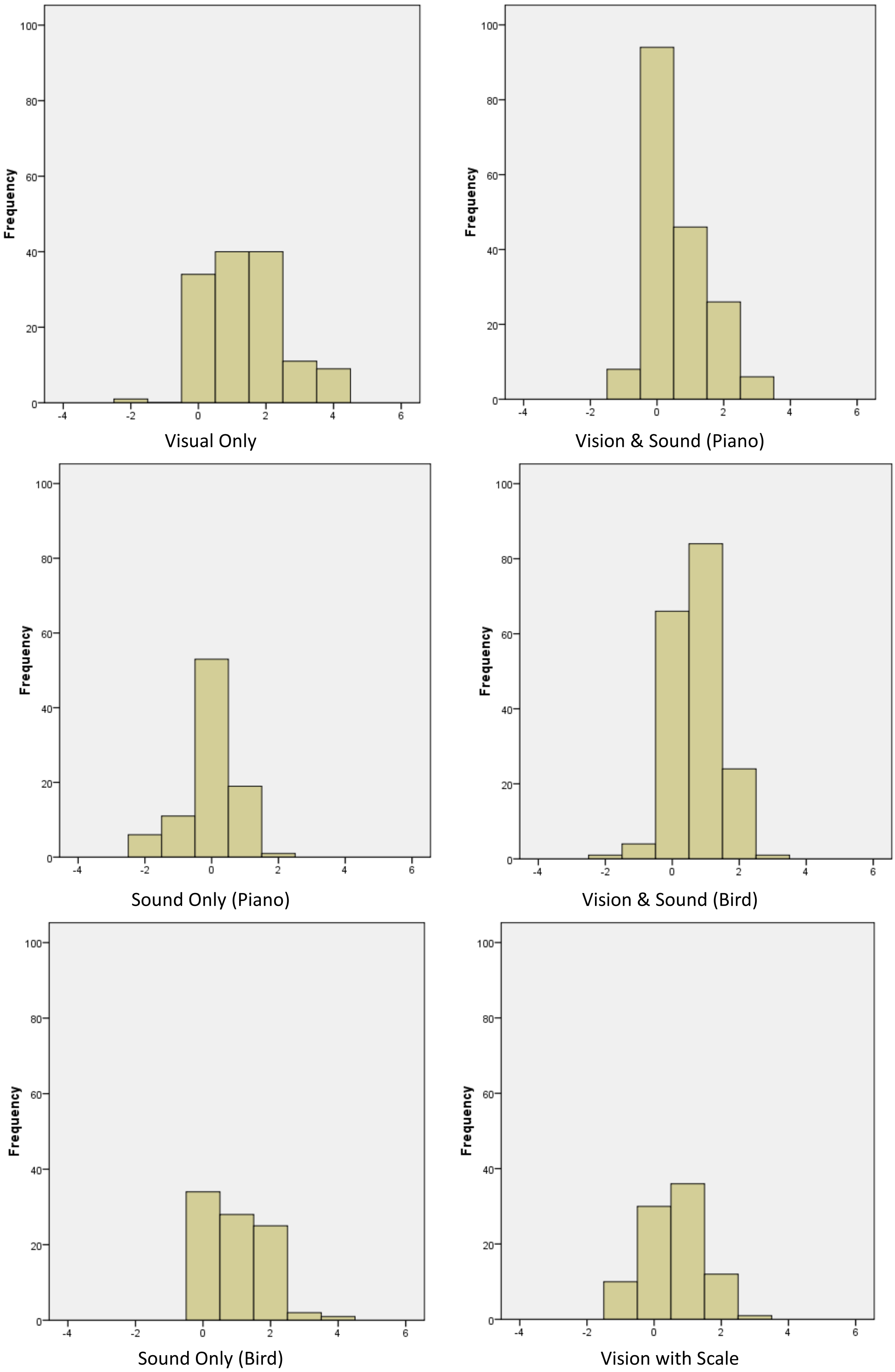


Figure 55. Histograms of the responses for each visualisation method. Note how overestimating the number of turbines is very common.

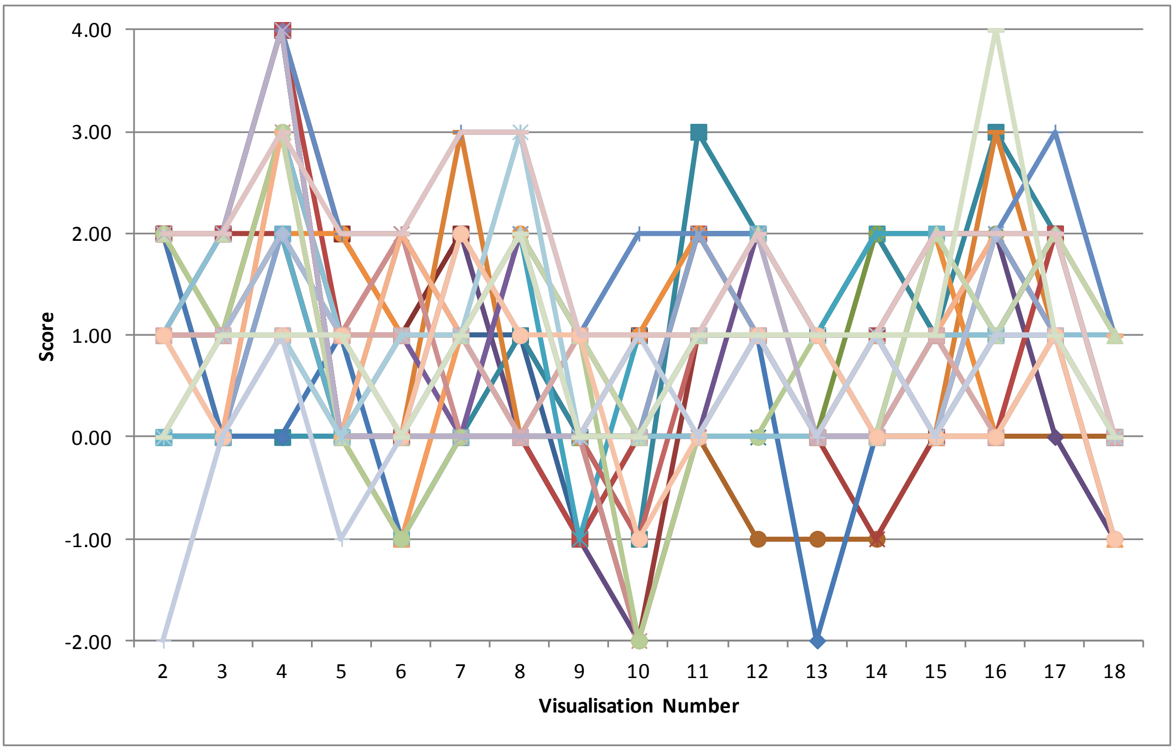


Figure 56. Graph of all responses in the evaluation. The visualisations are in the order completed, with 2-4 = Visual Only, 5-8 = Vision & Sound (Piano), 9-10 = Sound Only (Piano), 11-14 = Vision & Sound (Bird), 15-16 = Sound Only (Bird), 17-18 = Vision with Scale. There are no obvious groupings or clustering, but some particular methods tend to overestimate (e.g. Sound Only (Bird)) and some tend to underestimate (e.g. Sound Only (Piano)). Visualisation number 1 was a practice task to familiarise participants with the exercise.

6.4.5 Post visualisation question results

The participants were asked some questions after they had completed the visualisations, summarised in Table 27.

	Piano notes	Bird calls	No preference
Which sound did you find most helpful?	75.6% (34)	15.6% (7)	8.9% (4)
Which sound did you find more pleasant to listen to?	66.7% (30)	15.6% (7)	17.8 (8)

	Used sound exclusively	Used sound mainly, but also vision	Used sound and vision equally	Used vision mainly, but also sound	Used vision exclusively
When you could use both vision and sound, which did you use more?	0% (0)	44.4% (20)	26.7% (12)	26.7% (12)	2.2% (1)

Table 27. Responses to questions asked after the visualisations.

The majority of people found piano notes both more helpful and more pleasant to listen to, which matches what was said in the discussion session. There were no significant differences in performance between groups who preferred piano notes or bird calls, although the heavily skewed nature of the data limits the possibility of significance. Comparing piano notes with a sound that was preferred by more participants would allow a more detailed comparison.

Nearly everyone (97.8%) used sound and vision in some combination when both were available which shows that having both available is useful for this activity. The participants who said they used mainly sound, but also vision performed significantly worse with the Visual Only method (1.62) than any other methods (p = 0.017) as shown in Table 28. Participants who said they used vision mainly but also sound performed better with vision than those who said they used mainly sound but also vision, but this difference was not significant. The performance of those who used sound but also vision implies that as a group they will perform more effectively with a sonification than those who mainly used vision. However, the difference was not significant throughout all combinations.

	When you could use both vision and sound, which did you use more?		
Method	**Used sound mainly, but also vision**	**Used sound and vision equally**	**Used vision mainly, but also sound**
Visual Only	1.62	1.31	1.14
Vision & Sound (Piano)	0.58	0.90	0.40
Sound Only (Piano)	0.05	0.08	-0.21
Vision & Sound (Bird Call)	0.73	0.77	0.73
Sound Only (Bird Call)	0.95	0.92	1.12
Vision with Scale	0.63	0.54	0.68

Table 28. Participants performance (mean score, 0 = correct) split by preference for sound or vision.

6.4.6 Factors influencing the results

Very few people chose the lowest level of confidence (7) for their answers. The spread of confidence did not vary significantly between the different methods, but participants who had lower confidence were more likely to get the correct answer. The confidence subjective measure had an opposite relationship to the objective measure, compared to the relationship expected. This is further discussed in the regression section (6.4.7) of the analysis.

As a part of the evaluation, the location of the participants when they completed the evaluation was recorded (i.e. whether they were to the left or right of the centre of the screen and how far back they were) to see whether this had any impact on their performance. This data showed that location had no significant impact, given the limited variety of locations of participants.

The overall learning style score was calculated in a similar manner to the previous case study, with the auditory score subtracted from the visual score (see Table 29). However, this overall score had no relationship to the performance overall or for each individual method.

Learning Method	-5	-4	-3	-2	-1	0	1	2	3	4	5
Mean Score	0.47	0.88	1.10	0.71	0.66	0.95	0.86	0.73	0.46	0.35	0.59

Table 29. Mean score for each learning method value

However, considering each element individually it was clear that some elements were more useful than others, as shown in Figure 57.

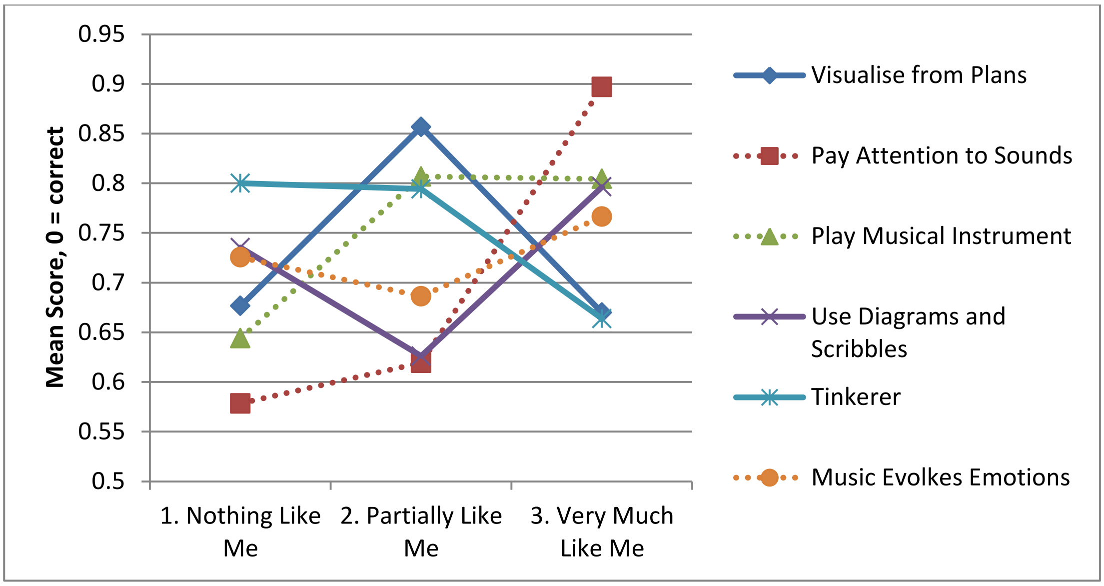


Figure 57. Graph showing the mean score for the different learning style elements. Aural methods have dotted lines, visual methods have solid lines.

There is a slight trend that participants who said the aural learning style elements (dotted in Figure 57) were 'Nothing Like Me' performed more effectively in the evaluation. Therefore, people who paid attention to sounds were more likely to be incorrect than those who did not. However, the message is not clear in this form so each individual statement was included in the regression analysis to see if it had a significant effect.

6.4.7 Regression

Given the wide variation in success of participants between the different methods and the few obvious patterns within the questions, a binary logistic regression was run to see which factors explained the variation seen in the results. Until this point the results have been the difference between the participants answer and the correct answer (ranging from –2 to +4 with a correct answer being 0), for the regression the responses were coded as 0 or 1, with 0 being incorrect and 1 being correct consistent with logistic regression conventions. A binary logistic regression was chosen rather than a multinomial regression because the answers could equally well be categorised as correct or incorrect and it would be much harder to predict multiple categories with the categorical independent variables available than to predict the binary options of either correct or incorrect.

A regression was run for all results overall and additionally it was run for each method individually, which provided much more useful output. A number of subgroups of the different methods were created, on which the regression was also run:

- Sound (All) – Vision & Sound (Piano), Sound Only (Piano), Vision & Sound (Bird Calls), Sound Only (Bird Calls)
- Sound Piano (All) - Vision & Sound (Piano), Sound Only (Piano)
- Sound Bird (All) - Vision & Sound (Bird Calls), Sound Only (Bird Calls)

The same independent variables were included in all of the regressions and a Forward Stepwise method was used because there was no previous information to base a decision on which variables to include. The variables included were:

- GIS Knowledge
- Landscape Visualisation Knowledge
- Planning Knowledge
- 'You pay attention to the sounds of various things' (element of Learning Style question)
- 'You play a musical instrument or sing on (or close to) key' (Learning Style)
- 'You use diagrams and scribbles to communicate ideas and concepts' (Learning Style)
- 'You can easily follow instructions represented in diagrams' (Learning Style)
- Attempt Number
- Confidence[11]

The regression run using the variables listed above gave very unreliable results because there were often a very small number (1 or 2) of cases with a particular combination of predictor attributes. Some of the Cox & Snell pseudo R^2 values were very high (such as 0.504 for the Visual Only group), but upon further investigation these were found not to be as reliable as some attribute combinations were related to small numbers of cases (or even individual ones).

To address this, the variables were reviewed and ones that did not have a significant impact on more than two models were removed. Also, the answers were grouped together to reduce the number of categories (as shown in Table 30). The majority of predictors were reduced to two categories, which made the regression more robust. The only variable that could not be reduced to two categories was GIS Knowledge, which was reduced to 4 categories (from 5). This was because the correct answer values did not follow any pattern to allow different responses to be combined. The R^2 values were fairly low (see Table 31) and some individual visualisation method

[11] See Appendix D.2 for the full list of questions the participants were asked.

models did produce significant results, while Sound Only (Piano) and Vision with Scale did not have any significant factors.

Question	Old Categories	New Categories
How much do you know about GIS Knowledge / Spatial Data?	1 (A lot)	1
	2	2
	3	3
	4, 5 (Nothing / Very Little)	4
You pay attention to the sounds of various things	1 (Nothing Like Me) 2 (Partially Like Me)	0
	3 (Very Much Like Me)	1
You use diagrams to communicate ideas	1 (Nothing Like Me) 2 (Partially Like Me)	0
	3 (Very Much Like Me)	1
Attempt Number	1st 2nd	0
	3rd 4th	1
How confident are you that your answer is correct?	1 (Low) 2, 3	0
	4, 5, 6, 7 (High)	1

Table 30. Table showing how the categories were combined for the regression.

Method	Factors with a significant influence	Best R^2	Predicted Correct
All Methods	'You pay attention to sounds' / Confidence	0.031	61.4%
Visual Only	Attempt No	0.077	74.8%
Sound (All)	'You pay attention to sounds' / Confidence	0.045	60.9%
Vision & Sound (Piano)	'You pay attention to sounds'	0.080	64.4%
Sound Only (Piano)	None	-	-
Vision & Sound (Bird Call)	Attempt No	0.063	63.3%
Sound Only (Bird Call)	'You use diagrams'	0.057	64.4%
Sound Bird (All)	Attempt No	0.029	63.0%
Sound Piano (All)	'You pay attention to sounds' / Confidence	0.061	61.1%
Visual with Scale	None	-	-

Table 31. Regression outputs (Cox and Snell R^2 and proportion of the observed results predicted correctly) and factors that had a significant influence for each regression model.

For All Methods the results are very similar to Sound (All) with 'You pay attention to sounds' and confidence having a significant impact. Visual Only also shows a trend where the later attempts score better, but this is not a particularly strong trend, as the probability increases from 0.09 to 0.33, which is not enough to break the 0.5 barrier to be predicted as a success.

For Vision & Sound (Piano), the tendency to pay attention to sounds had a negative effect on the result (i.e. people who did not pay attention to the sounds of things performed more effectively). This is likely to be because those who do pay attention to the sounds spend time and effort listening to the particular notes, rather than focusing on the task in hand. It could also be that those who pay attention to sounds found the visual and sonic elements contradicted each other in the visualisation (as mentioned in the discussion sessions), which may have resulted in a lower score. This effect is also seen in the Sound Piano (All) group. Additionally, confidence has an impact with a lower confidence producing a better result. This is perhaps because some people were over confident and chose the incorrect answer.

For Vision & Sound (Bird Call) and Sound Bird (All) the more attempts people had the better they got. This is as expected because using the bird calls in this way is a relatively unusual application, and further experience of using the sound gives better results. For Sound Only (Bird), people who tend not to use diagrams to communicate tended to score better, because they are less reliant on the visual aspect of the visualisation which is not available for this method.

An issue of attention could result in some of the effects seen in the models. People who do not pay attention to the sounds of things and have a lower confidence scored better. This is because people who do not pay attention to the sounds of things see themselves as ‘not aural’, so they try harder (concentrate more) in the tasks, and so perform better. Additionally, they would rate their confidence lower than people who perceived themselves as ‘aural’ who may think they have performed better, despite how they felt the evaluation went.

The prediction success rate varied between 60% and 70% for the models, but the majority of this success is achieved by assuming the result is incorrect, particularly for the less successful ones. The R^2 values were relatively poor, but the results from these regressions were more reliable compared to the results from the original regression. Higher R^2 values (up to 0.504) could be achieved by using the ungrouped input variables, but as explained previously these were much less reliable because of the very small number of cases in certain attribute combinations.

The regression from the reduced categories generally gave the same results as the regression with the full categories, apart from the Vision & Sound (Bird Calls) and Sound Only (Bird Calls) methods. They still featured the same factor (Attempt No), but the impact it had was opposite. In the full models, the second attempt (the only one which had a significant impact) was worse than the first one but in the reduced models the later attempts performed better than the earlier ones. This is because the finer differences in the different attempts for these methods are lost when the number of categories is reduced to two. While the reduced regression does limit the ability to predict whether users answers will be correct or not, it is much more reliable and useful than the full regression model, due to the observed number of users in each subgroup.

Overall it is possible to get an idea of whether participants will be able to use the sonification method but there is still a wide variation within the groupings that can be discerned. For the sound methods, some of the issues addressed by the learning style questions are important, as shown by the impact of ‘You pay attention to the sounds of various things’. However, as shown very clearly in Section 6.4.6 the combined learning style does not provide any information, and the individual elements only provide a limited insight. The ‘pay attention’ aspect does have an impact, but not in the expected way. A broader range of learning style questions related to sound should be asked in future research, to see whether the effect can be more effectively predicted. The Attempt No also features strongly in the regressions with more repetitions providing a better result, however some randomisation of the visualisations (and therefore a larger group of testers) would give more weight to this result.

6.4.8 Discussion sessions

Overall the results from the discussion sessions were similar to the questionnaire results. The discussion sessions did provide some additional information, but nothing that contradicted the previous results.

A number of participants said that they would have found the exercise much easier if they heard the 2000m sound (rather than just the 1000m and 3000m sounds), but understood that that would make the exercise too easy. People were divided about whether low sounds should indicate close distances or far distances; it is likely that this is something that varies between individuals, so would be something that should be customisable for individual users.

Many participants found the bird calls difficult to differentiate between, and some said that the warbling in the sound was distracting, while others found it useful as the warble cycle was extended with the lower pitched sounds. While piano notes were the most successful sound, they were by no means universal as three people said that they had trouble distinguishing the notes and one person said they could not hear any difference at all. The reaffirms the need to have individual options, and possible future research could include a customisable element.

During the Sound Only (both Piano Notes and Bird Calls) visualisations, around $^{2}/_{3}$ of people closed their eyes when the sounds were played. They said that they felt this helped them answer the questions and it is likely some of them imagined the visualisation in their head. The combination of vision and sound was not ideal for all, and about 10 participants said that the sound and visual aspects contradicted each other rather than providing complementary information. A majority of people commented that it would have been very helpful to have some other objects in the visualisation (such as trees or buildings) to help them see the scale of the wind turbines.

The number of repetitions was significantly higher than in the other case studies, and it was clear from the participants' behaviour and comments that this case study was very close to or at the limit of repetitions that could be completed without the participants loosing attention. As people progressed through the different visualisations, they felt they improved as more and more contextual information was provided. For each set of repetitions within each method, around half of the group did say they found it became easier as they did more, while the other half said it did not get easier.

At the end of the evaluation the sounds were played with the distances shown on screen at the same time, so participants could get an idea of whether they had answered questions correctly. As a result of this, six people commented that they thought they made a consistent systematic error throughout the evaluation, with some underestimating distance and some overestimating distance. Two people of the group of six said they felt they did not understand some of the questions properly.

The discussion sessions provided useful contextual information and they highlighted a few areas for future research, particularly with respect to the possible customisation that may be required to accommodate people's individual preferences.

6.4.9 Issues with evaluation sessions

During one of the sessions there was an issue with one of the keypads of the audience response system. The participants answer did not appear to register the first time they pressed the button. The participants were asked to press their button again which resulted in the correct number of responses being recorded. It is possible that this was caused by the respondent pressing their button too early, or from a low battery in the response pad. However, the pads were checked before each use and any found not to be working were replaced. The only effect this had was to make some of the timing of the answers unreliable. It did not affect the answers themselves.

In the session at Ordnance Survey one participant was confused by the first vision and scale visualisation. They were not sure whether the question (how many turbines are within 2000m) was referring to distance from the marked turbine or from the user viewpoint. It was from the position of the user, consistent with the other questions. In future evaluations, this was clarified, but this participant's answer was not collected for this question.

Ulla (2005) observed that when comparing an audience response system to a traditional paper and pen form, respondents tended to have greater doubts about the validity of their responses. It was hypothesised that this was because participants could not re-check their answers (whereas those using the paper and pen could). However, the participant's actual answers did not vary significantly, so would not have had an impact. It could, however, explain some of the low confidence responses given by the participants in this case study.

The issue of ordering the different evaluation methods was mentioned earlier, with the Vision with Scale method required to come last to prevent participants using information from that method with the other methods. Another approach would have been to show each quarter of the participants one method of the four possible ones, but this would have required a sample four times as large to generate the same statistical significance, which was not feasible for this study.

6.5 Conclusion

This section provides a general overview of the results, and then reviews them in relation to the specific research questions for this case study. Overall the best performing method was Sound Only (Piano) (with a mean score of -0.02), followed by Vision & Sound (Piano) (0.6), Vision & Sound (Bird Calls) (0.72), Vision with Scale (0.72), Sound Only (Bird Calls) (0.98) and Visual Only (1.39). Sound Only (Piano) was not expected to perform the best given its poor score in CS1, but participants said that they felt the sound on its own was easier, compared to sound and vision together for Piano Notes. This was because the exercise was just a comparison exercise, listening to different notes, rather than needing to relate the sounds to any other data. This was not the case with Sound Only (Bird Calls) because participants found the sounds much more difficult to tell apart and were more reliant on the visual element. Visual Only was significantly worse than all of the other methods, due to the very limited information available in the visualisation. The regression confirmed the previously observed trends, as well as highlighting some interesting additional aspects.

The specific research questions were:

- Can sound be used to represent additional data in a virtual reality / landscape visualisation setting?
- Is using sound to do this more effective than using existing visual methods?
- How does the participants learning style and/or previous knowledge of the data set impact on their ability to use the sonification?
- How does using an audience response system to collect evaluation responses compare to the methods used in CS1 and CS2?
- Do OS and UKCIP respondents perform better or worse with this data set compared to their expert data set? (i.e. AL2 for OS and UKCP09 for UKCIP).

Each will be addressed individually in the next sections.

6.5.1 Effectiveness of using sound

Sound can definitely be used to represent data in a landscape visualisation. The piano notes were particularly useful and using sound to supplement the visualisation provided a significant improvement in performance ($p < 0.001$). Piano notes alone gave even better results, but this may not be very useful in the context of providing additional information. The bird calls sound provided some improvement compared to vision alone, but not as much as the piano notes. The sound used is very important for the sonification and the preference and ability varied between participants.

