COMPARISONS ON THE PERCEPTIONS OF REPRODUCED URBAN SOUNDFIELDS AND URBAN SOUNDSCAPES: A MIXED MODEL APPROACH

Tobias Alexander ACKROYD

Acoustics Research Centre

School of Computing, Science and Engineering

College of Science and Technology

University of Salford, Salford, UK

Submitted in Partial Fulfilment of the Requirements of the

Degree of Doctor of Philosophy, November 2014

For Sarah

ACKNOWLEDGEMENTS

My warmest thanks go to Sarah, my parents – Mum, Andy and Dad, Granny and Charlie for their help and support in all forms of kindness. I also thank in particular my supervisors, Yiu Lam and Bill Davies, as well as Andy Elliot, Calum Sharp, Nanu Sudarsono, James Woodcock, Tim Hodgkinson, Paul Powers, Darius Satongar, Ian Drumm, Henric Mattsson, Ian Rattigan, Phil Kamara, Francis Lee, Jos Hirst, the Acoustics Research Centre, The Crescent, the EPSRC and everyone who has participated in this research.

ABSTRACT

Soundfield reproduction has a vast array of applications, yet the ecological validity – and external validity – has been little researched beyond the narrow limits of the physical approach. This mixed model research explored the differing perception of reproduced urban soundfields and urban soundscapes. The methodology analysed the influences of: *auralisation* technique; *soundscape*; *place* affordances; *time*; *space*; and memory – subject *agency* and *expertise*. Three methods of data collection and analysis were employed: a semantic differential analysis and two psycholinguistic methods – one existing and one novel.

The semantic differentials' soundscape dimensions of the public's response from in situ listening were replicated by participants' responses from ambisonic reproduction but not binaural. A generalised soundscape dimension model was proposed that integrates the structural model of appraisal theory with the interpretation of motivation-affordance fit and mediation dimensions.

Different soundscapes were evaluated differently and stimulated differing processes of perception, which in turn effected reproducibility. A focus group was used as well as members of the public and laboratory participants. It was found that experts responded more in terms of source identification.

Existing methods found no significant difference between in situ and ambisonic listening. The novel method found the ambisonic soundfield was described in terms of objects in motion or sound objects whilst the binaural soundfield and in situ soundscape were described in terms of sources or activity. As an assessment of external validity, the novel psycholinguistic method found that binaural reproduction held validity over ambisonic reproduction.

An in situ real-time binaural reproduction test sought to isolate the 'electroacoustic ear' – the findings were consolidated and discussed in terms of embodiment and ventriloquism. The novel psycholinguistic analysis provided a more accurate representation of the cognitive process of soundscape perception and is offered as a tool for the external validity assessment of urban soundfield reproduction.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Acoustic ecology	Study of sonic relationships between living beings and their environment
Auralisation	Purposeful sound creation, generally electroacoustic
Component	Factor produced by a PCA factor analysis
Earwitness	The subject of a soundscape
Ecological validity	Degree to which a study represents the 'real world' ecology being investigated
External validity	Degree to which a study represents the 'real world' phenomenon being investigated
In situ	In field, generally relating to an 'original' soundscape
KMO	Kaiser-Meyer-Olkin measure of sampling adequacy
Non-user	Person not actively using a space – under test conditions
Place	Sociocultural environment and its associated non-specific spatial context
Reproduction	Electroacoustic auralisation of a recording

PCA	Principal component analysis
Soundfield	Spatial sound pressure of a sound environment as presented to the ears of a listener
Soundscape	Sonic vista, a dimension of experience
Soundscape research	Soundscape-related research and practice motivated primarily by urban design and policy
Soundscape studies	Broad field of soundscape-related research and practice
Space	Cartesian space – specific
User	A person actively using a space

CHAPTER ONE

1 INTRODUCTION

This chapter begins with the **Research Aim**. The notion of **Soundscape** – and the development of its definition – is introduced. The **Thesis Structure** is outlined.

1.1 Research Aim

Soundfield reproduction has a vast array of applications, yet the validity of soundfield reproduction for the purposes of evaluating soundscapes has been under researched beyond the narrow limits of the physical approach. The aim of this research is focused on the perception of a ***soundfield out of context***, namely electroacoustic reproduction, and how such perception differs to that of an ***in situ soundscape***. The methodology was conceived around the identification of several abstractable variables: *auralisation* technique; *soundscape*; *place* affordances; *time*; *space*; and memory – subject *agency* and *expertise*. On the contrast of reproduction listening and in situ listening:

- Which existing method best observes differences in perception?
- How can one observe further differences in perception?
- What are the implications for the methods by which ecological validity is assessed?
- What is the influence, if any, of
    - The original soundscape itself?
    - The earwitness's place-relationship?
    - The reproduction technique?
- Which has a greater role in the differing perceptions, the new context or the electroacoustic ear?

1.2 Soundscape

A soundscape is a sonic vista, a dimension of experience.

The term ‘soundscape’ was first defined as a sonic environment, and the surrounding research pioneered, by R. Murray Schafer in the 1960s [1]. The American composer and artist John Cage declared: “music is sounds, sounds around us whether we’re in or out of concert halls” [2]. With acknowledgement to Cage, Schafer understood the world as ‘a macrocosmic composition’ of which we are all composers - transcending traditional definitions of music.

Current definitions and understanding of the term soundscape vary, and do so with the varying motives and disciplines of those who use it. A soundscape is often deemed synonymous with a *sound environment* – external, objective and readily accessible by natural science. A *soundfield* – retaining the objectivity of a sound environment whilst acknowledging the presence of a listener – is understood as the sound pressure at the ears of the listener (i.e. a spilt second before subjectivity). A *soundscape* incorporates both the objective and subjective. An ISO working group, WG 54 – grounded towards the environmental acoustics end of soundscape research – has only very recently overcome an objective bias: the definition has progressed from ‘a sound environment with an accent on perception’ to something experience centred. The notion of soundscape as *experience* is truly embraced by Husserlian perspectives, most commonly adopted by artists and critical

theorists[1]. In Husserlian perspectives the Cartesian framework for space-time is rejected in favour of a flexible experience-centric space-time. For example, a construct of place is not defined by metrics of space but by users' trajectories and perspectives – individual or collective interpretations of place identity. Given a soundscape is, in essence, a dimension of experience, this phenomenological stance is arguably the most conceptually sound. However, it does in practice alienate a large portion of soundscape researchers and practitioners. On balance, the (overarching and uniting) soundscape definition should be experience centred without prescribing a single philosophy of experience. This is achieved by the definition of a soundscape as a sonic vista.

As for the etymology of the word 'soundscape', the first syllable is self-evident. 'Landscape' was first introduced into the English language describing a picture representing a view or vista of land. Usage may later have evolved the word to describe land as an object but the original and dominant meaning refers to a framing or window of visual perception. A soundscape is the auditory equivalent: a sonic vista. The fact that landscape paintings were of outdoors (rather than portraits; still life etcetera) does give a skew towards the macro. (The origin of this bias can also be traced back to the naturalist sentiments of Murray

[1] Edmund Husserl believed that experience is the source of all knowledge and established the school of thought of phenomenology [97].

Schafer.) The word soundscape still carries this bias but this should be overcome since soundscapes can be macro, meso and micro.

The understanding maintained through this research is that a soundscape it is the framing of an auditory perception. Placing omen on the physical environment – or, conversely, the associated mindscape - enforces an epistemological bias and hence a disciplinary bias. It is experience that bridges individual and shared realities, and soundscapes are a dimension of that experience.

This work may refer to, say, the 'Salford soundscape'. To be particular, this is not a reference to an empirical array of sound sources and impulse responses but is instead a reference to the common (or communal) soundscape. The sonic experience will inevitably vary from person to person yet the Salford soundscape can be understood as that comprised of the *dwellers' perspective*[2] alongside other perspectives.