6.5.2 Learning style and data knowledge

The learning style as a combined measure was not particularly useful, but individual elements of the learning style did provide some relevant information. The element 'You pay attention to the sounds of various things' had a very strong association with the participant's ability to answer the questions, with those who did not pay attention to the sounds of things scoring better. This is because these participants were more focused on the task than the participants who did pay attention to the sound. They were concentrating on the sound and not the task in hand. 'You use diagrams to communicate ideas and concepts' was also associated with performance for the bird call methods, where those who did not use diagrams performed better. This shows a split between the visual and aural learning communities with the aural learners performing better than the visual learners. However this is by no means conclusive and is an indication of an area for future research.

Knowledge of GIS data was the only knowledge area that had any influence at all and this only had an influence in the full regression, not the reduced categories regression. There is a limited amount we can deduce from the full regression, but higher levels of GIS knowledge led to better performance. However this was not significant enough to be shown in the reduced categories regression so more specific questions regarding knowledge of the data set are required. Each case study was designed to have a particular expert group (planners in this case study) who may or may not perform better than the other participants (who are not experts in this field). In this case study there were no significant differences between the expert participants and non-expert participants. If the expert users were a more specialised group who used and designed advanced landscape visualisations on a regular basis, a difference may have been seen. A fuller review of the three different case studies is included in Chapter 7. Knowledge of uncertainty data is not as relevant to this case study as it is the others, which was why it was not included in the questionnaire.

6.5.3 Use of audience response system

Using the audience response system worked well and allowed the sonification to be used in a group setting, which previously was not possible. It also allowed for much faster development and piloting of the evaluation. Compared with the one-to-one sonification in CS1 and CS2, the group sonification was quite limited. Using a sonification on an individual basis provides a better user experience as the participant can interact with the sonification, which allows them to get more information from the sonification. Participants who had participated in all three case studies also agreed with this.

6.5.4 Regression results

The regression analysis showed how the important factors varied between the different methods although some had a much stronger influence than others. For the sonic methods 'You pay attention to the sounds of various things' and 'confidence' had a significant effect, but not in the expected way. Participants who did not pay attention to sounds were more likely to get the correct answer, and this is because people paying attention to the sounds were focusing too much on the sounds and not focusing enough on the exercise. Additionally, people with lower confidence were more likely to get the correct answer and this may be because participants who were highly confident with their result were actually overconfident and did not focus enough on the task. These factors were also seen in the regression for all the different methods. For methods using bird calls, the attempt number was important, with the later attempts more likely to be correct than the earlier ones. This is what would be expected, as the more familiar people are, the better they will do. It is also more pronounced with bird calls that piano notes, as in this context bird calls are more unusual then piano notes.

6.5.5 Future research

This study has highlighted how sound can be used in a virtual reality environment, and some aspects of the use of sound that need further research and refinement.

The learning style element developed from a simple two question section in the first case study and was extended and refined with each subsequent case study. For CS2, there were 20 statements for assessing the learning style, but the results for each statement were not stored – only the overall total, which reduced the amount of information available. As shown in this case study, some of the statements are much more useful than others at assessing whether the user will be able to make use of the sonification. From the six different statements used, it is not obvious why the ones that were useful were so effective. As far as the author can tell, there are no studies in the sonification literature about whether certain types of people are more successful at using sonifications than others. The only aspect that has been thoroughly investigated in the sonification literature is whether the ability to play a musical instrument has an impact and it has been shown to have no impact (Walker, 2000; Mauney and Walker, 2007). There is, however, a definite variation within the sample group in terms of the effectiveness of sonification, but this may be masked by participant's preferences for different sounds or sonification methods. However, some more research in this area, particularly how to measure whether people will be able to use a sonification, is required.

Related to this point, having some type of customisable sonification in future research would allow the user to choose their optimal sonification method. This would both provide additional information for the researcher, as well as allowing the participant to complete the evaluation task in the most effective possible manner. The participant may also be able to provide a reason for their preferred sonification, if they can experiment with and choose from a selection of different sounds.

Interactivity has been mentioned in the first two case studies as being very important for an effective sonification, referring to the participant's ability to interact with the sonification on a one to one basis. This could be achieved for this example by allowing participants to click on individual wind turbines and hear the sound related to them. Another, more intuitive option would be to use eye tracking technology, where the point on the landscape the participant is looking at could be tracked. A number of studies have used eye trackers to see how people explore different visual landscapes (de Lucio et al., 1996; Barbuceanu et al., 2011; Risko et al., 2011) and it would be an interesting possibility to see how the presence of a sonification would affect this movement. Potentially an eye tracker could be used to drive a sonification by the area that the participant is currently looking at, which may provide a more interactive or immersive environment for the participant.

Another way to develop the interaction side would be to create an interactive landscape, where the participant could move around and look at the landscape from different points of view. A higher level of interaction, with both the landscape and the sonification would increase the immersion feeling experienced by the user (Schroth, 2010) and would lead to greater understanding of the information. However both these ideas would require a significant level of resources.

6.6 Findings in Relation to the Research Hypotheses

This section will relate the findings from this case study to the wider research hypotheses of this thesis. The results from all three case studies will be compared and reviewed in terms of the research hypotheses in Chapter 7.

6.6.1 Specific research aims

The specific research aims were:

1. What form does uncertainty information need to be in for it to be represented using sound?
2. How effective are different means of representing uncertainty in spatial data?
3. Which types of sound variation best represent spatial data uncertainty to users?

This case study addresses all three specific research aims, particularly in comparison with the other case studies. The availability of uncertainty information (Aim 1) limited the applicability of this case study as landscape visualisations with explicit uncertainty information are very rare, and would have taken a significant amount of time to construct and implement. Distance was chosen as it is an explicit variable, it is easy to sonify and can provide some useful information for the participant. Different methods of representing distance were evaluated (Aim 2), comparing Visual Only, Vision & Sound, Sound Only and Vision with Scale. Sound Only and Vision & Sound were most effective but this did depend on the sound being used. Piano Notes were more effective than Bird Calls (Aim 3), with Sound Only (Piano Notes) significantly better than Sound Only (Bird Calls) ($p < 0.001$).

6.6.2 Research hypotheses

This section outlines the research hypotheses in reference to the results from this case study, both in terms of the sonification design and evaluation.

H1. The way uncertainty is conceptualised within the spatial data set will impact how it can be represented in a sonification.

As explained in Section 6.2.2, there is uncertainty within future landscape visualisations, but it is rarely represented explicitly within the data. Instead, for this case study distance was used as an example, but the principle could be applied to uncertainty data if it was available. The way uncertainty data are conceptualised and stored within the data is very important when considering how easy it is to implement. The conceptualisation will also impact on how uncertainty is stored in the final data product, which will shape the ability to represent it in a meaningful way in the final representation.

H2. The technology available to design, implement and evaluate the sonification of uncertainty in spatial data will impact the findings.

The technology used in this evaluation allowed for a very rapid development and testing of the sonification, but limited the complexity of the sonification. The evaluation method used in this case study was different to that used in the first two case studies, which does limit the findings. However, there are some similar trends which are applicable to wider areas than those evaluated in these case studies.

H3. The ability of the user to make effective use of the sonification will be effected by a variety of factors including their knowledge and previous experience.

Knowledge of the area of landscape visualisations did not impact the performance of the participants in the evaluations, due to the non-specialised nature of the data involved. However, elements of learning style and confidence did influence the results so some participant characteristics do influence their performance. Confidence has been shown previously to impact participant's performance and it is also more likely to have an impact in a real world setting than an academic evaluation.

H4. Providing redundancy (dual-coding) in a task by showing the same information both visually and sonically will improve performance.

Using sound in combination was significantly more effective than vision alone for both piano notes and bird calls. However, using piano notes alone was more effective than using piano notes and vision together, contrary to the expected result. This was because of the reported mismatch between the visual perception and sonic perception of the distance of the wind turbines. If this issue was solved, then it is expected that the performance of sound and vision would match sound only and possibly surpass it.

H5. Transferring the representation of uncertainty information from the visual to the auditory domain to address visual saturation will reduce cognitive load and therefore improve performance.

The nature of a landscape visualisation embeds distance information within it so it would be very difficult to try to represent this information using exclusively aural methods. The provision of additional information (such as uncertainty) was not tested in this evaluation so it is not known whether this would improve performance. However, it was possible for some participants to detect a difference in auditory and visual information so it is likely that additional information could be communicated effectively.

H6. Preference for one sound from an option of two will result in better performance with the preferred sound than performance with the other sound.

Two different sounds were compared in this case study (piano notes and bird calls) and there was a clear preference and performance difference. The majority of participants preferred the piano notes but a small minority preferred the bird calls as more effective ways of representing the information. However, there was no significant difference in performance between the two different preferences. This was due to the small number of participants who preferred bird calls. If a greater range of sounds were available, it is likely that the preference for sounds would be less skewed and then the difference between performance and preference for different sounds might have been significant. Participant confidence in their own performance was also important, as shown in the regression, although the relationship between confidence and performance was not as expected.

This case study fully addresses hypothesis 1, 2, 3 and 6 and partially addresses hypothesis 4 and 5. It has compared different sonic stimuli and shows a potential performance increase for individual participants based on their sound preferences. It also expands the information available on how learning style can influence performance as well as including participant confidence as an important factor. This case study will be reviewed alongside the other two case studies in the next chapter.

Chapter 7. Discussion

This section will review and discuss the results from the three case studies of this thesis. An overview of each case study and its findings will be provided. The results from each case study will be combined and reviewed in terms of the research aims and hypotheses. The advantages and disadvantages of the case study approach will also be discussed, followed by a discussion of the problems encountered with this research and a review of the potential future research. The final part of this chapter will revise the conceptual model presented in Chapter 2, updating it to reflect the results from this thesis.

7.1 Overview of Case Studies and Findings

7.1.1 CS1: Accuracy of address locations

This case study created a sonification tool within the ArcGIS environment, allowing the positional accuracy of address locations to be shown using visual and/or sonic stimuli. The positional accuracy information is a very important part of the AL2 data set, but it is often ignored by the data users for a variety of reasons. These reasons include not understanding the importance of the positional accuracy data, not being able to represent the positional accuracy data effectively and not being aware of the data's existence. A total of 49 participants from UEA and Ordnance Survey (OS) took part in the evaluation. The most successful representation method was using vision and sound to show the same variable (mean proportion correct 82%) with the least successful being when sound was used to show a different variable to that shown visually (mean proportion correct 65%). The logistic regression model showed the most important factor to be the frequency of the target feature to be identified, followed by the presentation method and knowledge of address data. This highlighted the importance that the type of data has on the success of the sonification and that knowledge of the dataset is an important factor, although knowledge of the data set did not make a significant difference to the regression model. Detailed questions about the participants' knowledge of the uncertainty element of the address data were not asked and it is these that are likely to be important. The results also highlighted the fact that learning style is important, but did not show a significant difference, so this factor in addition to data set and uncertainty knowledge was explored further in the next case study.

7.1.2 CS2: Uncertainty in UKCP09 data

The second case study used the Google Maps environment for the evaluation which was developed as a web based application. The recently released UKCP09 data set contains information on the uncertainty surrounding the predicted future climate, which can be very difficult to represent visually because of the amount and complexity of the data available. Visual, sonic and combined visual & sonic methods were evaluated to see which was most effective at representing the uncertainty in the data. Participants were recruited from UK Climate Impacts Programme (UKCIP) as well as UEA and OS (total n = 62). The most successful method was one of the training maps, where vision and sound were used to show the same variable on one map (S, 91% correct)[12]. The main evaluation for CS2 required two data sets to be shown (the temperature data and the uncertainty data) and the most successful method in the main evaluation was using one map to show both data sets, with one data set shown visually and one data set shown sonically (VS, 80%). However, this was not significantly different to using one map for each variable (shown side-by-side) using both sound and vision to show each data set (VSVS, 79%). Using vision (without sound) was significantly less successful (VV, 68%). Awareness of the data set had a significant impact on the ability to use the sonification, with those aware of the data set performing more effectively ($p = 0.031$). Participants with a more visual learning style appeared to perform better, but it was not clear whether this was related to the sonification specifically, or the fact that the spatial data in exercise made extensive use of the visual sense. The web based interface allowed the evaluation to be conducted in a wider variety of places than CS1 and also allowed the option of remote completion by participants who were unable to physically attend evaluation sessions.

7.1.3 CS3: Estimating distance in landscape visualisations

The third case study did not use an explicit GIS environment, but instead considered the use of sound to help participants estimate distances of wind turbines in a landscape visualisation. A total of 45 participants were recruited, including planners from Norfolk County Council and staff from the Broads Authority as well as participants from UEA, OS and UKCIP. The evaluation was run in a very different manner to the other case studies, using a group based sonification rather than individual sonifications as used in previous case studies. This time sound alone was the most effective method (59% correct), while sound and vision together still had acceptable performance (52%) and visual only the worst performance (25%). Two different sounds were compared and piano notes performed much more effectively than bird calls. The learning style variable was investigated much more thoroughly in this case study and certain elements were found to have a significant influence using a binary logistic regression. However, this was not seen across all representation methods in a consistent manner. Each visualisation was repeated more frequently

---

[12] See Section 7.2 for more details on how the results for the three case studies were compared.

in this case study to evaluate the learning effect and it was found that participants performed better on the later repetitions, particularly with the unfamiliar interaction methods.

This case study used a different evaluation method to the first two case studies. The first two used a computer based evaluation method where the participants worked through the evaluation at their own pace. They also had a one-to-one interaction with the sonification, where they could control it and explore it as much or as little as they liked. This was very different to CS3, where the evaluation was run as a group session, with the visualisations and sonifications shown to users on a projection screen. This meant that they had much less interaction with the sonification and comments from the discussion session showed that participants found the interactive method much more useful than the non-interactive method.

7.2 Overall Findings from the Three Case Studies

There are some strong messages that come across from the results of all three case studies, one being that using sound and vision to reinforce the same data is the most effective method. In addition, it appears that regardless of the method of sonification used, some participants find the sonification much more helpful and easier to use than others. This reinforces the fact that there is something about the participants that make them either able or unable to use the sonification, but it was not possible to fully clarify this in the research.

The results for each of the methods from the three case studies can be compared and Table 32 shows how the different methods relate to each other. Not all of the representation methods in each case study were fully duplicated in the other case studies, because of limitations deriving from the data or presentation techniques used. However many can be directly compared and are outlined below. All three cases studies included a basic visual only method as well as a visual and sonic method, where sound reinforced the information shown visually. The different data sets were also most effectively sonified in different ways, so not all of the different sonification methods used are directly comparable. CS2 is more complex to compare because some methods used two maps (side-by-side) to show the data, which was not the case with any of the other case studies. These results are of limited use when comparing with the other case studies, but did provide useful information in their own right.

Case Study	Method	Piano Notes?	Continuous or Discrete data?	One or two maps?	One or two data sets?	Comparable with:
CS1	V - Visual	-	D	1	1	CS2:V,VV CS3:V,VS
	S - Sonic	Y	D	1	1	CS2:S CS3:SOP,SOB
	VSS - VS Same	Y	D	1	1	CS2:VSVS CS3:VSP,VSB
	VSD - VS Different	Y	D	1	2	CS2:VS
CS2	V - Visual	-	C	1	1	CS1:V CS3:V,VS
	S - Sonic	N	C	1	1	CS1:S CS3:SOP,SOB
	VV - Visual Visual	-	C	2	1	CS1:VSD*
	VSVS - Visual Sonic on both maps	N	C	2	1	CS1:VSS, CS3:VSP,VSB
	VS - Visual and Sonic on one map, showing different data	N	C	1	2	CS1:VSD
CS3	V - Visual Only	-	D	1	1	CS1:V CS2:V
	VSP - Vision & Sound (Piano)	Y	D	1	1	CS1:VSS CS2:VSVS
	SOP - Sound Only (Piano)	Y	D	1	1	CS1:S CS2:S
	VSB - Vision & Sound (Bird)	N	D	1	1	CS1:VSS CS2:VSVS
	SOB - Sound Only (Bird)	N	D	1	1	CS1:S CS2:S
	VS - Vision with Scale	-	D	1	1	CS1:V CS2:V,VV*

Table 32. Grid showing which methods from each case study can be directly compared with each other. Most do have a direct link, but those marked with a * may not.

In order to compare the results from the three different case studies they need to be in a comparable format. The results for CS1 and CS2 could both be represented as the mean overall proportion correct, but CS3 could not. The results for CS3 were a mean difference from the correct answer, with the correct answer being 0 (see Section 6.4.2 for details). In order to allow CS3 results to be compared to the other case study results, the mean results for each method were recalculated into direct correct / not correct proportions (see Table 33 below).

Method	CS3 Mean Difference	Equivalent % Correct
V - Visual Only	1.39	25
VSP - Vision & Sound (Piano)	0.60	52
SOP - Sound Only (Piano)	-0.02	59
VSB - Vision & Sound (Bird)	0.72	37
SOB - Sound Only (Bird)	0.98	38
VS - Vision with Scale	0.72	33

Table 33. The mean correct answer for CS3 and the equivalent overall percentage correct.

Table 34 shows all the summary scores from each method in each case study. The difference in the case studies limits the direct comparisons that can be drawn between the results, but some comparisons can be made, particularly in terms of the relative effectiveness of the different methods within each case study. The order of these (i.e. most successful to least successful) can then be compared between the different case studies. While it has been mentioned previously that CS3 is probably the simplest task conceptually, the performance was less good across all methods than for the other case studies. This could be because CS1 and CS2 required interaction from the user, whereas CS3 was a passive experience, which meant that the user was less engaged with the exercise. For CS2, the results for Visual and Sonic are of limited use because in these first two maps not all of the data was shown, as the uncertainty data was not included at all. The results are included for completeness, but they are not given ranks as it is not meaningful to include them in the comparison.

Method	CS1	CS2	CS3
Visual Only	82 (V) (2nd)	86 (V)* / 68 (VV) (3rd)	25 (V) / 33 (VS) (3rd)
Sonic Only	67 (S) (3rd)	91 (S)*	59 (SOP) / 38 (SOB) (1st)
Visual and Sonic reinforcing the data	86 (VSS) (1st)	79 (VSVS) (2nd)	52 (VSP) / 37 (VSB) (2nd)
Visual and Sonic showing different data	65 (VSD) (4th)	80 (VS) (1st)	-

Table 34. Results from the three different case studies for each method compared, including the order (1st to 4th) of the different methods within each case study. The comparability of the different case studies is discussed on the previous pages. *For CS2 method V and S are not included in the comparison exercise as they did not show all the data.

One result that was consistently shown throughout all three case studies was that using sound to reinforce data shown visually tended to result in higher performance than either visual or sonic methods alone. The participants said that sound often worked as a confirmation of what they saw visually and this confirmation increased the confidence they had with their answer. This agrees with previous research, with the multimodal aspect of the interface improving both performance (Brewster, 1998; Constantinescu and Schultz, 2011) and confidence (Jeong and Gluck, 2003) in tasks.

Attempts to show two different data sets had varied success, with VSD performing poorly in CS1, but VSVS and VS in CS2 performing reasonably well. This is due to the different level of integration of the data sets, with the uncertainty data in CS2 being very well integrated with the temperature data, but the council tax data was not well integrated with the address positional accuracy data. The uncertainty data formed a key part of the UKCP09 data used in CS2 and it is not possible to make effective use of the UKCP09 data without taking into account the uncertainty. This is not the case with the council tax data and the Address Layer 2 data used in CS1 as they are two different data sets which can be used independently. The integration of the data sets is crucial for the participant to be able to use the representation effectively.

The Sound Only method featured in both CS1 and CS3, but provided very different results. In the first case study it was very unsuccessful, being the worst performing method. However, in CS3

Sound Only (Piano) was the best performer by far. The third case study was a much simpler exercise than the first one and it is this that appears to have made the difference. Equally, the Sound Only (Bird Calls) performed less effectively than the piano notes and sound & vision. This shows that both the task participants perform and the sound used have a significant impact on the effectiveness of the sonification. Preference for sound to be used for specific variables was also seen by Harding and Souleyrette (2010) who evaluated visual, sonic and haptic representation methods in a GIS highway planning scenario. Their users preferred terrain data in haptic mode and land use data in sonic mode. The relationship between objective and subjective results is less clear in the other case studies and is also not directly comparable because of different wording in the questions. However this is something that should be considered in future research as it can have an impact on performance.

7.3 Research Aims & Hypotheses

This section will discuss the combined results from the three case studies in terms of the original research aims and hypotheses.

7.3.1 Specific research aims

The initial research aims stated at the beginning of this research were:

1. What form does uncertainty information need to be in for it to be represented using sound?
2. How effective are different means of representing uncertainty in spatial data?
3. Which types of sound variation best represent spatial data uncertainty to users?

The main focus has been on the representation of uncertainty within spatial data and how this can be achieved using sound. The three aims are in a progressive order, with each one building on the previous one. The first element to be considered is what form the original geographic data needed to be in in-order to represent the uncertainty effectively. The uncertainty information has to be included in the data set for it to be shown and particularly for the first case study this was a stumbling block as many datasets do not include uncertainty on an object level basis within their metadata, if they have any metadata at all. If the uncertainty is included, it will be in different forms for raster and vector data. The different ways uncertainty is stored require a different approach for it to be represented using sound. The complexity of the uncertainty data is also important, with more complex data being more complicated and time consuming to sonify, making the subsequent sonification harder to use. However, the integration of the uncertainty information with the base data is also important, as a highly integrated data set is much easier to understand than a non integrated one. The second element considers different methods of representing uncertainty in spatial data. Overall, uncertainty information can be shown effectively using sound, but it is most effective when a limited amount of data are shown and when the users have a good level of understanding of the data set being shown. Using sound to reinforce data that is shown visually is also more effective than using sound alone (as stated by dual-coding theory, Paivio, 1991 and shown by Storms and Zyda, 2000), but this varies depending on the complexity of the dataset being shown. The third element considers which type of sound variation best represents uncertainty in spatial data to users. The type of sound that is most appropriate for a sonification varies with the data set and the participant, including the participants' skills and previous experience as explored in this research. The nature of the data is also important (i.e. field vs. object data models) with sonification methods for continuous data being different to sonification methods for object data. The direct comparison between continuous and object data (synonymous with raster and vector data) is limited in this research, but does highlight a number of issues that would be worth investigating further, such as the choice of a continuous sound for a continuous data set.

This thesis fully considers all three aims over the three case studies. The case study approach allowed the data collected and collection methods to be improved with each iteration and allowed a more effective and comprehensive answering of the aims than a single case study would permit. The results will now be explored in reference to the original research hypotheses.

7.3.2 Research hypotheses

This section outlines the six research hypotheses that are tested by this research in both the sonification design and evaluation phases. The results from the three case studies will be considered with reference to the hypotheses and will then be used to confirm and improve the conceptual model.

7.3.2.1 Sonification design

The hypotheses in the sonification design section are related to factors that impact on how the sonification is designed, including the way the uncertainty in spatial data is conceptualised and stored. This is limited by the data and technology available at the current time.

H1. The way uncertainty is conceptualised within the spatial data set will impact how it can be represented in a sonification.

The way uncertainty within spatial data is conceptualised by the data producer and represented within the data set affected all three case studies. The presence or absence of explicitly stated uncertainty data was a major factor that influenced which data sets could be used in the case studies. Decisions taken by the data producers on which elements of uncertainty to include in the data set will also impact the ability to represent this data (see Section 4.2). The level of integration of uncertainty with the data set is also important, with the spatial accuracy data within AL2 stored as an additional field, which could (and often was) ignored despite detailed available instructions (CS1). This is in contrast to the UKCP09 data where the uncertainty was integrated with the data (Jenkins et al., 2009), so the data could not be used without taking into account the uncertainty element (CS2). However, this also made the data much more difficult to use and it required a substantial amount of training for the users (see Section 5.2.1.1).

H2. The technology available to design, implement and evaluate the sonification of uncertainty in spatial data will impact the findings.

A range of technological solutions have been used to construct the evaluation tools used in the three case studies. The technologies chosen did have an impact on the requirements for the evaluation. CS1 using ArcMap was the most complex (see Section 4.3) and CS3 using PowerPoint was the least complex (see Section 6.3). The web based interface used in CS2 was designed to be easily used on a variety of machines in different locations, and was successful in this respect, performing generally as expected on the UKCIP and OS computers (see Section 5.3). It also allowed the interface to be adapted for remote completion, although not enough responses were collected using this method for the data to be used in the analysis. The comparison of different technologies also allowed for a comparison of interactive evaluation methods, such as CS1 and CS2 where the participant could directly interact with the sonification and CS3 where the participant was just passively listening to the sonification (see Sections 3.1.3 and 7.2). While only some of the results can be directly compared from the three case studies because of the different methods used, in general performance for CS3 was worse and the participants felt less involved with the evaluation (see Section 7.2).

7.3.2.2 Sonification use and evaluation

There are a range of factors that influence a participant’s performance in a sonification, including previous knowledge, presentation method and preference for different sounds (see Chapter 2).

H3. The ability of the user to make effective use of the sonification will be effected by a variety of factors including their knowledge and previous experience.

Knowledge of the data set and experience of uncertainty was only a statistically significant factor in CS2, but it was also a factor that showed an impact on performance in CS1. The type of data used in CS3 did not lend itself to showing an improvement for a set of participants with specific knowledge because the nature of the data was too generalised (see Section 6.4). As well as knowledge of the data set being sonified, knowledge about the representation of uncertainty within the data is important and should be evaluated separately as it has a distinct impact. The impact of knowledge of uncertainty was shown in CS2 for all methods using sound (see Section 5.4.8). However specific knowledge of uncertainty within UKCP09 was not tested and it is likely to be a significant factor which should be tested in future research.