A soundscape is inherently contextual. The power of electroacoustic reproduction is to disembody and re-contextualise sound. A soundfield recording may be taken from a

[2] The perspective influenced by the cultural norms and common past experiences of dwellers . This stance opposes that which supposes the world is wrapped in the cultural norms and truths that are prescribed to it [98].

soundscape, and when the soundfield is reproduced, a new soundscape is created - from the modified soundscape of the new space. This soundscape is experienced afresh with new motivation, agency and memory in the sociocultural backdrop of a new place.

1.3 Thesis Structure

The **LITERATURE REVIEW** begins by establishing **Soundscape Research as Cross-disciplinary** and voicing the demands for cross-disciplinary methodologies. **Electroacoustic Reproduction in the Broader Context of Sonic Practice and Study** is briefly summarised. This is followed with an overview of **Soundscape Perception**, including reference to Listening States and Multimodality as well as introducing Appraisal Theory. An overview of **Overview of** Soundscape Evaluation approaches introduces two methods of particular relevance, **Semantic Differential Analysis** and **a Psycholinguistic Approach**, which are then detailed in dedicated sections. These two methods are discussed further with regards to their application in the assessment of the **Validity of Reproduction** – in other words, a means of comparing the perceptions of *in situ* soundscapes and *reproduction* soundscapes. The literature review concludes with a **Summary of Research Motivations**.

The **DESIGN** chapter develops first the **Methodology,** establishing the **Development of the Research Question**, the **Research Paradigm: Pragmatism** and **Approach: Mixed Model Research** and the conceptualisation of the **Methodology to Method** design process. **The Three Phases of Experimentation** are introduced. The following sections detail the design considerations: **Reproduction Techniques**; **Soundscape Selection**; **Focus Group**; **Recruiting Participants**; and **Questionnaires**.