Musical ability and/or experience were tested in CS1 and CS3 and were found to have no impact (see Sections 4.4.1 and 6.4.1, consistent with findings from Walker and Nees, 2011). Learning style is a very difficult concept to test and did not show a significant difference when it was evaluated in CS1 and CS2. When the individual elements of learning style were evaluated in CS3, some of them had a significant impact and some did not. CS3 also showed that participant confidence had an impact on ability to use the sonification, with those who were more confident less likely to select the correct answer (for more details see Sections 4.4, 5.4 and 6.4).

H4. Providing redundancy (dual-coding) in a task by showing the same information both visually and sonically will improve performance.

Using sound to reinforce information shown visually created a significant improvement compared to visual only for CS2 in all cases and in some cases for CS1 and CS3, as you would expect to see according to dual-coding theory (Paivio, 1991). The sounds used effect the significance of this impact, but it always resulted in either better or the same performance. The varying impact of reinforcement is due to the differing levels of difficulty in the case studies (see Section 7.2). It is also due to the specific implementation of sonification, with some implementations more successful than others. The integration of the uncertainty data with the climate data in CS2 increases the coherence of the sonification which is one of the reasons for the better performance. When sonifying a dataset, its level of integration with the data being shown visually must be considered as this will have an impact on the effectiveness of the sonification (see also H2).

H5. Transferring the representation of uncertainty information from the visual to the auditory domain to address visual saturation will reduce cognitive load and therefore improve performance.

Using sound to show information not represented visually showed a significant performance improvement compared to the visual only representation method in CS2, showing that the multimodal representation was effective (see Section 6.4). Using sound to show additional information in CS1 was not successful, because the additional data was not directly related to the base data being represented. In CS2 the additional data (the uncertainty) was linked to the base data (the temperature) and the representation technique was effective. The link between the two data sets being represented is crucial in effectively representing data using this method. This type of multimodal representation for additional information was not tested in CS3.

H6. Preference for one sound from an option of two will result in better performance with the preferred sound than performance with the other sound.

Evaluating the option of different sounds would have added an additional layer of complexity to CS1 and CS2 which would have extended the evaluation to an unreasonable length, so different types of sound were not included. In the discussion sessions, when asked many participants said they would prefer different sounds. CS3 compared two different types of sound and piano notes were more effective than bird calls (see Section 6.4). This is due to the internal variation of the sound and consistency between sounds representing different values. Some participants did prefer bird calls and different participants will prefer different sounds, so there should be a choice for participants in a sonification in order for them to make the most effective use of the sonification. Participant's confidence with the answer is an important factor which impacts their performance and is a significant factor for CS1 (see Section 4.4). The impact of the measures of participants' confidence was not consistent throughout the case studies, as a result of inconsistent questioning.

7.4 Revised Conceptual Model

The conceptual model outlined in Section 2.5 showed the factors that were likely to influence the effectiveness of a sonification. Throughout the three case studies these factors have been tested, with some elements of the conceptual model confirmed, some found not to have an influence and some new factors found to be important, all of which need to be updated within the model (see Figure 58 for the revised version of the conceptual model).

The conception of uncertainty (A1b) was found to have a major impact on the ability to create a sonification, both in terms of how the data producers conceptualises uncertainty as well as how uncertainty is included within the data set (e.g. explicitly or implicitly) (e.g. see Section 4.2). The technology available impacts both the sonification technique and the evaluation process, which influences the data that can be collected (see Section 5.3). Multimodality is a very useful parameter mapping technique (Oviatt et al., 2004) and the need to be aesthetically pleasing is also very important.

In the forms tested, learning style does not significantly influence the ability to use a sonification. CS3 showed that specific elements within a learning style have some influence, but more research

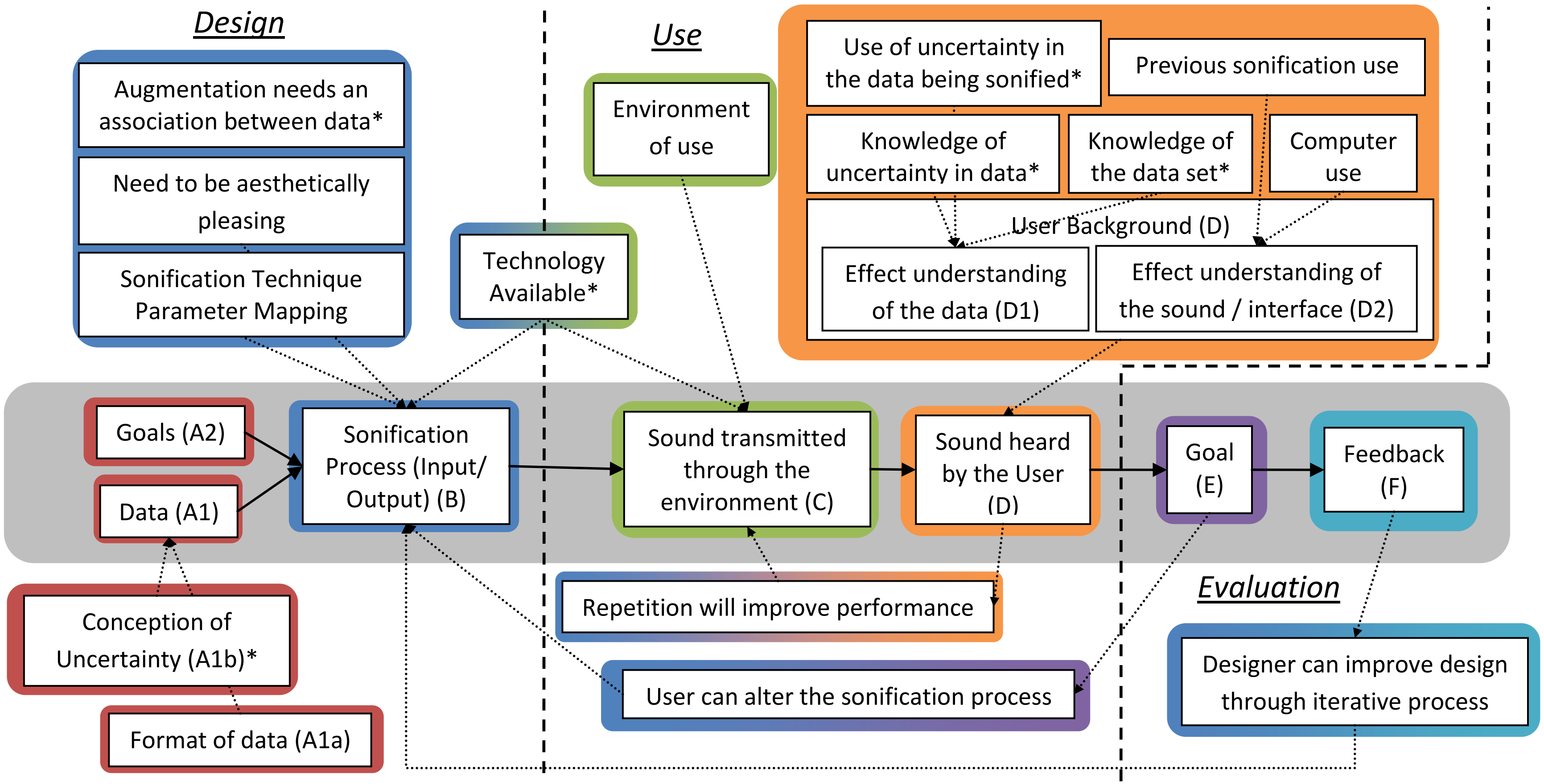


Figure 58. The revised conceptual model (see Figure 23 for original version). * indicates factors in the model introduced or confirmed by this research.

is required before a clear link can been seen so learning style is removed from the conceptual model. Musical ability, experience or exposure was also thought to have some influence, based on previous work (Jeong and Jacobson, 2002). However, in this study it was found to have no influence on performance with the sonification, consistent with other research (Walker and Nees, 2011), so it has been removed from the conceptual model. In CS2 academic experience was tested and found to have no significant impact on ability to understand a sonification. This is in agreement with research by Kidd et al. (2007) who found that there is no significant relationship between intellectual ability (as measured by SAT scores) and ability to discriminate specific sounds.

Several items can be added to the conceptual model which have been shown to influence effective use of the sonification. Firstly, knowledge of the data set has a significant impact, with greater knowledge resulting in better performance. While the participant's use of spatial data in general was already included (as suggested by Coluccia and Louse, 2004), this element refers to knowledge and use of the specific data set that is being sonified. A similar approach applies to knowledge of uncertainty, with knowledge of the uncertainty in the data set being sonified resulting in better performance.

Additionally, when using the parameter mapping technique of augmentation, where additional data are shown using sound and understood in combination with data shown visually, the link between the two data sets is crucial. If data are to be presented sonically and used in combination with data presented visually, then there must be a close link between the two data sets if the representation is to be effective. One example of this is used in CS2, projected future temperature and uncertainty of projected future temperature. This worked very well because the future temperature and uncertainty have a clear association. However similar example in CS1 using positional accuracy and council tax band did not work well, because there was no clear association between the two data sets.

Figure 58 shows the revised conceptual model with the alternations mentioned above. This conceptual model advances knowledge in the field of GIS and sonification with respect to the representation of uncertainty using sound. It also highlights the important factors that will influence the effectiveness of the sonification.

When creating an interactive sonification, the most important factor to consider is the user's knowledge of the data set that is being sonified (as initially stated by Brunet and Cowie, 2012 and extended and confirmed by this research). The users with more knowledge of a data set are more likely to make effective use of the sonification. This is also relevant to specific elements of the data, with the users in this research who had more experience of probabilistic data being able to make more effective use of the sonification. The technology available will also have an impact on the type of sonification that can be designed and therefore how effective the sonification is. The conception of uncertainty within the data set must also be considered, because this will impact what information can be represented in the sonification.

Chapter 8. Conclusion

This chapter will discuss the main headline results from this research, reflect upon the methodological approach used for this thesis, consider the limitations of and possible improvements to this study and look at what the research findings mean for the areas of uncertainty data presentation, uncertainty data use, GIS and sonification.

8.1 Summary of Findings

There are four main findings that can be drawn from this research. These are:

- When representing characteristics of spatial data, using sound to reinforce what is shown visually results in better performance than using vision or sound alone (see Section 7.3.1).
- Participants who have a higher level of knowledge about the data set used in the sonification perform better, showing that they have more capacity to interpret the additional uncertainty information (see Section 7.3).
- Using sounds with a clear and easily understood scale (e.g. piano notes) is more successful than using sound without an integral scale (e.g. bird calls). However, this does depend on the data type, interaction method and individual user preferences (see Section 7.3.2).
- A one-to-one interactive sonification (as used in CS1 and CS2) provides a better user experience and more in depth results than a non-interactive sonification shown to a group of people (as used in CS3, see Section 7.1.3 for more details).

There are also some additional findings, of less importance than the four main ones:

- Continuous data (e.g. UKCP09) and discrete data (e.g. address points and wind turbines) need different techniques to create an effective sonification (see Section 7.2).
- Some people can use sonifications more effectively than others. This may be to do with a person's learning style, or previous knowledge of uncertainty data, the data set or the uncertainty within it (see Section 7.3).
- More experience of using a sonification will result in better performance (see Sections 5.4 and 6.4).
- Using sound to show additional data (as a supplement to what is shown visually) can be done, but needs to be designed very carefully to ensure an effective tool and to not confuse or overwhelm the user (see Section 7.3.2).

The use of three case studies enabled sonification to be used for contrasting applications with different implementation techniques. The fact that some results are common across all three case studies adds weight to them and shows they are likely to be fundamental to any sonification of spatial data. The range of technologies used shows that an effective sonification can be generated in a range of different ways, depending on the data set and resources available.

8.2 Case Study Approach

A decision was taken to research the aims and hypotheses in this thesis using a case study approach, to allow them to be examined in detail and so the most could be gained from the research period. Each case study allowed sonifications of different types of spatial data to be compared. Each case study also took into account the results from the previous case study which allowed unclear results to be explored further. The nature of having three case studies did mean that each case study went into less depth than if the whole thesis was solely dedicated to one evaluation. However the multiple case study approach is suitable for this field because relatively little research and evaluation have been done on the sonification of spatial data to date and this approach allowed a greater variety of information to be gathered. The individual case studies have been reviewed previously and now the advantages and disadvantages of the case study approach will be discussed.

8.2.1 Advantages

The approach of using case studies has a number of advantages. In comparison to a single evaluation it has allowed a greater range of sounds and spatial data to be evaluated. Spatial data comes in a variety of different types, such as object based or continuous, which required different sonification techniques. The different case studies also allowed different sounds to be compared, although this is an element that could be developed further. As the case studies were developed and run sequentially, it allowed later case studies to be informed by earlier case studies which highlighted specific elements that needed to be investigated further. These included elements such as data set and uncertainty knowledge/experience, which were not identified in the literature review. Had a larger, single case study approach been used with only one iteration of testing, it is unlikely this would have been picked up. Multiple iterations of testing within a single case study may have allowed extra variables to be collected in later iterations, but this very much depends on the flexibility of the evaluation plan. Multiple case studies also allowed a number of different evaluation techniques to be used and developed, for example the decision to replace open questions with closed questions after CS1 to allow a greater amount of data to be collected in a limited time (see Section 5.3.4). Additionally, this approach allowed for the comparison of interactive and non-interactive sonifications. The area of sonification of spatial data has had relatively little research so this approach allowed a co-development of techniques and collection of useful data for research development.

While the three case studies covered very different data sets and methods of evaluation, they complemented each other by allowing wider research coverage in this developing area. The iterative design-evaluation process was particularly useful for determining respondent characteristics. The development of the 'learning style' questions progressed significantly, although a conclusive answer with respect to this issue was not identified. Information regarding the participant's previous knowledge was particularly well illustrated with CS2, and therefore the type (and complexity) of the sonification that can be used depends on the users' familiarity with the data being sonified.

8.2.2 Disadvantages

The choice to use three case studies to evaluate the use of sonification with uncertainty in spatial data has numerous advantages (see above) but also has some disadvantages. Using three case studies required significantly more research, development and testing than using a single case study approach. The effort required for this development limited the time that could be dedicated to data collection and evaluation for each case study. For example, time restrictions meant that it was not possible to conduct one-to-one interviews or use think-aloud protocols with individual participants completing the evaluations. The more quantitative approach to the data collection allowed data to be collected from more participants, which increased the statistical validity. The greater numbers and statistical validity are seen as more important than more detailed, in depth (but less statistically valid) results from fewer participants. This is because this area has relatively few evaluations with reasonably large sample sizes (see Section 2.3.2) and would benefit from the statistical validity that the larger sample size brings. While the sample sizes in this research were relatively large compared to previous studies (all cases studies had sample sizes of over 44), there was still on a limit on the amount of analysis that could be completed because of the sample size being too small. This is because there were many different factors that needed to be corrected for, so there were occasions where the number of observations with a particular combination of conditions became too small to run statistically valid tests (where $n < 5$). This would be addressed by a larger sample size in a future study. Some participants did participate in more than one case study, so there is the possibility for a learning effect over the three case studies. However previous participation was recorded in the results and there were no significant differences between those who had participated in previous case studies and those who had not. Additionally, if participants had been prevented from participating in multiple studies, the available pool of participants for each case study would have been much smaller reducing the statistical validity. Using three case studies also enabled a more exploratory approach to be taken to the research, which allowed the latter case studies to be informed by the earlier case studies. This is more appropriate for the area of sonification than one in-depth study because relatively little is known about how uncertainty in spatial data can be represented using sound and the exploratory approach allows a larger number of different methods to be evaluated.

Overall the advantages of the case study approach outweighed the disadvantages as outlined above. The approach is not perfect, but the undeveloped area of sonification of spatial data has benefited more from this approach, which has allowed a wider variety of data to be collected.

8.3 Limitations and Improvements

There are a number of limitations to the evaluations completed in this thesis, some of which could be improved in future research. There are also a number of areas that could be explored in the future, or explored differently as a result of developments in technology. This section will explore these areas, starting with limitations to this thesis.

8.3.1 Limitations

The limitations of this work are split into three sections. Firstly, those limitations resulting from the study design itself and secondly those related to the volunteers who participanted in the evaluations. Finally, issues arising from the use of technology are discussed for each case study.

The number of tasks in each case study limited the amount of information that could be gained from the evaluation sessions. This was a particular issue with the first case study, with only one measure of the participants understanding of the uncertainty information. Additional tasks for the participants would have provided more in-depth information, although there would still have been constraints on the length of the sessions. The other case studies used a different approach to gain more detailed information from the participants by using exclusively closed questions and asking more detailed questions, for example the questions used to ascertain learning style.

The sample size and variation within the sample was also a constraint in certain respects because of restricted variation in some factors. Due to the lack of data in some categories, the types of statistical analysis that could be performed were restricted and for techniques such as crosstabulation there were several occasions where catagories needed to be combined to enable meaningful analysis. The limited sample variation also restricted the amount the results could be generalised to a wider population.

The evaluation sessions were heavily reliant on the technology used working as expected. During the evaluation in CS1 one participant who was experienced with GIS attempted to use the zoom tools available in ArcMap to zoom in on some of the data. Unfortunately this turned off the sonification element and meant that the evaluation had to be restarted for that participant. Other than that interruption it did not affect the evaluation for that participant and no other participants tried to use the zoom controls. Using an existing GIS tool that participants are already familiar with has this disadvantage. CS2 was particularly susceptible to interruptions in Internet access and two evaluations sessions of CS2 suffered because of this. The issue was a major networking equipment failure at UEA, resulting in no network connectivity for two days for the whole university, which could not have been predicted. While this did not impact the results (the sessions were rescheduled successfully) it does highlight the potential problems with this type of evaluation and the reliance on factors beyond the evaluator's control. The other case studies had occasional issues with technology, but only one or two participants were affected and the problems they encountered were easily resolved. While using technology does have disadvantages, the advantages far outweigh these because of the much more detailed information that can be collected than if the technology was not used.

The methods used in the three case studies made the best use of the appropriate technology available at the time for the evaluation. Technology continues to develop at a rapid rate and if any of the case studies were repeated now (particularly CS1 and CS2) they would be completed differently because the techniques used are no longer current.

The interface for CS1 was developed with ArcGIS 9.2 using the VBA (Visual Basic for Applications) capacity to create a custom extension. ArcGIS is now on version 10.1 and while there have been a significant number of interface changes in this new version, the most relevant factor for the first case study is the withdrawal of support for VBA (ESRI, 2010). The VBA add-in still works with ArcGIS 10.x, but a specific license is required (available free from ESRI) to enable this. There is no technical support for VBA with version 10.x and VBA add-ins will not work at all from version 11. ESRI's suggested approach is to re-write the extension in VB .NET (using Visual Studio) and it is also possible to achieve a similar outcome using the Python scripting language within the ArcGIS framework. Currently, the ArcScript developed for this project works in ArcGIS 10.0 (with the appropriate license from ESRI) and a video is also included on the DVD to show how the interface works for when the ArcScript is no longer supported (see Appendix B.4 for details).

CS2 was heavily reliant on the Google Maps interface and as discussed in Chapter 3 the API for the Google Maps interface can change very rapidly (up to every 3 months). The version used in the development and evaluation (3.2) is no longer available, with the oldest available version currently 3.8 and the latest version 3.10. While there are no fundamental changes to the API when changing from one minor version to another and they should be backwards compatible, the application developed for the evaluation would only work with 3.2 and generated errors when used with 3.3. By the time 3.2 was no longer available, all the data had been collected so this was more of an issue for future access than a major problem for data collection. The issue preventing the application from working with version 3.3 appeared to be with the sound element of the program, and this code has been superseded by more up-to-date code as well as being supplemented by the provision of HTML5, as discussed earlier. Google Maps API is still appropriate to use for this type of evaluation tool, but the nature of the rapid turnover must be considered before development and evaluation work is carried out.

CS3 was a much simpler implementation using standard tools for the sonification (Microsoft PowerPoint) and an audience response system that could easily be substituted for a different technology if required. While its simplicity made the development phase of the tool much shorter, it did limit the depth of the data that could be collected. A one-to-one sonification interface would probably generate more useful results, because of the increased participant engagement (see Section 3.1.3).

8.3.2 Future improvements

A number of areas in the case studies could be improved to enable additional insights to be gained from the results and analysis. The study design and sample size could be improved to collect more detailed information. Additional information on learning style and musical experience from the participants are specifically required. In addition, the availability of uncertainty information in spatial data sets could be improved.

An increased sample size would allow for a greater number of repetitions of the evaluation and ensure there were enough participants in each sub-group for any analysis. There are a number of ways of increasing the sample size, for example use of a ‘captive audience’, such as all first year ENV students but this would require more time to collect the data and agreement from the relevant faculty. Alternatively the evaluation period could be longer, allowing time for recruitment of more participants; however this increases the chance of reaching saturation point in the test population, where no more volunteers would be available. Asking participants to give up more of their time (for a longer evaluation session) would require a financial incentive, and therefore require additional funding. A larger sample size and a more targeted sample selection would increase the variation of the sample which would allow statistically meaningful crosstab comparisons for a wider range of factors. It would also provide information on the degree that the results could be generalised to other potential users. Some of the wide variation in performance from participants is explained by different preferences for sonic and visual representation. A larger sample size would allow participants to pick from a selection of different representation methods so their individual preferences could be taken into account and their effectiveness could be evaluated using their preferred sonic and visual representation method.

A number of aspects of the study design could be altered with a larger sample size. For example, participants could be separated into different groups to evaluate different sonifications, which would eliminate any potential learning effects. The study design was limited by the resources available and could be improved by collecting more information on the participants themselves as well as providing a larger variety of tasks for them to complete. The impact of any potential learning effect could be better understood by randomising the order of the task presentations but this would need to be implemented in an effective way with sufficient explanation and training for the participants.

The learning style variable did have an influence on the success of individuals, but the differences were not significant. Certainly for CS2, the comparison between the visual and sonic methods was impacted by the nature of the task being very visual and this is partly relevant for the other two case studies as well. A wider group of participants would improve these results as the participants in this study were biased towards those who have experience of GIS data and software. In addition, a more extensive learning style test would be beneficial, as only two different styles were considered, compared to the five or six that are potentially available. CS3 showed than certain elements of the learning style questions were more useful than others (such as 'You pay attention to the sounds of various things') and it appears that these elements could be used to evaluate the potential effectiveness of sonification. Individual elements of the learning style should be evaluated and if possible, a set of learning styles that differentiate between speech and non-speech audio should be used.

While there was no impact of musical skill on the ability to use the sonification, Parbery-Clark et al. (2012) found that age related hearing loss was reduced for participants who played musical instruments. This study only relates to the perception of speech, but may result in better performance from older participants when using a sonification. Age and level of musical experience should be recorded in future studies to see if this observation can be repeated.

Different types of sonification are also worth investigating. Volume has been shown to be effective in previous research (Bearman, 2006; Jeong and Gluck, 2003; Dubus and Bresin, 2011), but it was decided not to use volume in this research because it is limited in terms of the resolution of the data that can be displayed (Neuhoff, 2011). It is also not independent from other sonic variables (pitch and timbre), which would make comparing the two variables more complex that initially imagined (Neuhoff et al., 2002). As well as volume, spatial sound, timbre, rate of change or duration amongst many others are worth investigating as effective methods of representing uncertainty in spatial data. However, the theoretical information on the different types of sound is very limited, with most previous research about pitch. Future research will push forward our understandings of the different types of sound which will be of great benefit to this type of research. If there is theoretical justification for thinking that factors other than pitch will be useful in this context, they should be tested. Using two sonic variables together to show one data set (e.g. pitch and volume) is also worth investigating although this must be compared to using one sonic variable to see if there is any significant improvement.

One of the main challenges faced by this research (particularly in CS1) was sourcing uncertainty data and/or converting it in to a usable form. The INSPIRE directive from the EU makes provision for metadata standards and this includes provision for uncertainty data (Comber et al., 2006). However, current implementation of this directive is limited in academic circles and a strict, well defined data gathering and recording method for spatial data is difficult to introduce within a fragmented, dispersed and varied academic environment. If the standards are implemented and used, it will make uncertainty data much more accessible and available and may provide a framework for including uncertainty as an element of data to be represented in the majority of data sets. Accessibility of uncertainty data is key and must be reviewed in any future research on this issue if the concept of sonification of uncertainty in spatial data is to be able to be applied in the real world in any meaningful way.

Sonification is a very young field, particularly with reference to spatial data as limited work has been completed in this area. The types of sound and interface used are very specific to the task and data in use and so far it has been quite difficult to establish broad rules for relationships between the two. This is in part due to the relatively high technical level required to create a sonification and the community previously being relatively insular with limited examples from outside the core music / psychology areas. However, this is beginning to change with a much wider spread of examples becoming available in the last few years. This is because the technology to create sonifications has become much simpler to use and more accessible, particularly in the last two decades. This will continue to grow and diversify in the future, which will only benefit the sonification and wider communities. At present, this technical requirement also restricts development of sonifications to those with the necessary programming skills.

8.4 Implications of the Research Findings

The implications of this research for the presentation and use of uncertainty data, for sonifying uncertainty data, for GIS and for sonification are discussed in detail in the next four sections. There are also some implications for the other areas considered in this research, specifically HCI, psychology and virtual environments. Researchers from the HCI community can take from this research the fact that the user's knowledge of the data set being used is very important in a sonification, as well as the fact that accommodating individual user preferences (in this research, different sounds) is likely to increase performance. There is also a need for more interaction between the sonification and HCI community, particually in terms of developing a conceptual model that is useful to both areas (see Section 1.3 for details) although this thesis starts the process of developing such a model. Psychology researchers may find the evaluation of the use of sonification in this setting useful as it is one of a few studies that evaluate the use of sound to represent uncertainty in spatial data with a larger sample. They may also be interested in the conceptual model that was created as a result of this research. The specific finding of the importance of participant knowledge of the data set being sonified is also important. Researchers working to create virtual environments should take note of the finding that the use of sound to reinforce data shown visually improves performance, and this agrees with other research in virtual environments that the use of sound and vision together will create a more realistic environment (Storms and Zyda, 2000).