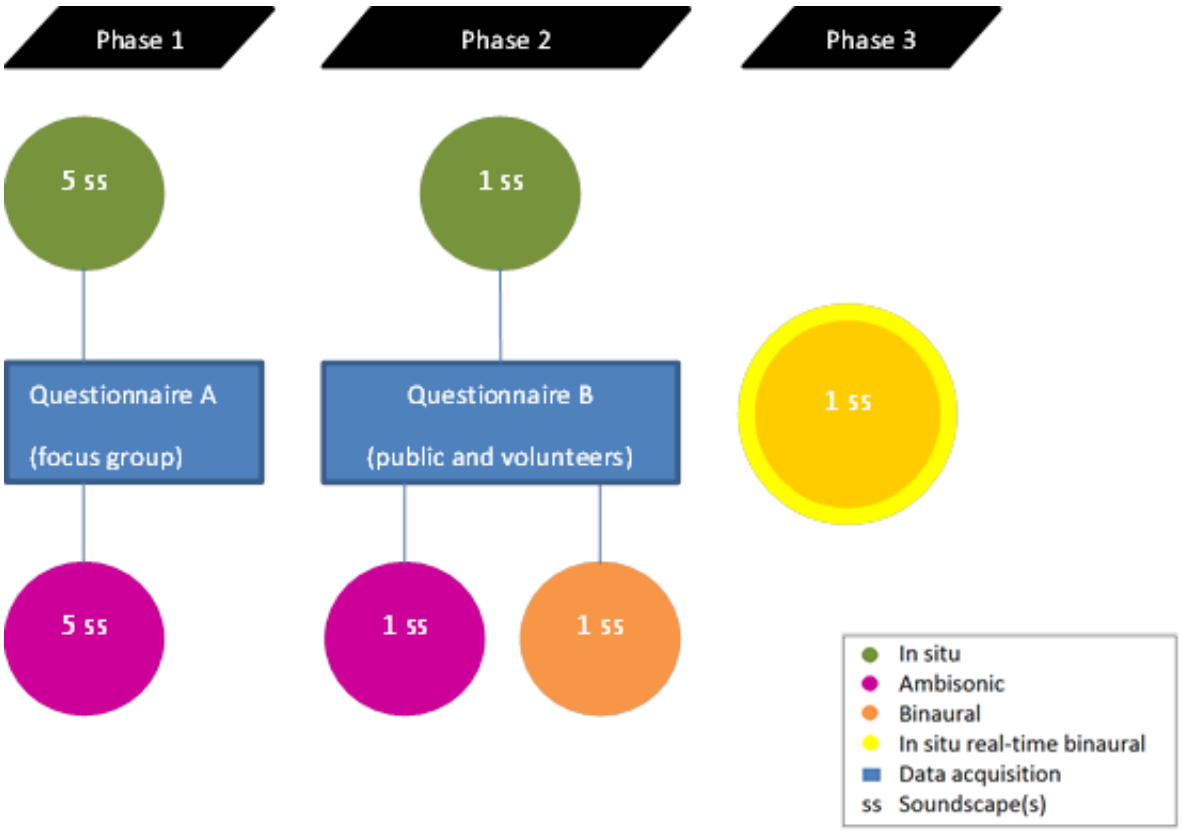


Figure 1 - The three phases of experimentation: focus group with 5 soundscapes and ambisonic reproductions; public at MediaCity soundscape and participants with ambisonic and binaural MediaCity soundfield reproductions; and public with in situ real-time binaural reproduction of MediaCity soundfield.

The **DISCUSSION AND RESULTS** chapter is structured by the three phases of experimentation. Each of the three sub-sections includes the within-phase analyses. The findings are later consolidated thematically in the fifth chapter.

PHASE ONE analysed the focus groups' responses at five in situ soundscapes and their ambisonic reproductions in a listening room. The semantic differential responses at one location were compared and 'soundscape dimensions' were procured. An existing method

of ‘ecological validity’ assessment (using three semantic categories) was assimilated with the focus groups’ open responses across all soundscapes. This was repeated with five semantic categories. The developed psycholinguistic analysis was performed on the all soundscapes’ responses individually. The developed psycholinguistic analysis was then re-appropriated as a method of validity assessment of soundfield reproduction.

PHASE TWO analysed public’s responses at the MediaCity soundscape in situ and participants’ responses in the listening room for an ambisonic reproduction and a binaural reproduction. The semantic differential responses were compared and ‘soundscape dimensions’ were produced. The open responses were categorised into four semantic categories and – prior to the principal component analysis of the developed psycholinguistic analyses – a Tukey (HSD) analysis was performed in conjunction with a one-way ANOVA.

PHASE THREE analysed the public’s responses from the real-time binaural reproduction listening test in situ at MediaCity. Semantic differentials’ soundscape dimensions were analysed and the developed psycholinguistic analysis was performed and discussed.

The **CONSOLIDATION** chapter discusses findings from all phases thematically. The semantic differentials’ soundscape dimensions are re-interpreted by relating the **Semantic Differentials and Appraisal Theory**. The **External Validity of Soundfield Reproduction** is discussed in relation to both semantic differentials’ soundscape dimensions and the

developed psycholinguistic analysis before comparing ambisonic and binaural reproduction. Phases Two and Three's results are discussed in terms of **Embodiment and Ventriloquism**. The roles and influences of the variables around which the methodology was inspired/designed – Error! Reference source not found. – are discussed in light of the esearch findings.

The **CONCLUSIONS Overview** the research project and provide a **Summary of Findings** and **Future Work and Recommendations**.

CHAPTER TWO

2 LITERATURE REVIEW

The literature review begins by establishing **Soundscape Research as Cross-disciplinary** and voicing the demands for cross-disciplinary methodologies. **Electroacoustic Reproduction in the Broader Context of Sonic Practice and Study** is briefly summarised. This is followed with an overview of **Soundscape Perception**, including reference to Listening States and Multimodality as well as introducing Appraisal Theory. An overview of **Overview of** Soundscape Evaluation approaches introduces two methods of particular relevance, **Semantic Differential Analysis** and **a Psycholinguistic Approach**, which are then detailed in dedicated sections. These two methods are discussed further with regards to their application in the assessment of the **Validity of Reproduction** – in other words, a means of comparing the perceptions of *in situ* soundscapes and *reproduction* soundscapes. The literature review concludes with a **Summary of Research Motivations**.