8.4.1 Implications for presentation and use of uncertainty data

Use of uncertainty data within spatial data applications can be very patchy, often because of the limited availability of uncertainty information in a usable, accessible format. Assuming the uncertainty data are available, this research is relevant to the areas of uncertainty data presentation and uncertainty data use, shown in the conceptual model as the design and use phases (see Section 7.4 for details). One important aspect to remember is that a users understanding of the uncertainty in the data presented will depend on their previous experience (of both the data set and any uncertainties surrounding it). Sonification is a new way of representing uncertainty in spatial data and while it has been tried before it has not been taken up in any significant way because of technological limitations and the complexity of sonification tools. This research does push forward the provision of tools for spatial data, but is not yet ready for more significant commercial or academic user take up, partly due to limits in understanding how different people use sonifications. The implementation of a sonification is still very technical and specific to a particular area, making it very difficult to create a sonification toolkit that can be used with many different data sets. The technical level is less of an issue for GIS users, who have a relatively high technical knowledge, but it still requires a significant investment of time to generate a sonification and not many users have this time available. This research shows that using sound to reinforce spatial information shown visually (including uncertainty) will result in better performance. There is also the potential to show additional information using sound, but the process to do this effectively is more complex and success is dependent on many more variables. If some of these sonification techniques are adopted, then it will make uncertainty information easier to understand and use. Along with increased availability of uncertainty data with developments such as those pioneered by UKCIP and those outlined within the EU INSPIRE programme, the use and understanding of uncertainty data should increase. This study does create a methodology for representing uncertainty in spatial data which can be applied to a number of different data sets easily. The technology used is easily accessible and will make use of sonification in a GIS environment easier for any subsequent user.

8.4.2 Implications for understanding sonified uncertainty data

This research highlights the importance of previous experience and understanding of uncertainty data, as shown in the evaluation section of the conceptual model. Participants who had a good level of understanding of the data, particularly those who understood the uncertainty element, performed more effectively in the evaluation and this is particularly relevant to the more complex data sets such as UKCP09 where uncertainty is an inherent part of the data set. This is a good example of the impact previous experience can have, as explained in a general way by Brunet and Cowie (2012, see also Figure 22 in this thesis). The study also shows that musical experience has no impact on ability to use sonifications, consistent with findings from Walker and Nees (2011). A participant's learning style may have some impact, but any particular trends that do exist were not seen in this research. A more detailed and targeted evaluation of learning styles would provide more information. A learning effect was also seen, with repeated usage improving performance, consistent with the literature (for example, Flowers et al., 1996). This study also developed a methodology for evaluating these factors that could be applied in a number of different testing environments.

8.4.3 Implications for GIS

GIS and cartography have always been about how data are shown to the user and in GIS technological developments have driven significant changes in the GIS interfaces and data availability (e.g. server-side to client-side, black and white to colour, increasing resolution and coverage of rasters, vector based data formats, web based interfaces for users and provision of data & information to the general public through tools like Google Earth). Sonification is another technological development that allows the representation of spatial data in new and different ways. This research shows how the use of other senses in combination with vision can enhance the experience of data by GIS users. While it is not the first GIS sonification that has been created, it is one of a handful with a comprehensive evaluation (see Section 2.3.2) which highlights a number of factors that are important when creating these types of sonifications (see Section 7.4). These include the importance of previous user experience of the data set, as well as specific user preferences for different types of sonifications. These both need to be considered when desiging a sonification for a GIS audience. This research also explores the different technological solutions to creating these sonifications and evaluations, including the use of the Google Maps API which has received significant subsequent interest. The research therefore shows that sonification can be used effectively with GIS and spatial data and provides information for subsequent users on how best to apply the sonification technology to uncertainty in spatial data.

8.4.4 Implications for sonification

Sonification is a rapidly developing area in which a wide range of work is taking place, much of this away from the traditional sonification areas such as psychology and music (Dubus and Bresin, 2011; Constantinescu and Schultz, 2011). This area of development has had very few pieces of work involving spatial data with an evaluation element such as the ones in this thesis, so the research represents a new contribution to the sonification field. Another major contribution of this thesis is the conceptual model (see Figure 58, page 229) which provides a framework of important aspects to consider when constructing these types of sonifications, as well as the evaluation methods developed for this research. The most important aspect is the user's knowledge of the data set being sonified. In the context of the examples in this thesis, as well as knowledge about the data set, knowledge about the uncertainty element of the data was very important, as well as how the uncertainty data is encoded within the data set. In addition, using sound to reinforce information shown visually will result in better performance, as will more experience of the sonification itself. The research also constitutes a contribution to the technology and methods of sonification in different contexts, which is growing rapidly. It particularly contributes to the evaluation of such techniques, which in the past has been limited because of few applications in this area. Throughout this thesis is has been noted that while many prototype sonification tools have been created, very few have had any significant user testing above 2 or 3 participants, often the author's office colleagues. The type of evaluations conducted in this thesis with larger evaluation groups are crucial to allow a meaningful literature to be constructed for the development of sonification applications.

Appendices

A Additional Material DVD

The DVD contains additional material for this thesis, including video and sound clips of the audio elements. They are referenced individually in the main text of the thesis, and a list of the contents is included below. This is also included in a readme.txt file on the DVD.

- `\Chapter 2 – CS1`
  - `\Program` – The ArcScript used to collect data for this evaluation, within Map Documents for ArcMap 9.2 (`Project92.mxd`) and ArcMap 9.3 (`Project93.mxd`).
  - `200911 Briefing Document UEA.doc` – Briefing Document given to UEA Participants
  - `200912 Briefing Document LA.doc` – Briefing Document given to Local Authority Participants
  - `20091026 Briefing Document OS.doc` – Briefing Document given to OS Participants
  - `Participant Information Sheet & Consent Form.doc` - Participant Information Sheet & Consent Form given to all participants on arrival
  - `Short Video.avi` – A video of the sonification element of the evaluation. A similar version is available at http://vimeo.com/14502978.
  - `Video Complete Run Through.avi` – A video of a participant working through the whole evaluation
- `\Chapter 3 – CS2`
  - `\Evaluation Code` – Code and related files for the CS2 web based survey. See Sonification_Readme.rtf for more details.
  - `\Videos`
    - `Complete Run Through Video.avi –` A video of a participant working through the whole evaluation
    - `S (visual & sonic, same data, 1 map).mwv –` A video of the S map interface
    - `VSVS (visual & sonic, 2 data sets, 2 maps).mwv –` A video of the VSVS map interface

        - `VS (visual & sonic, 2 data sets, 1 map).mwv` – A video of the VS map interface
    - `Briefing Document OS.doc` – Briefing Document given to OS Participants
    - `Briefing Document UEA.doc` – Briefing Document given to UEA Participants
    - `Briefing Document UKCIP.doc` – Briefing Document given to UKCIP Participants
    - `Participant Information Sheet & Consent Form.doc` - Participant Information Sheet & Consent Form given to all participants on arrival
- `\Chapter 4 - CS3`
    - `Briefing Document OS.doc` – Briefing Document given to OS Participants
    - `Briefing Document UEA.doc` – Briefing Document given to UEA Participants
    - `Briefing Document UKCIP.doc` – Briefing Document given to UKCIP Participants
    - `Participant Information Sheet & Consent Form.doc` - Participant Information Sheet & Consent Form given to all participants on arrival
    - `EvaluationDemo.pptx` – A PowerPoint file demonstrating some of the visualisations and all of the different methods used in the evaluation.
    - `WholeEvaluation.pptx` – The PowerPoint file used for the whole evaluation. Contains all the visualisations shown to the participants.
- `Readme.txt` – a file outlining the contents of the DVD

B Case Study 1

B.1. Briefing Document

A two page briefing document was emailed to all respondents at least 24 hours before the session. Contact details for the gatekeeper were provided as appropriate for the Ordnance Survey / UEA. A copy is available at \Chapter 2 – CS1 on the DVD.

<u>Briefing Document for evaluation at UEA during November 2009</u>

I am looking different ways of representing spatial data, particularly uncertainty data. I'm specifically interested in how the Address Layer 2 data set is used by people and how the uncertainty values within this dataset are handled and displayed. You do not need to know any details about the Address Layer 2 data set – all information will be provided.

Currently the vast majority of spatial data is represented visually, usually in some form of map, either paper or computer based. There can be limits to visual representation, particularly when trying to show additional information on the map as this can result in visual saturation, when no more data can be displayed on the map. One particular data set that suffers from this issue is uncertainty, and it is often not represented at all (Aerts et al. 2003).

One way of representing additional information is to use sound in combination with the traditional visual stimuli. There have been various pieces of research on using sound within a spatial data environment, both in terms of theoretical discussion (Brewster 1994; Krygier 1994) and practical applications (Fisher 1994; Lodha et al. 1996; MacVeigh & Jacobson 2007). I worked on a related area in my MSc GIS at University of Leicester, using sound with a number of different examples of data, and I gained very valuable feedback from volunteers at the Ordnance Survey.

For my PhD, I am looking at how visual and sonic methods can be used to represent uncertainty using three different types of data. These are vector based data, raster based data and virtual reality simulations. I am starting with the vector based data and using the example of the Ordnance Survey's Address Layer 2 product. This is a product that is used for a variety of purposes by many different users. I am specifically interested in comparing different ways of representing the uncertainty values (i.e. status flags / status values) particularly comparing visual and sonic methods.

I have created an evaluative prototype which will take you through a series of exercises, showing different ways of representing the uncertainty and asking questions about it. You do not need to have extensive knowledge about the Address products, as you will be told all you need to know as a part of the evaluation.

Your answers will be completely anonymous and there will be no way to identify your individual responses. You are free to withdraw from the study at any time before the questionnaire is completed. The results will be used for my PhD and any related publications, and will also be shared with the Ordnance Survey as they are part funding my PhD. No names (personal or company) will be reported in the results. You will also have the opportunity to be informed of the results of my research, and/or kept up to date with my ongoing research if you wish.

If you have any questions about the study, please ask either myself, or Jenny Harding at the Ordnance Survey, or my supervisor at UEA, Andrew Lovett (contact details below).

I hope I have given you a bit of background information on my project and the evaluation that will take place. I very much look forward to seeing you on the 26th / 27th October, and if you have any questions please feel free to call me on 07717745715 or email me on n.bearman@uea.ac.uk.

Contact Details:

Nick Bearman, PhD Student
E-mail: n.bearman@uea.ac.uk Tel: 07717745715 or 01603 591346
School of Environmental Science, UEA, Norwich, NR4 7TJ.

Andrew Lovett, Supervisor
E-mail: a.lovett@uea.ac.uk Tel: 01603 593126
School of Environmental Science, UEA, Norwich, NR4 7TJ.

Jenny Harding, Senior Research Scientist
E-mail: Jenny.Harding@ordnancesurvey.co.uk Tel: 023 8079 2052
Ordnance Survey Research, c530, Romsey Road, Southampton, SO16 4GU.

References:

Aerts, J.C.J.H., Clarke, K.C. & Keuper, A.D., 2003. Testing Popular Visualization Techniques for Representing Model Uncertainty. Cartography and Geographic Information Science, 30, 249-261.

Brewster, S.A., 1994. Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces. Human-Computer Interaction Group, Department of Computer Science, University of York, Available at:
http://www.dcs.gla.ac.uk/~stephen/ThesisIntroAbs/thesisIntro.html.

Fisher, P.F., 1994. Hearing the Reliability in Classified Remotely Sensed Images. Cartography and Geographic Information Systems, 21(1), 31-36.

Krygier, J.B., 1994. Sound and Geographic Visualization. In Visualization in Modern Cartography. Oxford, UK: Elsevier Science, pp. 149-166.

Lodha, S.K., Wilson, C.M. & Sheehan, R.E., 1996. LISTEN: sounding uncertainty visualization. Available at: http://portal.acm.org/citation.cfm?id=245053 [Accessed March 4, 2008].

MacVeigh, R. & Jacobson, R.D., 2007. Increasing the dimensionality of a Geographic Information System (GIS) Using Auditory Display. Available at: http://www.music.mcgill.ca/icad2007/proceedings.php [Accessed May 25, 2008].

B.2. Participant Information Sheet and Consent Form

This document was given to the users when they arrived for the evaluation. They were asked to read, sign one copy and return to me. They were given a second copy, to keep if they wished. A copy is available at \Chapter 2 – CS1 on the DVD.

<u>Participant Information Sheet & Consent Form</u>

This information sheet explains about the research I am conducting. Please read, sign one copy and return to me. If you have any questions, please ask.

I am interested in different ways of representing Ordnance Surveys Address Layer 2 data, specifically how the uncertainty / positional accuracy status flags within the data are shown. Typically this information is not well utilised and I want to see if using different ways of representing positional accuracy makes these details more accessible.

I am comparing different sonic and visual methods to see which is the most effective to represent the data. To measure this, I will be showing you a series of maps and ask you to say what levels of positional accuracy the data has. I will also ask a number of background questions to understand the representativeness of my sample and to see if those factors have an impact on preference for different methods.

If you have any questions at any point, or would like more information then please ask. Your results will be recorded electronically and stored anonymously with no way to link them to you. The results will be used in aggregate analysis and any text comments you make may be used, but there will be no way to link them to a specific participant.

If you wish to end your participation in the research, you may stop at any point throughout the questionnaire and your results will not be used in the analysis. However, after you leave the room it will not be possible to remove your results from the analysis as they are anonymous and cannot be identified.

This research will be used as a part of my PhD and any related publications or conference papers. The Ordnance Survey will also receive a copy as they partly fund my PhD. If you wish to be kept up to date with my research and where this data is used, there is the option to supply your email address in the evaluation. This will be stored separately to any results so that your email address cannot be linked to your results.

If you have any questions or would like further explanation, please ask.

Signed: ___________________________________ Date:__________________

If you would like further information both mine and my supervisors contact details are below:

Nick Bearman, PhD Student
E-mail: n.bearman@uea.ac.uk Tel: 07717745715 or 01603 591346
School of Environmental Science, UEA, Norwich, NR4 7TJ.

Andrew Lovett, Supervisor
E-mail: a.lovett@uea.ac.uk Tel: 01603 593126
School of Environmental Science, UEA, Norwich, NR4 7TJ.

Jenny Harding, Senior Research Scientist
E-mail: Jenny.Harding@ordnancesurvey.co.uk Tel: 023 8079 2052
Ordnance Survey Research, c530, Romsey Road, Southampton, SO16 4GU.

B.3. Questionnaire from the Evaluation

The introduction text and questions from CS1 follow. See Appendix B.4 for more information on the program itself.

Thank you for taking part in this research. You should have already received a participant information sheet, explaining the research I am doing and have signed a participant consent form. If you have not done both of these, please ask for help now.

You can progress through the survey by clicking the 'Next' button, which you will find at the bottom right of this window. Whilst you are working with the map, this window will shrink and move to the top of the screen. When you have finished with the map, click 'Continue' and this window will return to the previous position.

The survey should take no more than 30 minutes to complete. If, at any point, you have any problems or questions please ask.

Click 'Next' to continue.

All questions require an answer unless specified

2a Musical Experience
Can you play (or have you ever been able to play) any musical instruments?
Yes/No

2b Musical Exposure
Do you listen to music on a regular basis (more than approx 7 hours per week)?
Yes/No

2c Colour Blindness
Do you suffer from any type of colour blindness? If you don't know, then please choose no.
Yes/No

3 Learning Style

3a Please state which type of lecture out of these two options you would prefer to attend:
A lecture whose main content is presented visually through diagrams and pictures
A lecture whose main content is presented verbally by the lecturer

3b You own a camera and want to improve the photos you take. Please state which of these two options you would prefer:

Look at lots of examples of photos with information on how to improve them

Speak to someone who knows about photography and ask them for advice

4a How well do you know Ordnance Surveys Address data products? This could be Address Layer 2, Address Layer or Address Point.

1 Know it very well (use it on a daily basis or completed extensive work with it)

2 Know a reasonable amount about it

3 Know some information about it

4 Heard of it but know little / nothing about it

5 Never heard of it

4b Have you been involved with any of my previous work?

1 Yes - A telephone interview

2 Yes - An in-person interview with demo software

3 Yes - An in-person interview without demo software

4 No – Have not been involved before

5 Introduction to Demo

This evaluation uses Ordnance Survey MasterMap data, specifically Topography and Address Layer 2 data. The topography layer is a physical representation of the real world. It shows the roads, buildings, gardens, woods, fields and so on in Great Britain (see example of a street of terraced houses in Norwich, left).

The Address Layer 2 data consists of a postal address for each residential or commercial building. It shows the spatial location of this address as an Address Point, as well as linking it to a particular building in the Topography layer using the Ordnance Survey topographic identifier (TOID).

There is a level of uncertainty in the Address Point, called Positional Accuracy, which we will be exploring. It represents how likely it is that the spatial location of the address point is within the correct building polygon. It can be one of three values, which are (in order from highest accuracy to lowest accuracy):

- Surveyed (highest)
- Postcode Unit Mean
- Estimated (lowest)

These will be shown on the map using colour or sound. Click on the buttons to the right to see examples of which colours and sounds will be used. When you have seen them, click 'Next' to continue.

These accuracy values show how likely it is that the Address Points are located in the correct building polygon. For example, the spatial location of an Address Point with a Surveyed value is likely to be within the correct building, while the spatial location of an Address Point with Estimated value is unlikely to be within the correct building.

Below are a couple of questions to check you understand the different accuracy values. If you want to look at the information again, click 'Back' to view the previous page.

6a Is it likely or unlikely that an address point with a positional accuracy value of Surveyed is within the correct building polygon?
Likely/Unlikely

6b Is it likely or unlikely that an address point with a positional accuracy value of Estimated is within the correct building polygon?
Likely/Unlikely

When you click 'Next', you will be shown an example map that shows how sound can be used to represent different positional accuracy values. Each building has an positional accuracy value which is represented by a sound when you move the mouse cursor over the building.

A legend will appear, and plays the sound used for a particular positional accuracy level when you click the relevant button. Please experiment with this, and if you cannot see the legend or have any questions over accessing and understanding the legend or data, please ask.

For this exercise you need to familiarise yourself with the map, data and sounds.

When you click 'Next', this window will shrink and move to the side of the screen. It will contain a 'Continue' button. When you are happy with the legend and data, click 'Continue' to carry on with the evaluation.

When you click 'Next' you will be shown a series of maps, some with sound and some without. After each, you will be asked what proportion of the data has the highest accuracy value (i.e. Surveyed).

The options will be: 25%, 50% or 75%. Please choose whichever you think is nearest to the correct proportion. If you have absolutely no idea, then please choose don't know. The proportions will change for each map you see, and make sure you look at the whole of the screen before answering the question.

Visual

The next example map uses colour to show the positional accuracy values. Each building has a colour, and the legend is shown on the left hand side.

Sonic

The next example map uses sound to represent the positional accuracy values. Each building has a sound which will play when you move the mouse over it.
A legend will appear, and will show which sound is associated with which value when you click on the relevant button. This is the same method as shown in the demonstration.

Visual & Sonic (same)

The next example map uses both colour and sound to show the positional accuracy values.
Legends for both colour and sound are available as before.

Visual & Sonic (different)

The next example map uses sound to show the positional accuracy values for the Address Points, as before.
In addition, it uses colour to show different Council Tax bands, from Band A (cheapest) to Band H (most expensive). Bands A & B are grouped together, as are Bands F, G & H because there are not many houses in these bands.
As well as being asked about the proportion of values that have the highest positional accuracy value, you will also be asked what proportion of houses are in Council Tax Band A & B. The possible options are 25%, 50% or 75%.

10a, 13a, 16a, 20a Estimate the proportion of buildings that have the highest positional accuracy value (i.e. Surveyed)
25%/50%/75%/Don't know

20d Estimate the proportion of houses in Council Tax Band A&B
25%/50%/75%/Don't know

11b, 14b, 17b, 21b How well does this approach convey the positional accuracy levels of the data for a specific building?
Very well + (1)/(2)/(3)/(4)/(5)/(6)/Not well at all - (7)

11d, 14d, 17d, 21d How well does this approach convey the overall positional accuracy of the data in this area?
Very well + (1)/(2)/(3)/(4)/(5)/(6)/Not well at all - (7)

22a How well does this approach convey the Council Tax data overall?

Very well + (1)/(2)/(3)/(4)/(5)/(6)/Not well at all - (7)

11c, 14c, 17c, 22c How easy did you find it to answer the questions with this example?

Very easy + (1)/(2)/(3)/(4)/(5)/(6)/Very difficult - (7)

18a When looking at the previous example that utilised vision and sound, did you rely more on the visual, more on the sound, or use both equally?

Relied entirely on visual (v)

Relied on both vision and sound, but used visual more (vs)

Relied on both vision and sound equally (e)

Relied on both vision and sound, but used sound more (sv)

Relied entirely on sound (s)

That was the last map to look at. Now there are just a few more questions.

23a Overall, which representation method was hardest to understand?

1 Visual

2 Sonic

3 Visual & Sonic (both showing positional accuracy)

4 Visual & Sonic (visual showing Council Tax, sonic showing positional accuracy)

23b Overall, which representation method was easiest to understand?

1 Visual

2 Sonic

3 Visual & Sonic (both showing positional accuracy)

4 Visual & Sonic (visual showing Council Tax, sonic showing positional accuracy)

23c Why did you find these ones the easiest/hardest?

Open text response

24a How did the use of sound affect your interaction with the data compared to visual methods?

Open text response

24b This application altered the pitch of the sound to represent different information. Out of all of the possible options, which do you think would be the best aspect of sound to alter?

Pitch/Volume/Tempo/Other (specify in Why)

24c Why?

Open text response

25c As you progressed through the evaluation, did understanding the data and sounds get easier?
Yes/No – it got harder/No – there was no change/Other

25a Based on everything you have done and said so far, do you think it would be worth developing this as a commercial application?
Yes/No

25b Why?

26a Many users of the Address Layer 2 data set do not make use of the positional accuracy data that is available. What do you think could be done to improve the usage of this data?
Open text response

26b Would you like to be emailed a summary of the research when it is completed and/or kept up to date with the progress of my research? You can unsubscribe from this at any time.
Summary of research only/Summary and updates/No

26c Enter email address

Thank you very much for taking part in this research. Your input is very much appreciated and forms an important part of my project. Please click 'Finish' to end the evaluation.

B.4. ArcScript Details

The evaluation was run from within ArcMap itself, using the ArcScript ability to program a questionnaire interface. The Map documents are included but may not work as expected with ArcGIS 10. A VBA license will be required, but this is free to obtain from ESRI. A video is also included (in the \Chapter 2 – CS1\ folder on the DVD) which shows the sonification element of the evaluation (Short Video.avi) and a run-though of the whole evaluation (Video Complete Run Through.avi).

C Case Study 2

C.1. Flowcharts and Programming Details

Once the questionnaire was finalised and the map fully tested, the two elements were brought together. The web based section of the evaluation was designed to fit around the introduction at the beginning and the discussion session at the end. Introductory text was provided, allowing the users to work through the evaluation at their own rate, once the session was introduced.

The 'online' version was created to allow participants who were difficult to access geographically (including the UKCP09 users) to complete the evaluation independently. Contact details were provided if the users had any difficulty completing the evaluation.

The programming code used is available in the attached DVD in '\Chapter 3 – CS2\Evaluation Code'. The flow charts below show how the program operates.

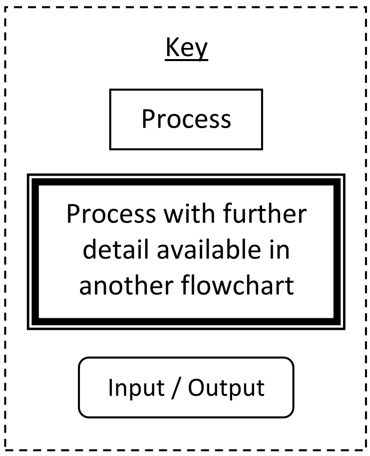


Figure 59. Key for the following flow charts.

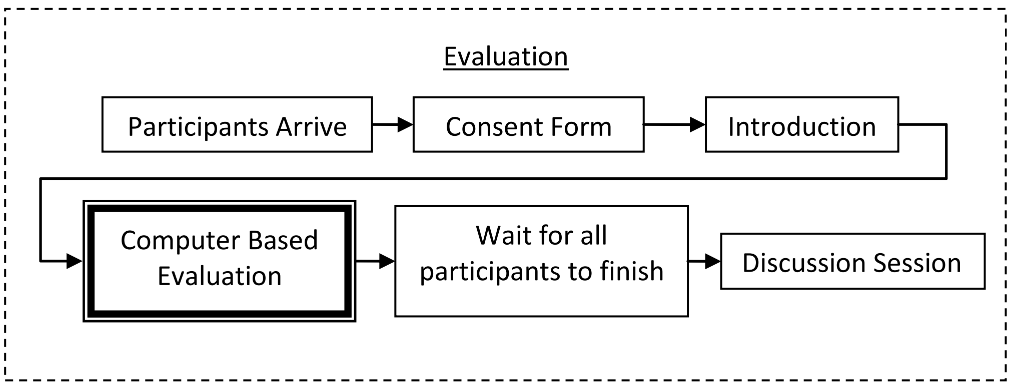


Figure 60. Flow chart of the overall evaluation process. A copy of the consent form is available in Appendix C.4 (page 262).