2.1 Soundscape Research as Cross-disciplinary

Soundscape studies, soundscape research and acoustic ecology are all inherently cross-disciplinary yet much research on the perception of soundscapes still lacks sufficiently cross-disciplinary methodologies. The design and use of cross-disciplinary methodologies is paramount – a matter that is frequently voiced in the research community.

Westerkamp aligns the realms of acoustic ecology and soundscape composition, both curating place with sound [3]. The former seeks soundscapes that are a positive extension of a place whilst the latter represents a place with sound. Expanding this theory to the practice of sound reproduction, one who reproduces a soundfield with the objective of simulation is simultaneously taking on the role of composer, by creating new context.

Acoustic ecology is analysed according to a variety of methodological parameters. '[This presents a case] for the necessity and utility of collaborative work between the sciences and humanities in such interdisciplinary fields' [4]. Truax intricately explained how 'environmental sound is problematic because it points to a blind spot in the dominant paradigm of nearly every discipline which touches on it' [5]. In light of this, the acoustic communicational model was conceived as an interdisciplinary solution [6].

Zhang and Kang highlight the range of aspects that need to be considered in the study of soundscapes.

> *"A number of aspects need to be considered, including acoustic- psychological- social characteristics of various sounds, acoustic effects of space boundaries and elements, social- demographic characteristics of users, and general physical and environmental conditions."* [7]

Schulte-Fortkamp comments on the established overlap of soundscape research and the field of community noise and sound quality, identifying the complexities brought to the latter field of research.

> *"It is now about [seventeen] years that Soundscape came into the field of community noise and sound quality. The multidimensional soundscape approach puts emphasis on the way the acoustic environment is perceived, experienced and understood by the individual and by society."* [8]

Mossberg seeks holistic comprehension. This is something that multi-disciplinarity cannot achieve: inter- and trans- disciplinarity is required.

> *"Soundscape research addresses many interdependent areas and touches upon health as well as philosophical, aesthetic, and technical issues. To get a holistic comprehension, these perspectives need to be synchronized."* [9]

Kull calls for the integration of different methods and models, designs which are not only mixed method but mixed model.

> *"There is an urgent need for the use of innovative designs which integrate the different levels of current analyses (qualitative and quantitative; individual and aggregate levels)."* [10]

Schulte-Fortkamp re-iterates the need for interdisciplinary mixed methods, particularly when dealing with perception.

> *"Interdisciplinary procedures are needed that include acoustics, physics, psychology, and sociology when a survey on perception of acoustic environments is carried out under the aspect of comfort."* [11]

2.2 Electroacoustic Reproduction in the Broader Context of Sonic Practice and Study

Soundscape research is of course not the only field concerned with electroacoustic reproduction. In the broader context, the motives for electroacoustic reproduction span from the replication (e.g. simulation) of a perception to the design of new perceptions (e.g. performance). Furthermore the sound- visual- weighting, or medium dominance, spans from auditory dominance (e.g. soundscape research) to visual dominance (e.g. film) to equal dominance (e.g. simulation and particular forms of sound art). All such practices require and seek understanding of re-contextualised sound and the perceptions it stimulates. This

section reviews along the two spectrums: motives (re-creation to creation); and medium dominance.

Research on spatial audio reproduction techniques is primarily concerned with envelopment and localisation. Practicality is also researched in the cases of loudspeaker arrays. For example, enlarging the sweet spot of ambisonic reproduction and hence the 'usable' space; improving focused sources with reproductions such as wavefield synthesis [12]; perceptual testing aimed at limiting the resources (e.g. loudspeakers, space) required for acceptable perceived sound quality [13].

In the case of virtual reality (and game design), the aim is to simulate environments and their perceptions, or create new environments that are perceived and navigated 'naturally' – as one would a 'real' environment, with the visual-auditory dominance reflecting that of the equivalent 'real' experience. Envelopment and localisation are both key. Further attention is given to the audio-visual relationship – aligning the perceived space and depth of environments and sources.