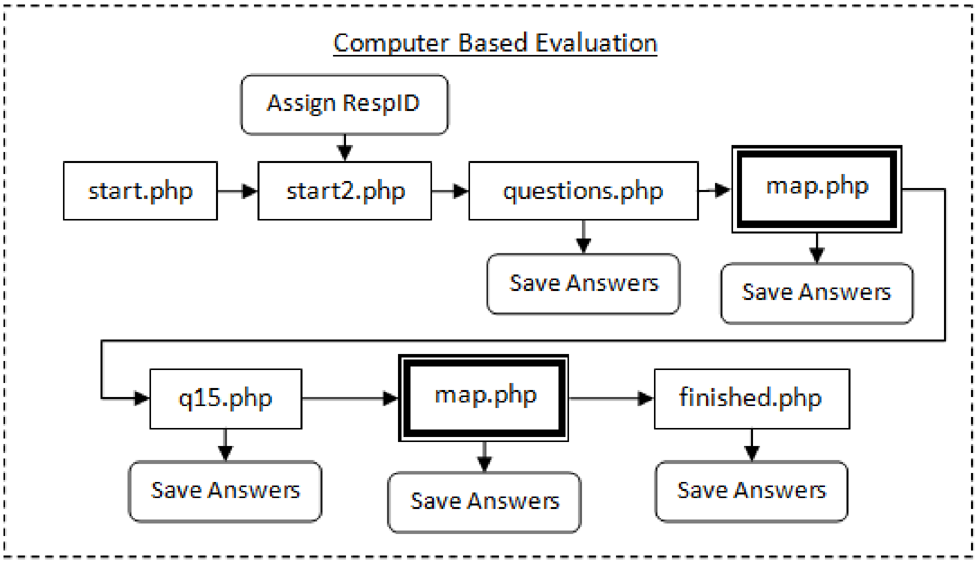


Figure 61. The flow of the computer based evaluation, operated by the participants on an individual basis. The individual PHP files are available on the DVD (\Chapter 3 – CS2\Evaluation Code) and a further flow diagram is provided for map.php.

Map.php

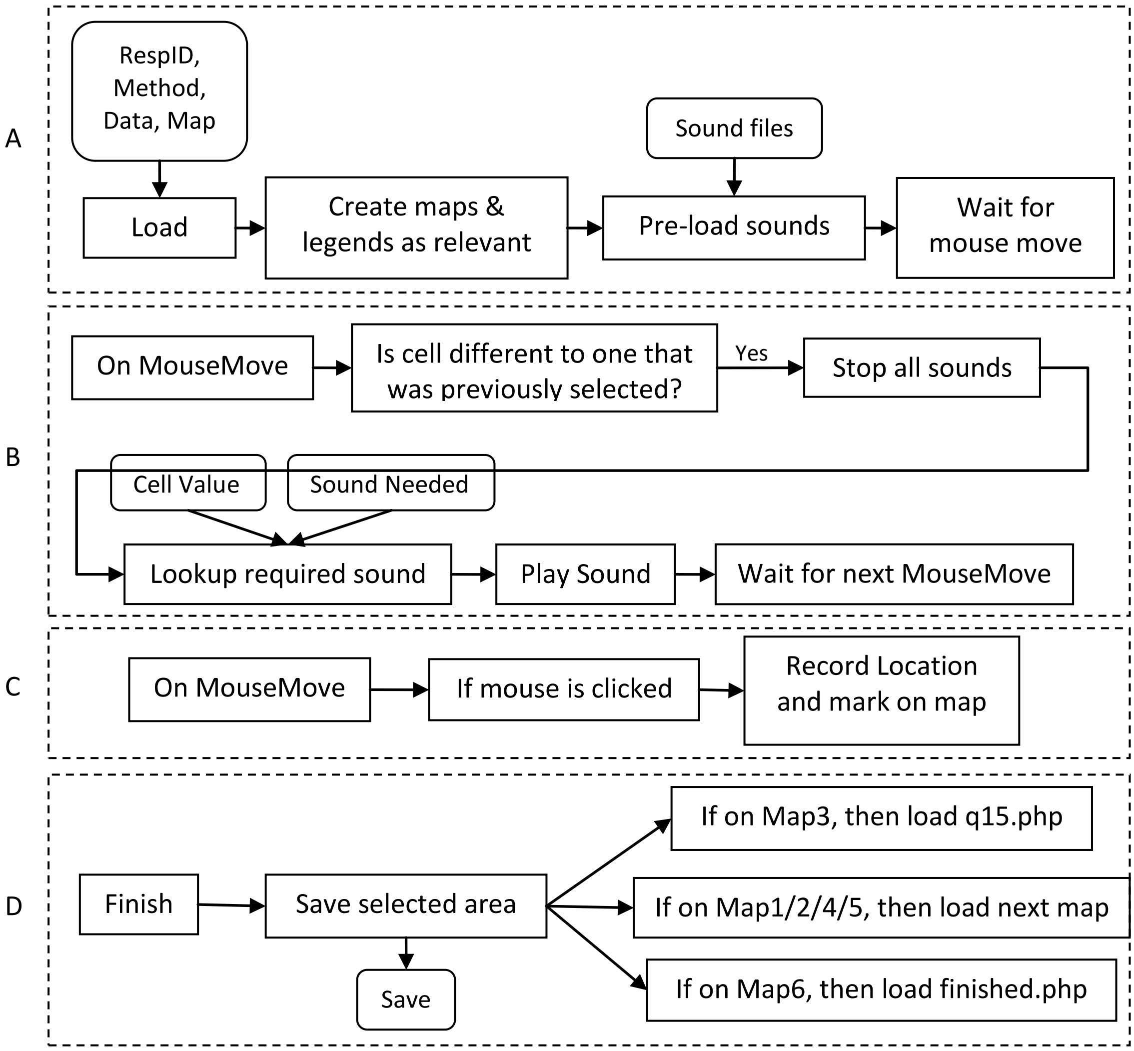


Figure 62. Flow charts for the main file in the evaluation, map.php. 'A' shows the process of loading the maps and pre-loading the sound files. 'B' shows the process for playing the sound when the mouse is moved over the map. 'C' is the process for recording the selected area. 'D' is the process when the user has finished the task, and moves on to the next map or ends the evaluation.

C.2. Google Maps API

The Google Maps API had a KML class, which was relatively under developed. The class was able to display KML layers, and could return the value in a KML object when it was clicked, but it was not possible to return a value when the mouse moved over the KML object, which was what was required for this software. Instead, the data was hard coded into the JavaScript, as a series of arrays (see code saved on the DVD at \Chapter 3 – CS2\Evaluation Code, specifically LoadArray.js).

C.3. Briefing Document

A two page briefing document was emailed to all respondents at least 24 hours before the session. Contact details for the gatekeeper were provided as appropriate for the Ordnance Survey / UEA / UKCIP. A copy is available at \Chapter 3 – CS2 on the DVD.

<u>Briefing Document</u>

<u>for UEA November & December 2010</u>

I am currently studying for a PhD in Environmental Science looking at how spatial data can be represented using sound. I'm particularly interested in how the uncertainty in the UKCP09 projections data is handled and displayed.

Currently the vast majority of spatial data is represented visually, usually in some form of map, either paper or computer based. There can be limits to visual representation, particularly when trying to show additional information on the map as this can result in visual saturation, when no more data can be displayed on the map. One particular data set that suffers from this issue is uncertainty, and it is often not represented at all (Aerts et al. 2003).

One way of representing additional information is to use sound in combination with the traditional visual stimuli. There have been various pieces of research on using sound within a spatial data environment, both in terms of theoretical discussion (Brewster 1994; Krygier 1994) and practical applications (Fisher 1994; Lodha et al. 1996; MacVeigh & Jacobson 2007).

For my PhD, I am looking at how visual and sonic methods can be used to represent uncertainty using three different case studies. The first case study looked at using sound to represent uncertainty in Ordnance Survey Address Layer 2 data last year. This is the second case study, where I am comparing different visual and sonic methods used to represent the uncertainty within the UKCP09 data set.

I have created an evaluative prototype which will take you through a series of exercises, showing different ways of representing the data and asking you to highlight specific areas. I am looking for people with different levels of knowledge about the UKCP09 data and GIS systems, and I am also collecting data from staff at the Ordnance Survey, the staff at UKCIP and users of the UKCP09 data. As a 'thank you' for taking part, you have the option of entering a prize draw for a £10 shopping voucher.

Your answers will be stored securely at UEA, will be completely anonymous and there will be no way to identify your individual responses. You are free to withdraw from the study at any time before the questionnaire is completed. The results will be used for my PhD and any related papers or publications. No personal or company names will be reported in the results. You will also have the opportunity to be sent a copy of the results of my research, if you wish.

If you have any questions about the study, please ask either myself or my supervisor, Andrew Lovett (contact details below).

I hope I have given you a bit of background information on my project and the evaluation that will take place. I very much look forward to seeing you at the session, and if you have any questions please feel free to call me on 07717745715 or email me on n.bearman@uea.ac.uk.

Contact Details:

Nick Bearman, PhD Student
E-mail: n.bearman@uea.ac.uk Tel: 07717745715 or 01603 591346
School of Environmental Science, UEA, Norwich, NR4 7TJ.

Andrew Lovett, Supervisor
E-mail: a.lovett@uea.ac.uk Tel: 01603 593126
School of Environmental Science, UEA, Norwich, NR4 7TJ.

References:

Aerts, J.C.J.H., Clarke, K.C. & Keuper, A.D., 2003. Testing Popular Visualization Techniques for Representing Model Uncertainty. Cartography and Geographic Information Science, 30, 249-261.

Brewster, S.A., 1994. Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces. Human-Computer Interaction Group, Department of Computer Science, University of York, Available at: http://www.dcs.gla.ac.uk/~stephen/ThesisIntroAbs/thesisIntro.html.

Fisher, P.F., 1994. Hearing the Reliability in Classified Remotely Sensed Images. Cartography and Geographic Information Systems, 21(1), 31-36.

Krygier, J.B., 1994. Sound and Geographic Visualization. In Visualization in Modern Cartography. Oxford, UK: Elsevier Science, pp. 149-166.

Lodha, S.K., Wilson, C.M. & Sheehan, R.E., 1996. LISTEN: sounding uncertainty visualization. Available at: http://portal.acm.org/citation.cfm?id=245053 [Accessed March 4, 2008].

MacVeigh, R. & Jacobson, R.D., 2007. Increasing the dimensionality of a Geographic Information System (GIS) Using Auditory Display. Available at: http://www.music.mcgill.ca/icad2007/proceedings.php [Accessed May 25, 2008].

C.4. Participant Information and Consent Form

This document was given to the users when they arrived for the evaluation. They were asked to read, sign one copy and return to me. They were given a second copy, to keep if they wished. A copy is available at \Chapter 3 – CS2 on the CD.

Participant Information Sheet & Consent Form

This information sheet explains about the research I am conducting. Please read, sign one copy and return to me. If you have any questions, please ask.

I am exploring different methods of communicating the uncertainty within the UKCP09 projection data. This uncertainty information is relatively new to the UKCIP data series, and the visual methods for communicating it are limited. I am looking to see whether sound can be used to communicate it, instead of the traditional visual methods.

The evaluation will compare different sonic and visual methods to see which is most effective at representing the data. To measure this, I will be showing you a series of maps and ask you to highlight particular areas. I will also ask a number of background questions to understand the representativeness of my sample and to see if those factors have an impact on whether participants highlight the correct area. After the computer based evaluation, we will have a short discussion session on your feedback and thoughts of the use of sound.

If you have any questions at any point, or would like more information then please ask. Your results will be recorded electronically and stored anonymously with no way to link them to you, and the results will only be used in aggregate analysis. The discussion session will be recorded to help me type up notes from the sessions. It will not be used for any other purpose and will be stored securely.

If you wish to end your participation in the research, you may stop at any point throughout the questionnaire and your results will not be used in the analysis. However, after you leave the room it will not be possible to remove your results from the analysis as they are anonymous and cannot be identified.

This research will be used as a part of my PhD and any related publications or conference papers. The Ordnance Survey will also receive a copy of my thesis as they partly fund my PhD. You will be offered an opportunity to receive the results of the research and/or enter a prize draw to win a £10 shopping voucher. Your email address will be stored separately to the evaluation answers, and there is no way to link them together.

If you have any questions or would like further explanation, please ask.

Signed: ___________________________________ Date:__________________

If you would like further information both mine and my supervisors contact details are below:

Nick Bearman, PhD Student
E-mail: n.bearman@uea.ac.uk Tel: 07717745715 or 01603 591346
School of Environmental Science, UEA, Norwich, NR4 7TJ.

Andrew Lovett, Supervisor
E-mail: a.lovett@uea.ac.uk Tel: 01603 593126
School of Environmental Science, UEA, Norwich, NR4 7TJ.

C.5. Questionnaire

The following questions were asked before the users were shown the evaluation maps.

1. On a scale of 1-5 (with 1 being high) how do you rate your knowledge of climate change in general?
   1 - I know a lot about climate change / 2 / 3 - I know a reasonable amount about climate change / 4 / 5 - I know nothing / very little about climate change
2. On a scale of 1-5 (with 1 being high) how do you rate your knowledge of GIS / computer based maps and spatial data in general?
   1 - I know a lot about GIS / computer based spatial data / 2 / 3 - I know a reasonable amount about computer based spatial data / 4 / 5 - I know nothing / very little about computer based spatial data
3. Are you a member of the Royal Meteorological Society?
   Yes / No
4. How long have you been aware of the UKCP09 data set?
   Not before today / 1 month / 6 months / 1 year / From when it was launched (June 2009) / From before it was launched

5. Have you used the UK Climate Projections website (image right) which contains information on what the projections are and how to use them? http://ukclimateprojections.defra.gov.uk/)
   Yes / No
6. Have you been to the PiP (Projections in Practice) or any other training organised by UKCIP?
   Yes / No
7. Have you looked at any of the online training material available at http://moodle.ukcip.org.uk from UKCIP?
   Yes / No
8. Have you downloaded any data from the User Interface site (apart from as a part of the training session? http://ukclimateprojections-ui.defra.gov.uk)
   Yes / No
   a. *If Yes:* Was this:
      Just to see how to download data / For a specific project you have worked or are working on

The UKCP09 data set includes projected changes in temperature, precipitation, air pressure, cloud and humidity for the UK over a range of time periods. It shows a range of possible outcomes and the probability of each outcome, based on how much evidence there is for different levels of future climate change.

9. Before today, were you aware that the UKCP09 data contains information on the probability of the changes of the future climate?
   Yes / No
10. On a scale of 1-5 (with 1 being high) how relevant is the uncertainty aspect of the climate projection data to the work you do? (If uncertainty is not relevant at all, or you have never used the data, choose 5)
    1 – Very relevant / 2 / 3 / 4 / 5 – Not relevant at all
11. On a scale of 1-5 (with 1 being high) how confident are you that you are using (or could use) the UKCP09 data correctly?
    1 – Very confident / 2 / 3 / 4 / 5 – Not confident at all
12. On a scale of 1-5 (with 1 being high) how would you rate your knowledge of probabilistic data?
    1 – I have a high level of knowledge / 2 / 3 / 4 / 5 – No knowledge at all
13. Which statement best describes your understanding of betting?

a. *Extensive knowledge about this area so I could calculate betting odds for a bookie on a race course*
b. *I know enough to place a bet confidently and understand the likelihood of winning*
c. *I know that 10/1 (10 to 1) is less likely that 5/1 (5 to 1) but not much more*
d. *I wouldn't be able to understand the odds to place a bet on a horse without help from someone else*

14. Learning Styles

Please answer each question by choosing whether the statement is nothing like you, partially like you or very much like you. You don't need to spend a long time thinking about these questions. Your immediate reaction is fine.

Choose from 'Nothing Like Me' / 'Partially Like Me' / 'Very Much Like Me'

a. Jingles, themes or parts of songs pop into your head at random
b. You can easily visualise objects, buildings, situations etc from plans or descriptions.
c. You navigate well and use maps with ease. You rarely get lost. You have a good sense of direction. You usually know which way North is.
d. You like listening to music - in the car, studying, at work (if possible!).
e. Music was your favourite subject at school
f. In school you preferred art, technical drawing, geometry.
g. You like using a camera or video camera to capture the world around you.
h. You use rhythm or rhyme to remember things, e.g. phone numbers, passwords, other little sayings.
i. You like books with lots of diagrams or illustrations.
j. You pay attention to the sounds of various things. You can tell the difference between instruments, or cars, or aircraft, based on their sound
k. You have a good sense of colour.
l. You can play a musical instrument or you can sing on (or close to) key
m. You occasionally realise you are tapping in time to music, or you naturally start to hum or whistle a tune. Even after only hearing a tune a few times, you can remember it.
n. You don't like the sound of silence. You would prefer to have some background music or other noises over silence.
o. You draw well, and find yourself drawing or doodling on a notepad when thinking.
p. You use diagrams and scribbles to communicate ideas and concepts. You love whiteboards (and colour pens).
q. You hear small things that others don't.

r. You are a tinkerer. You like pulling things apart, and they usually go back together OK. You can easily follow instructions represented in diagrams.

s. Music evokes strong emotions and images as you listen to it. Music is prominent in your recall of memories

t. You like visual arts, painting, sculpture. You like jigsaws and mazes.

The next question was asked after some of the evaluation maps (see Appendix C.1 for details)

15. For the previous map, did you generally rely more on the visual or the sound?
    a. Relied entirely on Visual
    b. Relied on both Vision and Sound, but used Visual more
    c. Relied on both Vision and Sound equally
    d. Relied on both Vision and Sound, but used Sound more
    e. Relied entirely on Sound

The final questions were asked at the end of the evaluation

16. Please rank the 3 different methods in order according to how easy you found them to use. Please only choose one method per rank:

    1 – Easiest / 2 / 3 – Hardest

    Map 3 – VV (Visual Visual) Both data sets represented visually

    Map 4 – VSVS (Visual Sonic Visual Sonic) Both data sets represented visually and sonically

    Map 5 & 6 – VS (Visual Sonic) 50th Percentile data represented visually and Range data represented sonically

17. Please rank the 3 different sounds in order according to how easy you found them to use. Please only choose one sound per rank:

    1 – Easiest / 2 / 3 – Hardest

    Low Sound / Medium Sound / High Sound

18. Have you been involved with my previous work?

    Yes at OS / Yes at UEA / Yes at LA / No

19. Do you suffer from any type of colour blindness? If you don't know, please choose No.

    Yes / No

20. (a) Which of the following are you currently studying for or already have? (Please choose the highest level qualification you are currently studying for or already have).

    High School or College Level (or equivalent)

    BSc/BA (or equivalent)

    MSc/MA

    PhD

(b) Which area is this qualification in?

Climate Change (using UKCP09 data)

Climate Change (not using UKCP09 data)

Environmental Science

Other Area / Non Area Specific

C.6. Videos

There are four videos available for this case study:

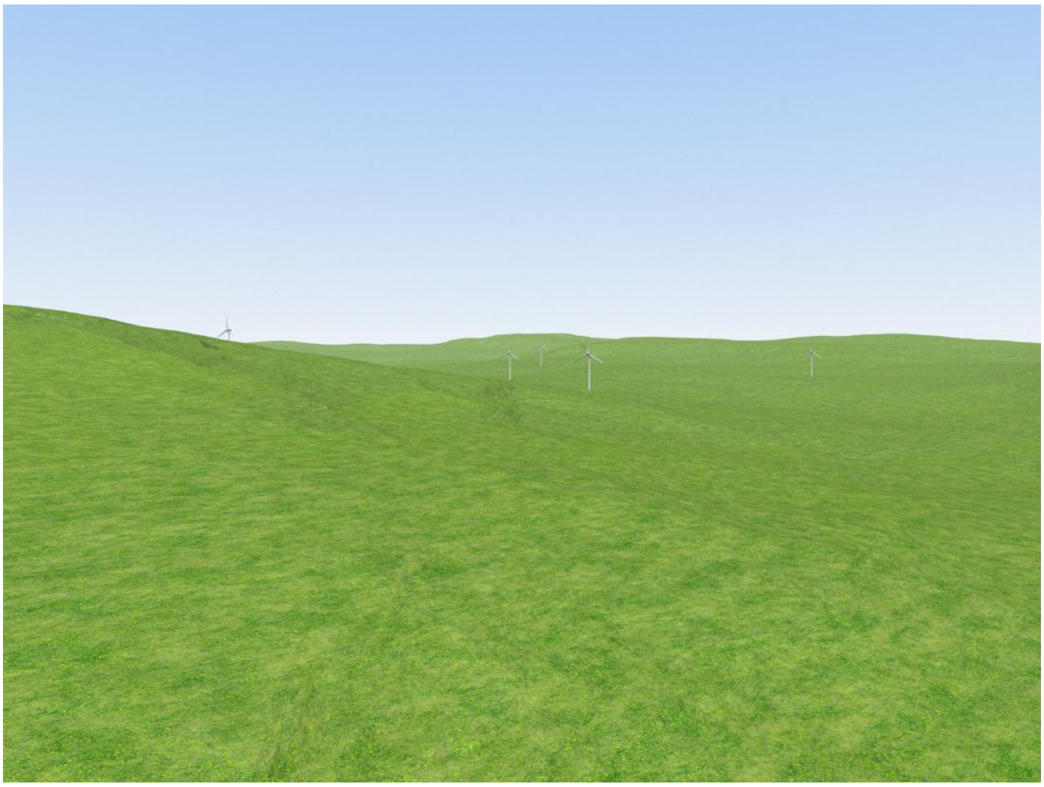
- `Complete Run Through Video.avi` – A video of a participant working through the whole evaluation. This is a very long video, and shows how the evaluation progresses from one stage to the next.
- `S (visual & sonic, same data, 1 map).mwv` – A video of the S map interface. This video shows how the S method works, where vision and sound are used to show one data set on one map.
- `VSVS (visual & sonic, 2 data sets, 2 maps).mwv` – A video of the VSVS map interface. This video shows how the VSVS method works, where data is shown using vision and sound, but with one data set on the left hand map and one data set on the right hand map. A similar clip is also available online at http://vimeo.com/17029341.
- `VS (visual & sonic, 2 data sets, 1 map).mwv` – A video of the VS map interface. This video shows how the VS method works, where two data sets are shown, but on one map. One data set is shown visually, and one data set is shown sonically. A similar clip is also available online at http://vimeo.com/17029358.

D Case Study 3

This evaluation was quite different to the first two case studies, as it did not require any programming at all. A full copy of all the Microsoft PowerPoint file used for the evaluation can be found at \Chapter 4 – CS3\Whole Evaluation.pptx on the DVD. Additionally, a shortened version shows the different sonification techniques (EvaluationDemo.pptx).

D.1. Examples of Visualisations

Below are some examples of the visualisations shown to the participants. All the visualisations are available in the file Whole Presentation.pptx.



D.2. Questions

The participants were asked a number of questions before they were shown the evaluations. These were:

- On a scale of 1-5 how do you rate your knowledge of GIS / spatial data in general?
  1 - I know a lot about it / 2 / 3 - I know a reasonable amount about it / 4/ 5 - I know nothing / very little about it
- On a scale of 1-5 how do you rate your knowledge of computer generated landscapes / landscape visualisations?
  1 - I know a lot about it / 2 / 3 - I know a reasonable amount about it / 4/ 5 - I know nothing / very little about it
- On a scale of 1-5 how do you rate your experience of planning applications involving wind turbines?
  1 - I know a lot about it / 2 / 3 - I know a reasonable amount about it / 4/ 5 - I know nothing / very little about it
- How much are the following statements like you? *Choose from 1 – Nothing Like Me / 2 – Partially Like Me / 3 – Very Much Like Me*
  - You can easily visualise objects, buildings, situations etc. from plans or descriptions.
  - You pay attention to the sounds of various things. You can tell the difference between instruments, cars or aircraft, based on their sound.
  - You can play a musical instrument or you can sing on (or close to) key.
  - You use diagrams and scribbles to communicate ideas and concepts. You love whiteboards (and colour pens).
  - You are a tinkerer. You like pulling things apart, and they usually go back together OK. You can easily follow instructions represented in diagrams.
  - Music evokes strong emotions and images as you listen to it. Music is prominent in your recall of memories.
- Have you been involved in any of my previous work?
  Yes / No

After each visualisation, participants were asked:

- How many turbines are within 2000m?
  1 / 2 / 3 / 4 / 5
- How confident are you that your answer is correct?
  1 – Very confident / 2 / 3 / 4 / 5 / 6 / 7 – Not confident

The questions at the end of the evaluation were:

- Which sound did you find most helpful?
  Piano Notes / Bird Calls / No Preference
- Which sound did you find more pleasant to listen to?
  Piano Notes / Bird Calls / No Preference
- When you could use both vision and sound, which did you use more?
  1 – Used sound exclusively / 2 – Used sound mainly, but also vision / 3 – Used sound and vision equally / 4 – Used vision mainly, but also sound / 5 – Used vision exclusively

D.3. Briefing Document

A two page briefing document was emailed to all respondents at least 24 hours before the session. Contact details for the gatekeeper were provided as appropriate for the Ordnance Survey / UEA / Planners. A copy is available at \Chapter 4 – CS3 on the DVD.

<u>Briefing Document</u>
<u>UEA, October 2011</u>

I am currently studying for a PhD in Environmental Sciences looking at how spatial data can be represented using sound. I'm particularly interested in how sound can be used to help represent distance in a landscape visualisation / computer generated landscape setting.

Landscape visualisations are frequently used in planning scenarios to show what a particular area may look like in the future with new developments. There are many different examples of this, and I am specifically using the examples of wind turbines in a rural landscape. There have been many studies looking at the perception of wind farms, both in the real world and in visualisations. In many of these situations user's perception of distance is often relatively poor, particularly over larger distances (Plumert et al. 2004).

One way of representing distances would be to use sound in combination with the visualisation. There have been various pieces of research on using sound within a general spatial data environment (for example Krygier 1994; Fisher 1994; MacVeigh & Jacobson 2007), but very little on the use of sound in a visualisation environment to represent additional data.

For my PhD, I am looking at how visual and sonic methods can be used to represent additional information using three different case studies. The first two case studies looked at using sound to represent uncertainty in a GIS setting, using Ordnance Survey Address Layer 2 (AL2) data and UK

Climate Projections 2009 (UKCP09) data as examples. This third case study will evaluate the use of sound to help represent distance in a landscape visualisation environment.

I have created an evaluative prototype which will take you through a series of exercises, showing a number of visualisations with different arrangements of wind turbines. You will be asked to estimate how many turbines are within a certain distance.

Your answers will be stored securely at UEA, will be completely anonymous and there will be no way to identify individual responses. You are free to withdraw from the study at any time before the evaluation is completed. The results will be used for my PhD and any related papers or publications. No personal or company names will be reported in the results. You will also have the opportunity to be sent a copy of the results of my research, if you wish.

If you have any questions about the study, please ask either myself or my supervisor, Andrew Lovett (contact details below).

I hope I have given you a bit of background information on my project and the evaluation that will take place. I very much look forward to seeing you at the session, and if you have any questions please feel free to call me on 07717745715 or email me on n.bearman@uea.ac.uk.

Contact Details:

Nick Bearman, PhD Student
E-mail: n.bearman@uea.ac.uk   Tel: 07717745715 or 01603 504943
School of Environmental Sciences, UEA, Norwich Research Park, Norwich, NR4 7TJ.

Andrew Lovett, Supervisor   E-mail: a.lovett@uea.ac.uk   Tel: 01603 593126
School of Environmental Sciences, UEA, Norwich Research Park, Norwich, NR4 7TJ.

<u>References:</u>

Fisher, P.F., 1994. Hearing the Reliability in Classified Remotely Sensed Images. Cartography and Geographic Information Systems, 21(1), 31-36.

Krygier, J.B., 1994. Sound and Geographic Visualization. In Visualization in Modern Cartography. Oxford, UK: Elsevier Science, pp. 149-166.

MacVeigh, R. & Jacobson, R.D., 2007. Increasing the dimensionality of a Geographic Information System (GIS) Using Auditory Display. Available at: http://www.music.mcgill.ca/icad2007/proceedings.php [Accessed August, 30 2011].

Plumert, J.M., Kearney, J.K. & Cremer, J.F., 2004. Distance perception in real and virtual environments. In Proceedings of the 1st Symposium on Applied perception in graphics and visualization - APGV '04. the 1st Symposium. Los Angeles, California, p. 27. Available at: http://dl.acm.org/citation.cfm?id=1012557 [Accessed August 30, 2011].

D.4. Participant Information Sheet and Consent Form

This document was given to the users when they arrived for the evaluation. They were asked to read, sign one copy and return to me. They were given a second copy, to keep if they wished. A copy is available at \Chapter 4 – CS3 on the DVD.

<u>Participant Information Sheet & Consent Form</u>

This information sheet contains information about the research I am conducting. Please read, sign one copy and return to me. A second copy is for you to keep if you wish. If you have any questions, please ask.

I am exploring different methods of using sound to help communicate distance within a landscape visualisation setting. I want to test whether the use of sound in this setting can improve people's ability to estimate distance.

The evaluation will compare different sonic and visual methods to see how effective they are at representing distance. To assess this, I will show you a series of landscapes and ask you to estimate how many turbines are within a certain distance. I will also ask a number of background questions to see which factors (if any) influence the usability of the sounds. After the evaluation, we will have a short discussion session covering your general feedback and thoughts on using sound to represent distance.

If you have any questions at any point, or would like more information then please ask. Your results will be recorded electronically and stored anonymously with no way to link them to you, and the results will only be used in aggregate analysis. The discussion session will be recorded to help me type up notes from the sessions. It will not be used for any other purpose and will be stored securely.

If you wish to end your participation in the research, you may stop at any point during the evaluation session and your results will not be used in the analysis. However, if you complete the evaluation and leave the room it will not be possible to remove your results at a later stage from the analysis as they are anonymous and cannot be identified.

This research will be used as a part of my PhD and any related publications or conference papers. Ordnance Survey will also receive a copy of my thesis as they partly fund my PhD. You have to option to receive a copy of the results from this work via email, and if you do your email address will be stored separately to the evaluation answers with no way to link them together.

If you have any questions or would like further explanation, please ask.

By signing below, I agree that I am happy to take part in the evaluation and have my data from this evaluation stored as described above.

Signed: ___________________________________ Date:__________________

Contact Details

Nick Bearman, PhD Student

E-mail: n.bearman@uea.ac.uk Tel: 07717745715 or 01603 504943

School of Environmental Sciences, UEA, Norwich Research Park, Norwich, NR4 7TJ.

Andrew Lovett, Supervisor

E-mail: a.lovett@uea.ac.uk Tel: 01603 593126

School of Environmental Sciences, UEA, Norwich Research Park, Norwich, NR4 7TJ.

References

1Spatial (2008). Radius Clarity, Accessed at http://www.1spatial.com/products/radius_clarity/ on 27/09/2012.

3D Nature (2011). What Is Visual Nature Studio? Accessed at http://3dnature.com/vnsinfo.html on 27/09/2012.

Ackerman, P.L., and Wolman, S.D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. Journal of Experimental Psychology: Applied *13*, 57–78.

Aerts, J.C.J.H., Clarke, K.C., and Keuper, A.D. (2003). Testing Popular Visualization Techniques for Representing Model Uncertainty. Cartography and Geographic Information Science *30*, 249–261.

Alexandrova, I.V., Teneva, P.T., Rosa, S. de la, Kloos, U., Bülthoff, H.H., and Mohler, B.J. (2010). Egocentric distance judgments in a large screen display immersive virtual environment. In Applied Perception in Graphics and Visualization, pp. 57–60.

Anderson, D.B., and Casey, M.A. (1997). The sound dimension. IEEE Spectrum *34*, 46–50.

Appleton, K., and Lovett, A. (2003). GIS-based visualisation of rural landscapes: defining `sufficient' realism for environmental decision-making. Landscape and Urban Planning *65*, 117–131.

Appleton, K., and Lovett, A. (2005). GIS-based visualisation of development proposals: reactions from planning and related professionals. Computers, Environment and Urban Systems *29*, 321–339.

Appleton, K., Lovett, A., Dockerty, T., and Sünnenberg, G. (2004). Representing Uncertainty in Visualisations of Future Landscapes. In Proceedings of the XXth ISPRS Congress, Istanbul, Turkey, Accessed at http://www.isprs.org/proceedings/XXXV/congress/comm4/papers/385.pdf on 27/09/2012.

Ashfield, S., Jarvis, C., and Kaduk, J. (2010). Serious Games for Geographical Field Skills: An Exploration. Second International Conference On Games and Virtual Worlds for Serious Applications (VS-GAMES), 2010, pp. 71 –74.

Aspinall, R.J., and Pearson, D. (1995). Describing and managing uncertainty of catagorical maps in GIS. In Innovations in GIS 2, (London, UK: Taylor & Francis),.

Axon, S., Speake, J., Crawford, K., Axon, S., Speake, J., and Crawford, K. (2012). "At the next junction, turn left": attitudes towards Sat Nav use, "At the next junction, turn left": attitudes towards Sat Nav use. Area *44*, 170-177.

Baddeley, A. (1992). Working Memory. Science *255*, 556–559.

Barbuceanu, F., Antonya, C., Duguleana, M., and Rusak, Z. (2011). Attentive user interface for interaction within virtual reality environments based on gaze analysis, Lecture Notes in Computer Science. In proceeding of: Human-Computer Interaction. Interaction Techniques and Environments - 14th International Conference, HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part II.

Barker, P.G., and Manji, K.A. (1989). Pictorial dialogue methods. International Journal of Man-Machine Studies *31*, 323–347.

Barrass, S. (2005). A comprehensive framework for auditory display: Comments on Barrass, ICAD 1994. ACM Trans. Appl. Percept. *2*, 403–406.

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

Batty, M. (2012). Smart cities, big data. Environment and Planning B: Planning and Design *39*, 191–193.

BBC News (2012). What would Higgs boson sound like? BBC. Accessed at http://www.bbc.co.uk/news/science-environment-18785232 on 14/07/2012.

Bearman, N. (2006). An Investigation into using Variations of Tempo and Volume in Sound to Represent Spatial Geographical Data. BSc Thesis, Department of Geography, University of Leicester.

Bearman, N. (2008). Using Sound to Represent Spatial Data in ArcGIS. MSc Thesis, Department of Geography, University of Leicester.

Bearman, N., and Appleton, K. (2011). Using Google Maps to collect spatial responses in a survey environment. GISRUK 2011, University of Portsmouth, UK.

Bearman, N., and Brown, E. (2012). Who's Sonifying Data and How Are They Doing it? A Comparision of ICAD and Other Venues since 2009. In Proceedings of the 18th International Conference on Auditory Display, (Atlanta, GA, USA, June 18-21 2012).

Bearman, N., and Fisher, P.F. (2012). Using sound to represent spatial data in ArcGIS. Computers & Geosciences *46*, 157–163.

Beekhuizen, J., Heuvelink, G.B.M., Biesemans, J., and Reusen, I. (2011). Effect of DEM Uncertainty on the Positional Accuracy of Airborne Imagery. IEEE Transactions on Geoscience and Remote Sensing *49*, 1567 –1577.

Benke, K.K., Pettit, C.J., and Lowell, K.E. (2011). Visualisation of spatial uncertainty in hydrological modelling. Journal of Spatial Science *56*, 73–88.

Bennett, C.A., and Rey, P. (1972). What's So Hot About Red? Human Factors: The Journal of the Human Factors and Ergonomics Society *14*, 149–154.

Bennett, D.J., and Edwards, A.D.N. (1998). Exploration of Non-seen Diagrams. International Conference on Auditory Display - ICAD 1998, Glasgow University, UK.

Benson, J.F. (2005). The Visualisation of Windfarms. In Visualization in Landscape and Environmental Planning : Technology and Applications, I.D. Bishop, and E. Lange, eds. (London, UK: Taylor & Francis).

Bertin, J. (1984). Semiology of Graphics: Diagrams, Networks, Maps, University of Wisconsin Press.

Beven, K. (2011). I believe in climate change but how precautionary do we need to be in planning for the future? Hydrological Processes *25*, 1517–1520.

Bishop, I.D. (2002). Determination of thresholds of visual impact: The case of wind turbines. Environment and Planning B: Planning and Design *29*, 707–718.

Bishop, I.D. (2003). Assessment of visual qualities, impacts, and behaviours, in the landscape, by using measures of visibility. Environment and Planning B: Planning and Design *30*, 677–688.

Bishop, I.D., and Miller, D.R. (2007). Visual assessment of off-shore wind turbines: The influence of distance, contrast, movement and social variables. Renewable Energy *32*, 814–831.

Bishop, I.D., and Stock, C. (2010). Using collaborative virtual environments to plan wind energy installations. Renewable Energy *35*, 2348–2355.

Bishop, I.D., Wherrett, J.R., and Miller, D.R. (2000). Using image depth variables as predictors of visual quality. Environment and Planning B: Planning and Design *27*, 865–875.

Bishop, M.J., and Cates, W.M. (2001). Theoretical foundations for sound's use in multimedia instruction to enhance learning. Educational Technology Research and Development *49*, 5–22.

Blades, M., Golledge, R.G., Jacobson, D., Kitchin, R.M., and Lippa, Y. (2002). The effect of spatial tasks on visually impaired peoples' wayfinding abilities. Journal of Visual Impairment & Blindness (JVIB) *96*, 407-419.

Blades, M., Ungar, S., and Spencer, C. (1999). Map Use by Adults with Visual Impairments. The Professional Geographer *51*, 539–553.

Blakemore, M. (1984). Part 4: Mathematical, Algorithmic and Data Structure Issues: Generalisation and Error in Spatial Data Bases. Cartographica: The International Journal for Geographic Information and Geovisualization *21*, 131–139.

Blattner, M.M., Sumikawa, D.A., and Greenberg, R.M. (1989). Earcons and Icons: Their Structure and Common Design Principles. Human-Computer Interaction *4*, 11-44.

Bonebright, T.L., and Flowers, J.H. (2011). Evaluation of Auditory Display. In The Sonification Handbook, T. Hermann, A. Hunt, and J.G. Neuhoff, eds. (Berlin, Germany: Logos Publishing House), 111–144.

Bonebright, T.L., and Nees, M.A. (2009). Most earcons do not interfere with spoken passage comprehension. Applied Cognitive Psychology *23*, 431–445.

Bovermann, T., Rohrhuber, J., and de Campo, A. (2011). Laboratory Methods for Experimental Sonification. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, pages 237-272. Logos Publishing House, Berlin, Germany.

Brazil, E. and Fernström, M. (2011). Auditory icons. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 13, pages 325–338. Logos Publishing House, Berlin, Germany.

Brent Council (2007). Interactive Maps (tMaps). Accessed at http://www.brent.gov.uk/maps on 18/12/2010.

Brewster, S.A. (1994). Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces. PhD Thesis, Human-Computer Interaction Group, Department of Computer Science, University of York.

Brewster, S.A. (1996). A sonically enhanced interface toolkit, International Conference on Auditory Display - ICAD 1996, Xerox PARC, USA.

Brewster, S.A. (1997). Using Non-Speech Sound to Overcome Information Overload. Displays *17*, 179–189.

Brewster, S.A. (1998). Sonically-Enhanced Drag and Drop. In International Conference on Auditory Display 1998, Glasgow.

Brewster, S.A. (2002). Overcoming the Lack of Screen Space on Mobile Computers. Personal Ubiquitous Computing *6*, 188–205.

Brewster, S.A., and Clarke, C.V. (1997). The design and evaluation of a sonically-enhanced tool palette, International Conference on Auditory Display - ICAD 1996, Xerox PARC, USA.

Brewster, S.A., McGookin, D., and Miller, C.A. (2006). Olfoto: Designing a Smell-Based Interaction. CHI 2006, Montréal, Québec, Canada.

Broad, K., Leiserowitz, A., Weinkle, J., and Steketee, M. (2007). Misinterpretations of the "Cone of Uncertainty" in Florida during the 2004 Hurricane Season. Bulletin of the American Meteorological Society *88*, 651-667.

Brown, A. (2006). Archiving websites : a practical guide for information management professionals, London: Facet Pub.

Bruce, C., and Walker, B.N. (2009). Developing Effective Real-time Audio Interpretation to Enhance Accessibility of Dynamic Zoo and Aquaria Exhibits. In Proceedings of the Association for the Advancement of Assistive Technology in Europe Conference (AAATE 2009), (Florence, Italy, 31 August - 02 September).

Bruckner, H.-P., Bartels, C., and Blume, H. (2011). PC Based real time sonification of human motion captured by internal sensors. In International Conference on Auditory Display, Budapest, Hungary.

Brunet, P., and Cowie, R. (2012). Towards a conceptual framework of research on social signal processing. Journal on Multimodal User Interfaces 1–15.

Brunsdon, C. (2011). Reproducible Research: What Can Geocomputation Achieve? GISRUK, University of Portsmouth, UK.

Bujacz, M., Skulimowski, P., and Strumiłło, P. (2011). Sonification of 3D scenes using personalized spatial audio to aid visually impaired persons. In Proceedings of the 17th International Conference on Auditory Display, Budapest, Hungary, 20-24 June 2011.

Bujacz, M., and Strumiłło, P. (2007). Synthesizing a 3D auditory scene for use in an electronic travel aid for the blind. SPIE *6937*.

Bunch, R.L., and Lloyd, R.E. (2006). The Cognitive Load of Geographic Information. The Professional Geographer *58*, 209–220.

Burrus, C. (2009). There's Math behind the Music! Accessed at http://www.charlieburrus.com/MathInMusic/Index.htm on 17/02/2010.

Business Green (2011). Half of wind farms in England and Wales rejected - 11 Jul 2011 - News from BusinessGreen. Accessed at http://www.businessgreen.com/bg/news/2092863/half-wind-farms-england-wales-rejected on 25/11/2011.

Buxton, W., Bly, S.A., Frysinger, S.P., Lunney, D., Mansur, D.L., Mezrich, J.J., and Morrison, R.C. (1985). Communicating with sound (panel session). SIGCHI Bull. *16*, 115–119.

de Campo, A., Daye, C., Frauenberger, C., Vogt, K., Wallisch, A., and Eckel, G. (2006). Sonification as an Interdisciplinary Working Process. In Proceedings of the 12th International Conference on Auditory Display, London, UK, June 20-23.

Candas, V., and Dufour, A. (2005). Thermal Comfort: Multisensory Interactions? Journal of Physiological Anthropology and Applied Human Science *24*, 33–36.

Caquard, S., Brauen, G., Wright, B., and Jasen, P. (2008). Designing sound in cybercartography: from structured cinematic narratives to unpredictable sound/image interactions. International Journal of Geographical Information Science *22*, 1219 – 1245.

Carles, J.L., Barrio, I.L., and de Lucio, J.V. (1999). Sound influence on landscape values. Landscape and Urban Planning *43*, 191–200.

Carlile, S. (2011). Psychoacoustics. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 3, pages 41–61. Logos Publishing House, Berlin, Germany.

Casey, M.B., Nuttall, R.L., and Pezaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: A comparison of spatial skills with internalized beliefs and anxieties. Developmental Psychology *33*, 669–680.

Cassettari, S., and Parsons, E. (1993). Sound as a Data Type in a Spatial Information System. In Proceedings of the 4th European Geographical Information Systems Conference (EGIS ’93), 194–202.

Cheshire, J., and Batty, M. (2012). Visualisation tools for understanding big data. Environment and Planning B: Planning and Design *39*, 413–415.

Clarke, K. (2010). Google Maps Mania. Accessed at http://googlemapsmania.blogspot.com/ on 31/07/2011.

Coffield, F., Moseley, D., Hall, E., and Eccleston, K. (2004). Learning styles and pedagogy in post-16 learning : A systematic and critical review (London, UK). Accessed at https://crm.lsnlearning.org.uk/user/order.aspx?code=041543 on 09/02/2010.

Collins, M., Booth, B., Harris, G., Murphy, J., Sexton, D., and Webb, M. (2006). Towards quantifying uncertainty in transient climate change. Climate Dynamics *27*, 127–147.

Coluccia, E., Iosue, G., and Brandimonte, M.A. (2007). The relationship between map drawing and spatial orientation abilities: A study of gender differences. Journal of Environmental Psychology *27*, 135–144.

Coluccia, E., and Louse, G. (2004). Gender differences in spatial orientation: A review. Journal of Environmental Psychology *24*, 329–340.

Comber, A., Fisher, P.F., and Wadsworth, R. (2003). Actor–network theory: a suitable framework to understand how land cover mapping projects develop? Land Use Policy *20*, 299–309.

Compound Security (2012). Mosquito Device. Accessed at http://www.compoundsecurity.co.uk/mosquito-mk4-multi-age on 28/03/2012.

Constantinescu, A., and Schultz, T. (2011). Redundancy versus complexity in auditory displays for object localization – A pilot study. In Proceedings of the 17th International Conference on Auditory Display, (Budapest, Hungary, 20-24 June 2011),.

Coombes, E.G., and Jones, A.P. (2010). Assessing the impact of climate change on visitor behaviour and habitat use at the coast: A UK case study. Global Environmental Change *20*, 303–313.

Couclelis, H., Golledge, R.G., Gale, N., and Tobler, W. (1987). Exploring the anchor-point hypothesis of spatial cognition. Journal of Environmental Psychology *7*, 99–122.

Cranston, G., and Lock, G.D. (2010). Who wants to be an Aerospace Engineer? Use of an Audience Response System to stimulate student learning in engineering lectures. Engineering Education: Journal of the Higher Education Academy Engineering Subject Centre *5*, 23–29.

Crawford, P.V. (1971). Perception of Grey-Tone Symbols. Annals of the Association of American Geographers *61*, 721–735.

Davison, B.K., and Walker, B.N. (2007). Sonification Sandbox Reconstruction: Software for Auditory Graphs. In Proceedings of the 13th International Conference on Auditory Display, Montréal, Canada, 26-29 June 2007.

Design Museum (2006). Jock Kinneir + Margaret Calvert / Designing Modern Britain - Design Museum : - Design/Designer Information. Accessed at http://designmuseum.org/design/jock-kinneir-margaret-calvert on 03/07/2012.

Devillers, R., Bedard, Y., and Jeansoulin, R. (2005). Multidimensional Management of Geospatial Data Quality Information for its Dynamic Use Within GIS. Photogrammetric Engineering and Remote Sensing *71*, 205–215.

Devillers, R., Bédard, Y., Jeansoulin, R., and Moulin, B. (2007). Towards spatial data quality information analysis tools for experts assessing the fitness for use of spatial data. International Journal of Geographical Information Science *21*, 261–282.

Dillman, D.A., Smyth, J.D., and Christian, L.M. (2009). Internet, Mail and Mixed-Mode Surveys: The Tailored Design Method (New Jersey, USA: John Wiley & Sons).

Diniz, N., Coussement, P., Deweppe, A., Demey, M., and Leman, M. (2012). An embodied music cognition approach to multilevel interactive sonification. Journal on Multimodal User Interfaces *5*, 211–219.

Dombois, F. (2002). Auditory Seismology: on free oscillations, focal mechanisms, explosions and synthetic seismograms. In Processing of the 2002 International Conference on Auditory Display, Kyoto, Japan, July 2-5.

Dombois, F. and Eckel, G. (2011). Audification. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 12, pages 301–324. Logos Publishing House, Berlin, Germany.

Dorling, D., and Openshaw, S. (1992). Using computer animation to visualize space - time patterns. Environment and Planning B: Planning and Design *19*, 639–650.

Dowdey, S. (2012). HowStuffWorks “Smell and Memory.” Accessed at http://science.howstuffworks.com/environmental/life/human-biology/smell.htm on 31/07/2012.

Drummond, J. (1995). Positional Accuracy. In Elements of Spatial Data Quality, (Oxford, UK: Elsevier Science).

Dubus, G., and Bresin, R. (2011). Sonification of physical quantities throughout history: a meta-study of previous mapping strategies. In Proceedings of the 17th International Conference on Auditory Display, (Budapest, Hungary),.

Duckham, M., and McCreadie, J.E. (2002). Error-aware GIS Development. In Spatial Data Quality, Eds Shi, W., Fisher, P.F. and Goodchild, M.F., London, UK: Taylor & Francis.

Ecotricity (2011). Ecotech Centre, Norfolk - Our Green Energy - Ecotricity. Accessed at http://www.ecotricity.co.uk/our-green-energy/our-green-electricity/from-the-wind/wind-parks-gallery/ecotech-centre-norfolk on 25/11/2011.

Edwards, A.D.N., Challis, B.P., Hankinson, J.C.K., and Pirie, F.L. (2000). Development of a standard test of musical ability for participants in auditory interface testing. In International Conference on Auditory Display, (Georgia Institue of Technology, Atlanta, Georgia, USA).

Edwards, A.D.N., Evreinov, G.E., and Agranovski, A.V. (1999). Isomorphic Sonification of Spatial Relations. In Proceedings of HCI International (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I - Volume I, (Hillsdale, NJ, USA: L. Erlbaum Associates Inc.), 526–530.

Edworthy, J., and Hellier, E. (2000). Auditory warnings in noisy environments. Noise Health *2*, 27–40.

Efstathiou, N., and Bailey, C. (2011). Promoting active learning using Audience Response System in large bioscience classes. Nurse Education Today.

Ehlschlaeger, C.R., and Shortridge, A.M. (1996). Modeling Elevation Uncertainty in Geographical Analyses. In Proceeding of the International Symposium on Spatial Data Handling, (Delft, Netherlands).

Ehlschlaeger, C.R., Shortridge, A.M., and Goodchild, M.F. (1997). Visualizing spatial data uncertainty using animation. Computers & Geosciences *23*, 387–395.

Eldridge, A. (2006). Issues in Auditory Display. Artifical Life *12*, 259–274.

Environment Agency (2009). Communication and dissemination of probabilistic flood warnings - literature review (Environment Agency). Accessed at http://publications.environment-agency.gov.uk/PDF/SCHO0909BQYG-E-E.pdf on 29/08/2012.

Erbacher, R.F., Grinstein, G., Lee, J.P., Levkowitz, H., Masterman, L., Pickett, R., and Smith, S. (1995). Exploratory visualization research at the University of Massachusetts at Lowell. Computers & Graphics *19*, 131–139.

ESRI (2011). ArcGIS Server. Accessed at http://www.esri.com/software/arcgis/arcgisserver/index.html on 30/11/2011.

Evans, G.W., Skorpanich, M.A., Gärling, T., Bryant, K.J., and Bresolin, B. (1984). The effects of pathway configuration, landmarks and stress on environmental cognition. Journal of Environmental Psychology *4*, 323–335.

Evans, J.R., and Mathur, A. (2005). The value of online surveys. Internet Research *15*, 195–219.