Sound design for film is generally a congruous embellishment of the visual. This could be the sounding of objects, action, environments, emotions, and transition [14]. Whilst acknowledging the importance of the image-sound relationship in the process of sound design, it still remains that the image comes first (and indeed did, historically).

Over the last century art became less and less synonymous with visual art, owing not least to the development and growth of sound art. The broadly accepted definition of sound art has been worded by Angus Carlyle as art on the subject of sound, regardless of medium(s). (Visual) artists – most notably the futurist Luigi Russolo [15] and the followers of Marcel DuChamp – grew jealous of composers' ability to paint on the time dimension, and turned to sound.

Schulz commented that 'in the course of the past two decades, on the frontier between the visual arts and music, an art form has developed in which sound has become material within the context of an expanded concept of sculpture' [16]. This is, in particular, a reference to Klangkunst, a genre of art that is principally concerned with the relationship between sound and site (with equal dominance status). 'Space and site has through modern technology developed from being a container of music to something that might be articulated through sound' [17]. Klangkunst installations curate sound-space relationships. The relocation of sound, by electroacoustic reproduction, stimulates perception relating to the sound itself, sound origin, space (acoustic), place (sociocultural) and memory (of the subject). Put crudely, the whole point is electroacoustic reproduction creates new perceptions.

2.3 Soundscape Perception

A large portion of research on environmental sound is concerned with the trauma it produces [18]. Within the field of acoustical engineering this trauma is considered an effect – *the* perception, *the* 'human response' – and termed 'annoyance'. Such an approach is useful in some specific scenarios and does lend itself well to informing policy given the quantitative, and extremely simplified, nature of its findings. The native SPL->annoyance model is increasingly being extended to an acoustic-parameters->annoyance model (e.g. [19]) in light of past findings that a reduction in sound pressure level (SPL) does not necessarily lead to increased comfort in urban areas [20].

Whilst acoustic engineering focuses on the environment, the audiology and psychoacoustics realm focuses on the ear and auditory processing. These fields offer models of perceptual organization such as cognitive approaches to audition [21], auditory scene analysis [22] and listening theory [23]. In brief, the process by which the auditory system perceives sound is open to the influence of the individuals' attentional set – that is, what they are motivated to focus upon [22]. As a consequence, a less rigorously sequential, and more schematic, organization occurs, increasingly based on 'expectation, prior knowledge and experience' [22]. The ear-witnesses' own subjectivity and memory come into play to form individual associations [24]. The focus upon particular auditory streams can be based upon the individual's knowledge, familiarity, memory, context, expectation and associated meanings of the sound [25][26]. This encapsulates the social and cultural elements of environment –

elements that often shape 'common rules' of perception of sounds [27][28]. Payne performed a study that found free form responses largely varied in their reference to a) the place, soundscape, or individual sounds, and b) the individual's moods and desires, or the temporality of the sound(scape) [29].

2.3.1 Listening States

Various models of listening states exist. The three most relevant are introduced: *musical* and *everyday*; *holistic* and *descriptive*; *listening-in-search*, *listening-in-readiness*, and *background listening*. Gaver introduced *musical* listening and *everyday* listening [30]. Musical listening observes sound as an entity (e.g. timbre), whilst everyday listening focuses not on the sound itself but the associated source (be it action or object).

Raimbault et al introduced two states of listening: *holistic* listening and *descriptive* listening. Holistic listening occurs when the individual processes the soundscape as a whole - focussing upon their own subjectivity in relation to their current activity [31]. Descriptive listening focuses upon the meaning of the sound in relation to objects – physical sources within the environment. The perception of singular sounds differs to the combination of a number of sounds as the acoustic properties are altered and the perceptual processing of the stimuli varies [32].