Evreinov, G.E. (2001). Spotty: Imaging Sonification Based on Spot-Mapping and Tonal Volume. In Proceedings of the 2001 International Conference on Auditory Display, (Espoo, Finland, 29 July - 1 August 2001).

Fairbairn, D., Andrienko, G., Andrienko, N., Buziek, G., and Dykes, J. (2001). Representation and Its Relationship with Cartographic Visualization: A Research Agenda. Cartography and Geographic Information Science *28*, 13–28.

Felder, R.M., and Silverman, L.K. (1988). Learning and Teaching Styles in Engineering Education. Engineering Education *78*, 674–681.

Fernstrom, M., and Taylor, T. (2010). Exploring ambient sonification of water toxicity. In International Conference on Auditory Display, (Washington, D.C., USA).

Fisher, P.F. (1994). Hearing the Reliability in Classified Remotely Sensed Images. Cartography and Geographic Information Systems *21*, 31–36.

Fisher, P.F. (1999). Models of uncertainty in spatial data. In Geographical Information Systems, pp. 191–205.

Fisher, P.F. (2000). Sorites paradox and vague geographies. Fuzzy Sets and Systems *113*, 7–18.

Fisher, P.F. (2009). The Representation of Uncertain Geographic Information. In Manual of Geographic Information Systems, M. Madden, ed. (Bethesda Md.: American Society for Photogrammetry and Remote Sensing).

Fisher, P.F., Dykes, J., and Wood, J. (1993). Map design and visualization. The Cartographic Journal *30*, 136–142.

Fisher, P.F., and Orf, T.M. (1991). An investigation of the meaning of near and close on a university campus. Computers, Environment and Urban Systems *15*, 23–35.

Fisher, P.F., and Pathirana, S. (1990). The evaluation of fuzzy membership of land cover classes in the suburban zone. Remote Sensing of Environment *34*, 121–132.

Flannery, J.J. (1971). The Relative Effectiveness of Some Common Graduated Point Symbols in the Presentation of Quantitative Data. Cartographica: The International Journal for Geographic Information and Geovisualization *8*, 96–109.

Flowers, J.H. (2005). Thirteen Years of Reflection on Auditory Graphing: Promises, Pitfalls, and Potential New Directions, Proceedings of the 11th International Conference on Auditory Display (ICAD2005), Limerick, Ireland.

Flowers, J.H., Buhman, D.C., and Turnage, K.D. (1996). Data Sonification from the Desktop: Should Sound be part of Standard Data Analysis Software? ICAD 1996, Palo Alto, California, USA.

Flowers, J.H., Buhman, D.C., and Turnage, K.D. (1997). Cross-Modal Equivalence of Visual and Auditory Scatterplots for Exploring Bivariate Data Samples. Human Factors: The Journal of the Human Factors and Ergonomics Society *39*, 341–351.

Flowers, J.H., and Hauer, T.A. (1995). Musical versus visual graphs: cross-modal equivalence in perception of time series data. Human Factors *37*, 553–569.

Flowers, J.H., Whitwer, L.E., Grafel, D.C., and Kotan, C.A. (2001). Sonification of daily weather records: Issues of perception, attention and memory in design choices. In Proceedings of the 7th International Conference on Auditory Display (ICAD2001), J. Hiipakka, N. Zacharov, and T. Takala, eds. (Espoo, Finland: Laboratory of Acoustics and Audio Signal Processing and the Telecommunications Software and Multimedia Laboratory, Helsinki University of Technology), pp. 222–226.

Foody, G.M., and Atkinson, P.M. (2002). Uncertainty in Remote Sensing and GIS (Chichester, UK: John Wiley & Sons).

Fortin, M., Voss, P., Lassonde, M., and Lepore, F. (2007). Sensory loss and brain reorganization. M S-Med. Sci. *23*, 917–922.

Frauenberger, C., and Barrass, S. (2008). Recycling Auditory Displays, International Community for Auditory Display 2008, Paris, France.

Frauenberger, C., and Stockman, T. (2009). Auditory display design-An investigation of a design pattern approach. International Journal of Human-Computer Studies *67*, 907–922.

Frauenberger, C., Stockman, T., and Bourguet, M.-L. (2007). A survey on common practice in designing audio in the user interface. In Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but Not as We Know It - Volume 1, (Swinton, UK, UK: British Computer Society), 187–194.

Gagnon, D. (1985). Videogames and spatial skills: An exploratory study. Educational Technology Research and Development *33*, 263–275.

Galea, L.A.M., and Kimura, D. (1993). Sex differences in route-learning. Personality and Individual Differences *14*, 53–65.

Gardiner, A., and Perkins, C. (2005). “It”s a sort of echo...’: Sensory perception of the environment as an aid to tactile map design. British Journal of Visual Impairment *23*, 84–91.

Gaver, W.W. (1993). What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. Ecological Psychology *5*, 1–29.

Genov, A., Keaveny, M., and Zazelenchuk, T. (2009). Usability Testing with Real Data. Journal of Usability Studies *4*, 85–92.

GeoDa Center (2012). OpenGeoDa (School of Geographical Sciences and Urban Planning, Arizona State University, USA). Accessed at http://geodacenter.asu.edu/ on 11/05/2012.

Ghosh, P. (2010). Sound of “God particle” simulated. BBC News. Accessed at http://www.bbc.co.uk/news/10385675 on 14/07/2012.

Gilmartin, P., and Lloyd, R. (1991). The effects of map projections and map distance on emotional involvement with places. International Journal of Radioactive Materials Transport *28*, 145–151.

Gilmartin, P.P. (1981). The Interface of Cognitive And Psychophysical Research In Cartography. Cartographica: The International Journal for Geographic Information and Geovisualization *18*, 9–20.

Gluck, M., Yu, L., Ju, B., Jeong, W., and Chang, C.-T. (1999). Augmented seriation: usability of a visual and auditory tool for geographic pattern discovery with risk perception data. In Geocomputation 99, (Mary Washington College Fredericksburg, Virginia, USA).

Godbout, A., and Boyd, J.E. (2010). Corrective sonic feedback for speed skating: A case study. In International Conference on Auditory Display, (Washington, D.C., USA, 9-15 June 2010), 23–30.

Goddard, J. (Danny) (2012). Climate Variations: Solar Insolation, Ecological Abundance and Stable Isotopes. In Proceedings of the 18th International Conference on Auditory Display, (Atlanta, GA, USA, June 18-21 2012),.

Goldsberry, K., and Battersby, S. (2009). Issues of Change Detection in Animated Choropleth Maps. Cartographica: The International Journal for Geographic Information and Geovisualization *44*, 201–215.

Golledge, R.G. (1993). Geography and the Disabled: A Survey with Special Reference to Vision Impaired and Blind Populations. Transactions of the Institute of British Geographers *18*, 63–85.

Golledge, R.G., Dougherty, V., and Bell, S. (1995). Acquiring Spatial Knowledge: Survey versus Route-Based Knowledge in Unfamiliar Environments. Annals of the Association of American Geographers *85*, 134–158.

Goodchild, M.F. (1995). Sharing Imperfect Data. In Sharing Geographic Information, pp. 413–425.

Goodchild, M.F., and Hunter, G.J. (1997). A simple positional accuracy measure for linear features. International Journal of Geographical Information Science *11*, 299–306.

Google (2011a). Google Maps API v3 - Versioning. Accessed at http://code.google.com/apis/maps/documentation/javascript/basics.html#Versioning on 16/07/2011.

Google (2011b). Google User Experience. Accessed at http://www.google.com/about/corporate/company/ux.html on 08/07/2011.

Google Groups thread (2008). Flash API Vs Javascript. Accessed at http://groups.google.com/group/google-maps-api-for-flash/browse_thread/thread/d7a5dc53d863f828 on 02/08/2011.

Gosling, S., McGregor, G., and Lowe, J. (2012). The benefits of quantifying climate model uncertainty in climate change impacts assessment: an example with heat-related mortality change estimates. Climatic Change *112*, 217–231.

GRaBS (2011). GRaBS - Adaptation for urban areas. Accessed at http://grabs-eu.org/assessment.php on 24/11/2011.

Graham, D., and Waterman, S. (2005). Underenumeration of the Jewish Population in the UK 2001 Census. Population, Space and Place *11*, 89–102.

Grechkin, T.Y., Nguyen, T.D., Plumert, J.M., Cremer, J.F., and Kearney, J.K. (2010). How does presentation method and measurement protocol affect distance estimation in real and virtual environments? ACM Trans. Appl. Percept. *7*, 26:1–26:18.

Grohn, M., and Lokki, T. (2003). Comparison of auditory, visual, and audio-visual navigation in a 3d space. (Boston, MA, USA).

Grond, F. and Berger, J. (2011). Parameter mapping sonification. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 15, pages 363–397. Logos Publishing House, Berlin, Germany.

Guo, Z. (2011). Mind the map! The impact of transit maps on path choice in public transit. Transportation Research Part A: Policy and Practice *45*, 625–639.

Haase, H., Göbel, M., Astheimer, P., Karlsson, K., Schröder, F., Frühauf, T., and Ziegler, R. (1994). How Scientific Visualization Can Benefit from Virtual Environments. CWI Quarterly *7*, 159–174.

Haklay, M. (2010). The source of the assertion that 80% of all organisational information is geographic. Accessed at http://povesham.wordpress.com/2010/02/22/the-source-of-the-assertion-that-80-of-all-organisational-information-is-geographic/#comments on 25/02/2010.

Haklay, M. (2012). Citizen Science as Participatory Science. In Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice, Eds Sui, D., Elwood, S. and Goodchild, M., Springer.

Haklay, M., Singleton, A., and Parker, C. (2008). Web Mapping 2.0: The Neogeography of the GeoWeb. Geography Compass *2*, 2011–2039.

Hall, J. (2007). Probabilistic climate scenarios may misrepresent uncertainty and lead to bad adaptation decisions. Hydrological Processes *21*, 1127–1129.

Harding, C., and Souleyrette, R.R. (2010). Investigating the use of 3D graphics, haptics (touch), and sound for highway location planning. Computer-Aided Civil and Infrastructure Engineering *25*, 20–38.

Harrower, M. (2007). The Cognitive Limits of Animated Maps. Cartographica: The International Journal for Geographic Information and Geovisualization *42*, 349–357.

Harrower, M., and Brewer, C. (2003). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. The Cartographic Journal *40*, 27–37.

Hart, R.A., and Moore, G.T. (1973). The Development of Spatial Cognition: A Review. In Image & Environment: Cognitive Mapping and Spatial Behavior, R. M. Downs D. Stea, ed. (New Brunswick, NJ, US: AldineTransaction), pp. 246–288.

Hawk, T.F., and Shah, A.J. (2007). Using Learning Style Instruments to Enhance Student Learning. Decision Sciences Journal of Innovative Education *5*, 1–19.

Healey, M., and Jenkins, A. (2000). Kolb's Experiential Learning Theory and Its Application in Geography in Higher Education. Journal of Geography *99*, 185.

Healey, M., Kneale, P., and Bradbeer, J. (2005). Learning styles among geography undergraduates: an international comparison. Area *37*, 30–42.

Hegarty, M., Richardson, A.E., Montello, D.R., Lovelace, K., and Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. Intelligence *30*, 425–447.

Hermann, T. (2008). Taxonomy and definitions for sonification and auditory display. In Proceedings of the 14th International Conference on Auditory Display (ICAD2008), (Paris, France),.

Hermann, T. (2011). Model-based sonification. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 16, pages 399–427. Logos Publishing House, Berlin, Germany.

Hermann, T., Hunt, A., Neuhoff, J. G. (2011). Introduction. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 1, pages 1–6. Logos Publishing House, Berlin, Germany.

Hermann, T., Ungerechts, B., Toussaint, H., and Grote, M. (2012). Sonification of Pressure Changes in Swimming for Analysis and Optimization. In Proceedings of the 18th International Conference on Auditory Display, (Atlanta, GA, USA, June 18-21 2012), pp. 60–67.

Hermann, T., and Zehe, S. (2011). Sonified Aerobics - Interactive Sonification of Coordinated Body Movements. In Proceedings of the 17th International Conference on Auditory Display, (Budapest, Hungary, 20-24 June 2011).

Hetherington, J., Daniel, T.C., and Brown, T.C. (1993). Is motion more important than it sounds?: The medium of presentation in environment perception research. Journal of Environmental Psychology 283–291.

Heuvelink, G.B.M. (1998). Error propagation in environmental modelling with GIS (London; Bristol, PA: Taylor & Francis).

Hickson, I. (2011). HTML5 -A vocabulary and associated APIs for HTML and XHTML. Accessed at http://dev.w3.org/html5/spec/Overview.html on 20/07/2011.

Hitwise (2010). Hitwise United Kingdom » - the power of competitive intelligence. Accessed at http://www.hitwise.com/uk/datacentre/main/dashboard-7323.html on 18/12/2010.

Holding, C.S., and Holding, D.H. (1989). Acquisition of Route Network Knowledge by Males and Females. Journal of General Psychology *116*, 29.

Hopkins, P.E. (2007). Thinking critically and creatively about focus groups. Area *39*, 528–535.

Horowitz, M. (2007). A mouse with ears explores maps. In Proceedings of the 13th International Conference on Auditory Display, (Montréal, Canada, 26-29 June 2007).

Hoskins, B., Lowe, J.A., Howard, T., Pardaens, A., Tinker, J., Holt, J., Wakelin, S., Milne, G., Leake, J., Wolf, J., et al. (2009). UKCP09 Science Review.

Huang, C.-F., Lu, H.-P., and Ren, J. (2012). Algorithmic approach to sonification of classical Chinese poetry. Multimedia Tools and Applications *61*, 489–518.

Hunt, A., Hermann, T., and Pauletto, S. (2004). Interacting with sonification systems: closing the loop. In Eighth International Conference on Information Visualisation, 2004. IV 2004. Proceedings, 879 – 884.

Hunter, G.J. (2002). Understanding Semantics and Ontologies: They're Quite Simple Really – If You Know What I Mean! Transactions in GIS *6*, 83–87.

Hunter, G.J., and Goodchild, M. (1995). Dealing with Error in Spatial Databases - A Simple Case-Study. Photogrammetric Engineering and Remote Sensing (PE&RS) *61*, 529-537.

Ibrahim, A.A.A., and Hunt, A. (2007). An HCI Model for Usability of Sonification Applications. In Task Models and Diagrams for Users Interface Design, K. Coninx, K. Luyten, and K.A. Schneider, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), 245–258.

IPCC (2000). IPCC Special Report Emissions Scenarios, Summary for Policy Makers, IPCC.

Jacobson, D. (1998). Cognitive Mapping Without Sight: four preliminary studies of spatial learning. Journal of Environmental Psychology *18*, 289–305.

Jacobson, D., and Kitchin, R.M. (1997). GIS and people with visual impairments or blindness: Exploring the potential for education, orientation, and navigation. Transactions in GIS *2*, 315–332.

Jallouli, J., and Moreau, G. (2009). An immersive path-based study of wind turbines' landscape: A French case in Plouguin. Renewable Energy *34*, 597–607.

Jeong, W. (2005). Multimodal trivariate thematic maps with auditory and haptic display. Proceedings of the American Society for Information Science and Technology *42*, NA.

Jeong, W., and Gluck, M. (2003). Multimodal geographic information systems: Adding haptic and auditory display. Journal of the American Society for Information Science and Technology *54*, 229–242.

Jeong, W., and Jacobson, D. (2002). Exploratory user study of haptic and auditory display for multimodal information systems. In Touch in Virtual Environments: Haptics and the Design of Interactive Systems. IMSC Series in Multimedia, (New York: Prentice Hall), 194–204.

Jeung, H., Chandler, P., and Sweller, J. (1997). The Role of Visual Indicators in Dual Sensory Mode Instruction. Educational Psychology *17*, 329–345.

Joslyn, S.L., Nadav-Greenberg, L., Taing, M.U., and Nichols, R.M. (2009). The effects of wording on the understanding and use of uncertainty information in a threshold forecasting decision. Applied Cognitive Psychology *23*, 55–72.

Josselin, D. (2011). Spatial Analysis Using Sonorous Cartography. Some Propositions. International Conference for Auditory Display, Paris, France, 3-8 July 2011.

Kalmus, H., and Fry, D.B. (1980). On tune deafness (dysmelodia): frequency, development, genetics and musical background. Annals of Human Genetics *43*, 369–382.

Kardos, J., Benwell, G., and Moore, A. (2005). The Visualisation of Uncertainty for Spatially Referenced Census Data Using Hierarchical Tessellations. Transactions in GIS *9*, 19–34.

Kasperson, R. (2008). Coping with Deep Uncertainty: Chanllenges for Environmental Assessment and Decision-making. In Uncertainty and Risk, (Earthscan), 337–348.

Keller, S.A., Koonin, S.E., and Shipp, S. (2012). Big data and city living – what can it do for us? Significance *9*, 4–7.

Kitchin, R.M. (1994). Cognitive maps: What are they and why study them? Journal of Environmental Psychology *14*, 1–19.

Kitchin, R.M. (1996). Increasing the Integrity of Cognitive Mapping Research: Appraising Conceptual Schemata of Environment-Behaviour Interaction. Progress in Human Geography *20*, 56–84.

Kitchin, R.M., Bartley, B., Gleeson, J., Cowman, M., Fotheringham, S., and Lloyd, C. (2007). Joined-up Thinking across the Irish Border: Making the Data more Compatible. Journal of Cross Border Studies *2*, 22–33.

Kitchin, R.M., Blades, M., and Golledge, R.G. (1997). Understanding Spatial Concepts at the Geographic Scale Without the Use of Vision. Progress in Human Geography *21*, 225–242.

Kitchin, R.M., Gleeson, J., and Dodge, M. (2012). Unfolding mapping practices: a new epistemology for cartography. Transactions of the Institute of British Geographers no–no.

Kitchin, R.M., Jacobson, D., Golledge, R.G., and Blades, M. (1998). Belfast Without Sight: Exploring Geographies of Blindness - IQDA. Irish Geography *31*, 34–46.

Klein, E., II, J.E.S., Schmidt, G.S., Livingston, M.A., and Staadt, O.G. (2009). Measurement Protocols for Medium-Field Distance Perception in Large-Screen Immersive Displays. In Virtual Reality, IEEE Annual International Symposium, pp. 107–113.

Kolb, D.A. (1981). Learning Styles and Disciplinary Differences. In The Modern American College, (San Francisco: Jossey-Bass), pp. p232–255.

Kramer, G. (1994). Auditory Display: Sonification, Audification, and Auditory Interfaces (Perseus Books).

Krygier, J.B. (1994). Sound and Geographic Visualization. In Visualization in Modern Cartography, (Oxford, UK: Elsevier Science), 149–166.

Kurzweil, R. (1992). The Age of Intelligent Machines, MIT Press.

Lachaux, J.-P., Axmacher, N., Mormann, F., Halgren, E., and Crone, N.E. (2012). High-frequency neural activity and human cognition: Past, present and possible future of intracranial EEG research. Progress in Neurobiology.

Lammers, W.J., and Badia, P. (2005). Fundamentals of behavioral research (Australia; Belmont, CA: Thomson/Wadsworth).

Lange, E. (2005). Issues and Questions for Research in Communicating with the Public through Visualizations. In Trends in Real-Time Landscape Visualization and Participation, (Anhalt University of Applied Sciences, Dessau, Germany: Herbert Wichmann Verlag),.

Lange, E. (2011). 99 volumes later: We can visualise. Now what? Landscape and Urban Planning *100*, 403–406.

Lee, J., Snyder, P.K., and Fisher, P.F. (1992). Modeling the Effect of Data Errors on Feature Extraction from Digital Elevation Models. Photogrammetric Engineering and Remote Sensing *58*, 1461–1467.

Leplatre, G., and Brewster, S.A. (2000). Designing non-speech sounds to support navigation in mobile phone menus.

Leung, Y., Goodchild, M.F., and Lin, C. (1992). Visualization of fuzzy scenes and probability fields. (Charleston), pp. 280–490.

Levitin, D.J. (2007). This is your brain on music : understanding a human obsession (London: Atlantic).

Levitin, D.J., and Rogers, S.E. (2005). Absolute pitch: perception, coding, and controversies. Trends in Cognitive Sciences *9*, 26–33.

Li, X.-B., and Reeves, G.R. (1999). A multiple criteria approach to data envelopment analysis. European Journal of Operational Research *115*, 507–517.

Liljedahl, M., Papworth, N., and Lindberg, S. (2007). Beowulf: A game experience built on sound effects. In Proceedings of the 13th International Conference on Auditory Display, (Montréal, Canada, June 26-29, 2007).

Liu, I., Levy, R.M., Barton, J.J.S., and Iaria, G. (2011). Age and gender differences in various topographical orientation strategies. Brain Research *1410*, 112–119.

Lloyd, R., and Steinke, T. (1985). Comparison of Qualitative Point Symbols: The Cognitive Process. Cartography and Geographic Information Science *12*, 156–168.

Lodha, S.K., Heppe, T., Beahan, J., Joseph, A., and Zane-Ulman, B. (1997). MUSE: A Musical Data Sonification Toolkit, Proceedings of the International Conference on Auditory Display (ICAD), Palo Alto, California.

Longley, P., Goodchild, M., Maguire, D.J., and Rhind, D.W. (2010). Geographic Information Systems and Science (John Wiley & Sons).

Longley, P.A., Goodchild, M.F., Maguire, D.J., and Rhind, D.. (2001). Geographic Information Systems and Science (Chicester, London: John Wiley & Sons).

Loomis, J.M., Golledge, R.G., and Klatzky, R.L. (1998). Navigation System for the Blind: Auditory Display Modes and Guidance. Presence: Teleoperators and Virtual Environments *7*, 193–203.

de Lucio, J.V., Mohamadian, M., Ruiz, J.P., Banayas, J., and Bernaldez, F.G. (1996). Visual landscape exploration as revealed by eye movement tracking. Landscape and Urban Planning *34*, 135–142.

Lunney, D., and Morrison, R.C. (1990). Auditory presentation of experimental data. Proceedings of SPIE *1259*, 140–146.

Lunnon, B. (2012). Postcode forecasting. Weather *67*, 192–192.

MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., and Hetzler, E. (2005). Visualizing geospatial information uncertainty: what we know and what we need to know. Cartography and Geographic Information Science *32*, 139(22).

MacLean, A.L., D'avello, T.P., and Shetron, S.G. (1993). The use of variability diagrams to improve the interpretation of digital soil maps in a GIS. Photogrammetric Engineering and Remote Sensing *59*, 223–228.

MacVeigh, R., and Jacobson, D. (2007). Increasing the dimensionality of a Geographic Information System (GIS) Using Auditory Display. In Proceedings of the 13th International Conference on Auditory Display, (Montréal, Canada), 530–535.

Madge, C., and O'Connor, H. (2002). On-line with e-mums: exploring the Internet as a medium for research. Area *34*, 92–102.

Mainwaring, E. (2011). Colour Vision Deficiency Inclusive Cartographic Design. (Macclesfield, UK, 8th - 10th June 2011).

Manyoky, M., Wissen Hayek, U., Klein, T.M., Pieren, R., Heutschi, K., and Gret-Regamey, A. (2012). Concept for collaborative design of wind farms facilitated by an interactive GIS-based visual-acoustic 3D simulation. In Peer Reviewed Proceedings of Digital Landscape Architecture 2012, (Anhalt University of Applied Sciences), 297–306.

MapTube (2010). MapTube. Accessed at http://www.maptube.org/ on 18/12/2010.

Mark, D.M. (1993). Human spatial cognition. In Human Factors in Geographical Information Systems, (Belhaven Press), 51–60.

Mark, D.M., Freksa, C., Hirtle, S.C., Lloyd, R., and Tversky, B. (1999). Cognitive models of geographical space. International Journal of Geographical Information Science *13*, 747–774.

Mark, D.M., and McGranagham, M. (1986). Effective provision of navigation assistance for drivers: A cognitive science approach. In Proceedings, Auto-Carto London, England, September 14-19, 399–408.

Mauney, L.M., and Walker, B.N. (2007). Individual Differences in Cognitive, Musical, and Perceptual Abilities. In Proceedings of the American Psychological Association Annual Conference, (San Francisco, August 17-20, 2007).

McCartney, J. (1996). SuperCollider: a New Real Time Synthesis Language. In Proceedings of the 1996 International Computer Music Conference, (Hong Kong).

McCrindle, R.J., and Symons, D. (2000). Audio space invaders. In Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies, pp. 59–65.

McGookin, D. and Brewster, S. (2011). Earcons. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 14, pages 339–361. Logos Publishing House, Berlin, Germany.

McGookin, D., Robertson, E., and Brewster, S.A. (2010). Cluthcing at Straws: Using Tangiable Interatcion to Provide Non-Visual Access to Graphs. In In Proceedings of ACM CHI 2010 (Atlanta, GA, USA), ACM Press Addison-Wesley, pp. 1715–1724.

Medvedev, A. (2011). Music Representation of Surface, Proceedings of the 25th International Cartographic Conference, Paris, France, 3-8 July 2011.

Meijer, P.B.L. (1992). An experimental system for auditory image representations. IEEE Transactions on Biomedical Engineering *39*, 112–121.

de Mendonça, A.L.A., and Delazari, L.S. (2012). Remote Evaluation of the Execution of Spatial Analysis Tasks with Interactive Web Maps: A Functional and Quantitative Approach. Cartographic Journal, The *49*, 7–20.

MIDI Manufacturers Association (2011). GM1 Sound Set. Accessed at http://www.midi.org/techspecs/gm1sound.php on 05/09/2011.

Milios, E., Kapralos, B., Kopinska, A., and Stergiopoulos, S. (2003). Sonification of range information for 3-D space perception. Neural Systems and Rehabilitation Engineering, IEEE Transactions On *11*, 416 –421.

Miller, C.C. (2006). A Beast in the Field: The Google Maps Mashup as GIS/2. Cartographica: The International Journal for Geographic Information and Geovisualization *41*, 187–199.

Milne, R. (2011). Planning Portal - Onshore wind farms refusals increase. Accessed at http://www.midi.org/techspecs/gm1sound.php on 25/11/2011.

Minnery, J., Knight, J., Byrne, J., and Spencer, J. (2009). Bounding Neighbourhoods: How Do Residents Do It? Planning Practice and Research *24*, 471–493.

Monmonier, M. (1992). Authoring Graphic Scripts: Experiences and Principles. Cartography and Geographic Information Science *19*, 247–260.

Monmonier, M. (1996). How to Lie with Maps (London, UK: The University of Chicago Press).

Montello, D.R. (1991). The measurement of cognitive distance: Methods and construct validity. Journal of Environmental Psychology *11*, 101–122.

Montello, D.R., and Freundschuh, S. (2005). Cognition of Geographic Information. In A Research Agenda For Geographic Information Science, (CRC Press).

Montello, D.R., Lovelace, K.L., Golledge, R.G., and Self, C.M. (1999). Sex-related differences and similarities in geographic and environmental spatial abilities. Annals of the Association of American Geographers 515–534.

Morgan, E., Gill, L., Lange, E., and Dallimer, M. (2012). Integrating Bird Survey Data into Real Time 3D Visual and Aural Simulations. In Peer Reviewed Proceedings of Digital Landscape Architecture 2012, (Anhalt University of Applied Sciences), 492–498.

Morse, D., Tardival, G.M., and Spicer, J. (1996). A Comparison of the Effectiveness of a Dichotomous Key and a Multi-Access Key to Woodlice. Accessed at http://kar.kent.ac.uk/21343/ pn 09/08/2012.

Mousavi, S., Low, R., and Sweller, J. (1995). Reducing Cognitive Load by Mixing Auditory and Visual Presentation Modes. Journal of Educational Psychology *87*, 319–334.

Munday, P., Jones, A.P., and Lovett, A. (2010). Utilising Scenarios to Facilitate Multi-Objective Land Use Modelling for Broadland, UK, to 2100. Transactions in GIS *14*, 241–263.

Murphy, J.M., Sexton, D.M.H., Jenkins, G.J., Booth, B.B.B., Brown, C.C., Clark, R.T., Collins, M., Harris, G.R., Kendon, E.J., Betts, R.A., et al. (2009). UK Climate Projections Science Report: Climate change projections (Exeter, UK).

Nasir, T. (2009). Geo-sonf: Spatial sonification of contour maps. In 2009 IEEE International Workshop on Haptic Audio Visual Environments and Games, (Lecco, Italy), 141–146.

Nelson, E.S. (1996). A cognitive map experiment: Mental representations and the encoding process. Cartography and Geographic Information Systems *23*, 229–248.

Nesbitt, K.V. (2006). Modelling human perception to leverage the reuse of concepts across the multi-sensory design space. In Proceedings of the 3rd Asia-Pacific Conference on Conceptual Modelling - Volume 53, (Hobart, Australia: Australian Computer Society, Inc.), 65–74.

Neuhoff, J. G. (2011). Perception, cognition and action in auditory display. In Hermann, T., Hunt, A., Neuhoff, J.G., editors, The Sonification Handbook, chapter 4, pages 63–85. Logos Publishing House, Berlin, Germany

Neuhoff, J.G., Kramer, G., and Wayand, J. (2002). Pitch and loudness interact in auditory displays: can the data get lost in the map? Journal of Experimental Psychology: Applied *8*, 17–25.

Newcombe, N., and Liben, L.S. (1982). Barrier effects in the cognitive maps of children and adults. Journal of Experimental Child Psychology *34*, 46–58.

Nielsen, J., and Levy, J. (1994). Measuring usability: preference vs. performance. Communications of the ACM *37*, 66–75.

Nørretranders, T. (1999). The User Illusion: Cutting Consciousness Down to Size (Penguin Books Ltd).

O'Berine, J. (2010). 41Latitude - A Brief Comparison of Google Maps, Bing Maps, & Yahoo! Maps. Accessed at http://www.41latitude.com/post/557224600/map-comparison on 18/12/2010.

Oliveira, M.A. (2008). Electric Piano 2.5. Accessed at http://www.pianoeletronico.com.br/index-en.html on 27/09/2012.

OpenLayers (2011). OpenLayers. Accessed at www.openlayers.org on 30/11/2011.

Openshaw, S. (1984). The Modifiable Areal Unit Problem, CATMOG:38 (Norwich: Geo Books).

Openshaw, S., Waugh, D., and Cross, A. (1994). Some Ideas About the Use of Map Animation as a Spatial Analysis Tool. In Visualisation in Geographic Information Science, (Chichester, UK: Wiley),.

Ordnance Survey (2008). Data Specific for MSc Thesis, including MasterMap Topography & ITN Layer.

Ordnance Survey (2009). OS MasterMap Address Layer 2 Technical Specification. Accessed at http://www.ordnancesurvey.co.uk/oswebsite/products/osmastermap/layers/addresslayer2/detailedproductinfo/OSMMALandAL2userguide.pdf on 28/01/2010.

Ordnance Survey (2012a). AddressBase suite of products - the official site. Accessed at http://www.ordnancesurvey.co.uk/oswebsite/products/addressbase-products.html on 11/08/2012.

Ordnance Survey (2012b). AddressBase technical specification. Accessed at http://www.ordnancesurvey.co.uk/oswebsite/docs/technical-specifications/addressbase-technical-specification-csv.pdf on 11/08/2012.

Ordnance Survey, http://www. ordnancesurvey.co.uk/oswebsite/support/contact-us. htm. (2012c). OS OpenData - high quality map data to download or order today.

Oviatt, S., Coulston, R., and Lunsford, R. (2004). When do we interact multimodally?: cognitive load and multimodal communication patterns, Proceedings of the 6th international conference on Multimodal interfaces.

Paar, P. (2006). Landscape visualizations: Applications and requirements of 3D visualization software for environmental planning. Computers, Environment and Urban Systems *30*, 815–839.

Pagulayan, R.J., Keeker, K., Wixon, D., Romero, R.L., and Fuller, T. (2003). The human-computer interaction handbook. In Handbook for Human-Computer Interaction in Interactive Systems, J.A. Jacko, and A. Sears, eds. (Hillsdale, NJ, USA: L. Erlbaum Associates Inc.), pp. 883–906.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology/Revue Canadienne De Psychologie *45*, 255–287.

Paolino, L., Romano, M., Sebillo, M., Tortora, G., and Vitiello, G. (2009). Audio-Visual Information Clues about Geographic Data on Mobile Interfaces. In Advances in Multimedia Information Processing - PCM 2009, 1156–1161.

Parbery-Clark, A., Anderson, S., Hittner, E., and Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. Neurobiology of Aging.

Parsons, E. (1995). GIS visualisation tools for qualitative spatial information. In Innovations in GIS 2, P.F. Fisher, ed. 169–187.

Pashler, H., McDaniel, M., Rohrer, D., and Bjork, R. (2008). Learning Styles Concepts and Evidence. Psychological Science in the Public Interest *9*, 105–119.

Pasqualotto, A., and Newell, F.N. (2007). The role of visual experience on the representation and updating of novel haptic scenes. Brain and Cognition *65*, 184–194.

Pattison, P., and Grieve, N. (1984). Do spatial skills contribute to sex differences in different types of mathematical problems? Journal of Educational Psychology *76*, 678–689.

Pauletto, S., and Hunt, A. (2009). Interactive sonification of complex data. International Journal of Human-Computer Studies *67*, 923–933.

Peng, Z.-R. (2001). Internet GIS for public participation. Environment and Planning B: Planning and Design *28*, 889–905.

Peres, S.C., and Lane, D.M. (2005). Auditory Graphs: The Effects of Redundant Dimensions and Divided Attention. In Proceedings of the 11th International Conference on Auditory Display, Limerick, Ireland.

Peretz, I., and Hyde, K.L. (2003). What is specific to music processing? Insights from congenital amusia. Trends in Cognitive Sciences *7*, 362–367.

Perkins, C. (2002). Cartography: Progress in Tactile Mapping. Progress in Human Geography *26*, 521–530.

Pidgeon, N. (2008). Risk, Uncertainty and Social Controversy: From Risk Perception and Communication to Public Engagement. In Uncertainty and Risk, (Earthscan), 349–361.

Plumert, J.M., Kearney, J.K., and Cremer, J.F. (2004). Distance perception in real and virtual environments. In Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization - APGV ’04, (Los Angeles, California), 27.

Pollack, I., and Ficks, L. (1954). Information of Elementary Multidimensional Auditory Displays. The Journal of the Acoustical Society of America *26*, 155–158.

Porter, J.H., Hanson, P.C., and Lin, C.-C. (2012). Staying afloat in the sensor data deluge. Trends in Ecology & Evolution *27*, 121–129.

Pow, C.P. (2002). “Sense and Sensibility”: Social-spatial Experiences of the Visually-impaired in Singapore. Singapore Journal of Tropical Geography *21*, 166–182.

Rabinovich, A., and Morton, T.A. (2012). Unquestioned Answers or Unanswered Questions: Beliefs About Science Guide Responses to Uncertainty in Climate Change Risk Communication. Risk Analysis *32*, 992–1002.

Radburn, R., Dykes, J., and Wood, J. (2010). vizLib: Using The Seven Stages of Visualization to Explore Population Trends and Processes in Local Authority Research. In GISRUK2010,.

Ramirez, G., Gunderson, E.A., Levine, S.C., and Beilock, S.L. (2012). Spatial anxiety relates to spatial abilities as a function of working memory in children. Quarterly Journal of Experimental Psychology *65*, 474–487.

Rébillat, M., Boutillon, X., Corteel, É., and Katz, B.F.G. (2012). Audio, visual, and audio-visual egocentric distance perception by moving subjects in virtual environments. ACM Transactions on Applied Perception *9*, 19:1–19:17.

Reivo (2011). TurningPoint. Accessed at http://www.turningtechnologies.co.uk/ on 05/11/2011.

Renewable UK (2010). RenewableUK - Onshore Wind. Accessed at http://www.bwea.com/onshore/index.html on 25/11/2011.

Renewable UK (2011). RenewableUK Press Release: Wind deployment increases, but planning system needs reform. Accessed at http://www.bwea.com/media/news/articles/pr20111025-2.html on 25/11/2011.

Ribeiro Amigo, G. (2010). Javascript Sound Kit. Accessed at http://jssoundkit.sourceforge.net/ on 18/12/2010.

Rice, M., Jacobson, D., Golledge, R.G., and Jones, D. (2005). Design Considerations for Haptic and Auditory Map Interfaces. Cartography and Geographic Information Science *32*, 381–391.

Riding, R., and Cheema, I. (1991). Cognitive Styles—an overview and integration. Educational Psychology *11*, 193–215.

Riecke, B.E., and Wiener, J.M. (2007). Can People Not Tell Left from Right in VR? Point-to-origin Studies Revealed Qualitative Errors in Visual Path Integration. (IEEE), 3–10.

Rigas, D., and Alty, J. (2005). The rising pitch metaphor: an empirical study. International Journal of Human-Computer Studies *62*, 1–20.

Risko, E.F., Anderson, N.C., Lanthier, S., and Kingstone, A. (2011). Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. Cognition.

Rivara, F.P., Koepsell, T.D., Wang, J., Durbin, D., Jaffe, K.M., Vavilala, M., Dorsch, A., Roper-Caldbeck, M., Houseknecht, E., and Temkin, N. (2011). Comparison of Telephone with World Wide Web-Based Responses by Parents and Teens to a Follow-Up Survey after Injury. Health Services Research *46*, 964–981.

Robinson, A.H. (1952). The Look of Maps (Madison, Wisconsin, USA: The University of Wisconsin Press).

Robinson, A.H. (1995). Elements of cartography (New York: Wiley).

Rousell, A., Jarvis, C., and Li, J. (2008). The Effectiveness of Artificial Landmarks for Orientation Tasks Within a Large-Scale Virtual Environment. In Proceedings of GISRUK 2008,.

Sanders, M.S. (1992). Human Factors in Engineering and Design ~Seventh Edition (McGraw-Hill Higher Education).

Sang, N., Ode, Å., and Miller, D. (2008). Landscape metrics and visual topology in the analysis of landscape preference. Environment and Planning B: Planning and Design *35*, 504–520.

Saue, S. (2000). A Model for Interaction in Exploratory Sonification Displays. In Proceedings of International Conference for Auditory Display.

Schafer, R.M. (1994). The Soundscape (Destiny Books, USA).

Schaffert, N., Mattes, K., and Effenberg, A.O. (2011). The sound of rowing stroke cycles as acoustic feedback. (Budapest, Hungary),.

Schröder, F. (1995). apE - the original dataflow visualization environment. ACM SIGGRAPH Computer Graphics - Special focus: modular visualization environments (MVEs) *29*, 5–9.

Schroth, O. (2010). From Information to Participation: Interactive Landscape Visualization as a Tool for Collaborative Planning (vdf Hochschulverlag AG, Switzerland).

Schweizer, D.M., and Goodchild, M.F. (1992). Data quality and choropleth maps: An experiment with the use of color. In Proceedings, GIS/LIS ’92, (San Jose, California, USA), 686–699.

Seagull, F.J., Wickens, C.D., and Loeb, R.G. (2001). When Is Less More? Attention and Workload in Auditory, Visual, and Redundant Patient-Monitoring Conditions. Proceedings of the Human Factors and Ergonomics Society Annual Meeting *45*, 1395–1399.

Self, C.M., Gopal, S., Golledge, R.G., and Fenstermaker, S. (1992). Gender-related differences in spatial abilities. Progress in Human Geography *16*, 315–342.

Sevenant, M., and Antrop, M. (2011). Landscape Representation Validity: A Comparison between On-site Observations and Photographs with Different Angles of View. Landscape Research *36*, 363–385.

Sexton, D.M.H., and Murphy, J.M. (2011). Multivariate probabilistic projections using imperfect climate models. Part II: robustness of methodological choices and consequences for climate sensitivity. Climate Dynamics *38* 2543-2558.

Sexton, D.M.H., Murphy, J.M., Collins, M., and Webb, M.J. (2011). Multivariate probabilistic projections using imperfect climate models part I: outline of methodology. Climate Dynamics *38* 2513-2542.

Shang, H., and Bishop, I.D. (2000). Visual Thresholds for Detection, Recognition and Visual Impact in Landscape Settings. Journal of Environmental Psychology *20*, 125–140.

Shepherd, I.D.H. (1995). Putting time on the map: time-varying symbolism, data visualization and GIS. In Innovations in GIS 2, P.F. Fisher, ed. 169–187.

Shepherd, I.D.H., and Bleasdale-Shepherd, I.D. (2008). Multi-sensory Data Representation in Virtual Worlds: abstraction or pragmatism? International Conference on Developments in Visualization and Virtual Environments in Geographic Information Science.

Shepherd, I.D.H., and Bleasdale-Shepherd, I.D. (2009). Videogames: the new GIS? Accessed at http://eprints.mdx.ac.uk/6589/ on 09/07/2012.

Sheppard, S.R.J. (2001). Guidance for crystal ball gazers: developing a code of ethics for landscape visualization. Landscape and Urban Planning *54*, 183–199.

Sheppard, S.R.J., and Cizek, P. (2009). The ethics of Google Earth: Crossing thresholds from spatial data to landscape visualisation. Journal of Environmental Management *90*, 2102–2117.

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Visual Languages, IEEE Symposium, 336.

Shortridge, B.G., and Welch, R.B. (1980). Are We Asking the Right Questions? Comments on Instructions in Cartographic Psychophysical Studies. Cartography and Geographic Information Science *7*, 19–23.

Simkhada, P., Van Teijlingen, E., Kadel, S., Stephens, J., Sharma, S., and Sharma, M. (2009). Reliability of national data sets: Evidence from a detailed small area study in rural Kathmandu Valley, Nepal. Asian Journal of Epidemiology *2*, 44–48.

Sinclair, P. (2012). Sonification: what where how why artistic practice relating sonification to environments. AI & Society *27*, 173–175.

Smith, J., Beran, M., Couchman, J., and Coutinho, M. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. Psychonomic Bulletin & Review *15*, 679–691.

Smith, S., Bergeron, R.D., and Grinstein, G.G. (1990). Stereophonic and surface sound generation for exploratory data analysis. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People, (New York, NY, USA: ACM), 125–132.

SoundMAX (2011). SoundMAX Home Page. Accessed at http://www.soundmax.com.tw/index.html on 05/09/2011.

Spear, M., Hall, J., and Wadsworth, R. (1996). Communication of Uncertainty in Spatial Data to Policy Makers, Spatial Accuracy Assessment in Natural Resources and Environmental Sciences: Second International Symposium.

Speeth, S.D. (1961). Seismometer Sounds. The Journal of the Acoustical Society of America *33*, 909.

Stevens, A., and Coupe, P. (1978). Distortions in judged spatial relations. Cognitive Psychology *10*, 422–437.

Stewart, R.R., and Brough, S. (2008). Log jammin': transforming well logs to music.

Steynor, A., Gawith, M., and Street, R. (2012). Engaging users in the development and delivery of climate projections: the UKCIP experience of UKCP09. (UKCIP, Oxford, UK).

Storms, R.L., and Zyda, M.J. (2000). Interactions in Perceived Quality of Auditory-Visual Displays. Presence: Teleoperators and Virtual Environments *9*, 557–580.

Supper, A. (2012). "Trained Ears" and "Correlation Coefficients": A Social Science Perspective on Sonification. In Proceedings of the 18th International Conference on Auditory Display, (Atlanta, GA, USA, June 18-21 2012), 29–35.

Susini, P., Misdariis, N., Lemaitre, G., and Houix, O. (2012). Naturalness influences the perceived usability and pleasantness of an interface's sonic feedback. Journal on Multimodal User Interfaces *5*, 175–186.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. Learning and Instruction *4*, 295–312.

Thorndyke, P.W., and Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation. Cognitive Psychology *14*, 560–589.

Tindall-Ford, S., Chandler, P., and Sweller, J. (1997). When Two Sensory Modes Are Better Than One. Journal of Experimental Psychology: Applied *3*, 257–287.

Tiuraniemi, M.-K. (2011). Guiding Divers with Sound. In Proceedings of the 17th International Conference on Auditory Display, (Budapest, Hungary, 20-24 June 2011),.

Tobler, W.R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography *46*, 234–240.

Tufte, E.R. (1983). The Visual Display of Quantitative Information (Cheshire, Connecticut, USA: Graphics Press).

Tufte, E.R. (1990). Envisioning Information (Connecticut, USA: Graphics Press).

Tufte, E.R. (1997). Visual explanations : images and quantities, evidence and narrative (Cheshire Conn.: Graphics Press).

Tufte, E.R. (2001). The Visual Display of Quantitative Information (Cheshire, Connecticut, USA: Graphics Press).

Tzelgov, J., Srebro, R., Henik, A., and Kushelevsky, A. (1987). Radiation Search and Detection by Ear and by Eye. Human Factors: The Journal of the Human Factors and Ergonomics Society *29*, 87–95.

UKCIP (2011). Response to UKCP09 Science Review.

Ulla, B. (2005). Using scantron versus an audience response system for survey research: Does methodology matter when measuring computer-mediated communication competence? Computers in Human Behavior *21*, 343–359.

University of Newcastle (2002). Visual Assessment of Windfarms Best Practice (Scottish Natural Heritage Commissioned Report F01AA303A).

Unwin, D.J. (1994). Cartography, ViSC and GIS. Progress in Human Geography *18*, 516–522.

Unwin, D.J. (1995). Geographical information systems and the problem of “error and uncertainty.”Progress in Human Geography *19*, 549–558.

Veregin, H., Krause, P., Pandya, R., and Roethlisberger, R. (1993). Design and Development of an Interactive “Geiger Counter” for Exploratory Analysis of Spatial Data Quality. GIS/LIS 93 701–710. Annual Conference November 2-4, 1993. Minneapolis, Minnesota, USA.

Vogt, K., Goudarzi, V., and Holdrich,, R. (2012). SysSon - A Systematic Procedure to Develop Sonifications. In Proceedings of the 18th International Conference on Auditory Display, (Atlanta, GA, USA, June 18-21 2012), pp. 229–230.

Vogt, K., Holdrich, R., Pirro, D., and Gattringer, C. (2010). Spin quartets - Sonification of the XY model. In Proceedings of the 16th International Conference on Auditory Display, (Washington, D.C., USA),.

Volk, A., and Honingh, A. (2012). Mathematical and computational approaches to music: challenges in an interdisciplinary enterprise. Journal of Mathematics and Music *6*, 73–81.

Walker, B.N. (2000). Magnitude Estimation of Conceptual Data Dimensions for Use in Sonification. PhD Thesis. Rice University, U.S.A.

Walker, B.N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. J Exp Psychol Appl *8*, 211–221.

Walker, B.N. (2007). Consistency of magnitude estimations with conceptual data dimensions used for sonification. Applied Cognitive Psychology *21*, 579–599.

Walker, B.N., and Kramer, G. (1996). Mappings and metaphors in auditory displays. In Proceedings of the Third International Conference on Auditory Display ICAD96, (Palo Alto, California, USA, 4-6 November 1996),.

Walker, B.N., and Lane, D.M. (2001). Psychophysical scaling of sonification mappings: A comparison of visually impaired and sighted listeners. In Proceedings of the 2001 International Conference on Auditory Display, (Espoo, Finland, 29 July - 1 August 2001),.

Walker, B.N., and Lowey, M. (2003). Sonification sandbox: a graphical toolkit for auditory graphs. (Boston, MA, USA),.

Walker, B.N., and Nees, M.A. (2011). Theory of Sonification. In The Sonification Handbook, In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook. Logos Publishing House, Berlin, Germany.

Wang, X., Song, B., Chen, J., Zheng, D., and Crow, T.R. (2006). Visualizing forest landscapes using public data sources. Landscape and Urban Planning *75*, 111–124.

Ware, C. (2004). Information visualization : perception for design (San Francisco, CA: Morgan Kaufman).

Watson, M. (2006). Scalable Earcons: Bridging the Gap Between Intermittent and Continuous Auditory Displays. In Proceedings of the 12th International Conference on Auditory Display, (London, UK, June 20-23),.

Weber, P., Ellul, C., and Jones, C.E. (2012). Opening up GIS Training and Education – Building an IIGLU, GISRUK2012, 11th - 13th April 2012, Lancaster University, Lancaster, UK

Wechsung, I., Engelbrecht, K.-P., Kühnel, C., Möller, S., and Weiss, B. (2012). Measuring the Quality of Service and Quality of Experience of multimodal human–machine interaction. Journal on Multimodal User Interfaces 1–13.

Westminster City Council (2010). Find it in Westminister. Accessed at http://www.westminster.gov.uk/ on 18/12/2010.

Wherrett, J.R. (1999). Issues in using the Internet as a medium for landscape preference research. Landscape and Urban Planning *45*, 209–217.

Wieczorek, J., Guo, Q., and Hijmans, R. (2004). The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. International Journal of Geographical Information Science *18*, 745 – 767.

Winberg, F., and Hellstroem, S.O. (2001). Qualitative aspects of auditory direct manipulation. A case study of the towers of Hanoi. In Proceedings of the International Conference on Auditory Display, (Espoo, Finland, 29 July - 1 August 2001).

Wood, D., and Fels, J. (2009). The Natures of Maps: Cartographic Constructions of the Natural World (University Of Chicago Press).

Wood, J., Hearnshaw, H., and Unwin, D.J. (1994). Visualizing Contour Interpolation Accuracy in Digital Elevation Models. In Visualisation in Geographic Information Science, (Chichester, UK: Wiley),.

Worrall, D. (2009). Sonification and Information Concepts, Instruments and Techniques. University of Canberra. Communication.

Wright, J.K. (1944). The Terminology of Certain Map Symbols. Geographical Review *34*, 653–654.

Wright, K.B. (2005). Researching Internet-Based Populations: Advantages and Disadvantages of Online Survey Research, Online Questionnaire Authoring Software Packages, and Web Survey Services. Journal of Computer-Mediated Communication *10*.

Yeshurun, Y., Lapid, H., Dudai, Y., and Sobel, N. (2009). The Privileged Brain Representation of First Olfactory Associations. Current Biology *19*, 1869–1874.

Yeung, E.S. (1980). Pattern recognition by audio representation of multivariate analytical data. Anal. Chem. *52*, 1120–1123.

Yinnan, Z., Gang, C., and Zhigang, F. (2008). Study on Application of CAD Sonification. International Conference on Computer Science and Software Engineering.

Zeiler, M. (2000). Modeling Our World: The ESRI Guide to Geodatabase Design (ESRI, Inc.).

Zerger, A. (2002). Examining GIS decision utility for natural hazard risk modelling. Environmental Modelling & Software *17*, 287–294.

Zhang, J., and Goodchild, M.. F. (2002). Uncertainty in Geographical Information (London, UK: Taylor & Francis).

Zhang, R., Nordman, A., Walker, J., and Kuhl, S.A. (2012). Minification affects verbal- and action-based distance judgments differently in head-mounted displays. ACM Trans. Appl. Percept. *9*, 14:1–14:13.

Zhao, H., Plaisant, C., Shneiderman, B., and Lazar, J. (2008). Data Sonification for Users with Visual Impairment: A Case Study with Georeferenced Data. ACM Transactions on Computer-Human Interaction *15*, 1–28.