Truax introduced three states of listening: listening-in-search, listening-in-readiness and background listening [6]. Listening-in-search is an analytical listening that has the ability to focus on one sound to the exclusion of others. Listening-in-readiness depends on associations being built up over time, so that the sounds are familiar and can be readily identified even by "background" processing in the brain [6]. Background listening is where the sounds remain in the background of our attention [6] – when one is not seeking a sound or is in a position where a sound is unlikely to have significance.

2.3.2 Multimodality

Sound is not perceived in isolation, and is merely one facet of multimodal experience. Tim Ingold, for one, criticises the isolation of the aural sense inherent in soundscape studies [33]. In defence of soundscape studies, it provides cross-disciplinary – and hence robust – understandings of a sense and dimension of experience which have been 'neglected since the Enlightenment' [2] [3].

[3] It is merely a question of disciplinary boundaries and their temporality – regular reshuffling is necessary. When the sonic begins to carry the understanding and representation – relative to the visual – that is warranted, new 'cross-disciplines' will form and a sonic-visual divide will give way to new disciplinary boundaries.

Indeed the perception of an environment is multisensory – visual, auditory, tactile, olfactory, and proprioceptive [31]. Other senses – and also other sounds – can influence which sounds are perceived, by directing the attentional set of the listener [34]. This is an example of cognitive interaction between the senses. The processes of intermodal coordination can also occur pre-cognition [35] and at a peripheral, as well as central, cognitive level [36].

The perceptual output is the product of a unique combination of visual and auditory information, in which both vie for perceptual dominance – this is known as the McGurk effect [37]. Dependencies of perception in one modality on stimuli in another modality demonstrate inter-sensory bias [38]. It is the relative uncertainty of modalities that governs the extent of such bias (or audio-visual interaction) [39]. Cases of visual bias have been greatly researched and more recently there have been studies demonstrating cases of auditory bias. Such research is generally concerned with the localisation and the perception of motion, but also extends to appraisal.

Visual parameters have a significant influence on the perception and appraisal of sound. For example, the research of Maffiolo et al indicated that a positive evaluation of the landscape reduces annoyance directed at the soundscape and a negative evaluation of the landscape increases such annoyance [40]. Carles et al. conducted a laboratory study where 75 subjects were presented with six corresponding sound recordings and images in all 36 possible

combinations (after first rating each stimulus separately) [41]. It was found that the congruency between sound and image influenced preferences: coherent combinations were rated higher than the mean of the component stimuli.

So the evaluation of a soundscape varies depending on how sounds are perceived – i.e. alongside other stimuli – and evaluation of the soundscape and overall environment varies depending on the congruency between the perceived sensory information [41][42][43]. This is not only of interest for in situ perceptions, but for the perception of reproduction soundscapes.

2.3.3 Appraisal Theory

Appraisal theory in psychology is the theory that seeks to understand how people react to things differently. Although appraisal theory may occur in environmental psychology approaches it has not been discussed with emphasis on audition.

There are two main model types: structural models based on the notion that evaluations or appraisals (on a continuous spectrum) determine (discrete) emotions; and process models that develop the structural models to capture the dynamic nature of emotion (i.e. to form a sequential process) [44]. Smith and Kirby's model of appraisal proposes associative processing and reasoning work in parallel in reaction to perceptual stimuli [45]. Associative

processing is a memory-based process that makes quick connections and provides appraisal information based on activated memories that are quickly associated with the given stimulus; whilst reasoning is a slower, more deliberate, and thorough process that involves logical, critical thinking about the stimuli [44].

Scherer's multi-level sequential check model "includes: 1) a relevance (novelty and relevance to goals) check, 2) followed by an implication check (cause, goal conduciveness, and urgency), then 3) coping potential check (control and power), and finally the 4) check for normative significance (compatibility with one's standards)" [46].

Roseman discussed "the superiority of the cognitive approach to the causation of emotions over the behavioural and physiological approaches" [47]. The cognitive approach focuses on the extraction of emotion from interpretations of events rather than the events themselves. Five cognition-emotion dimensions were proposed: 1) motivational state: appetitive/aversive; 2) situational state: present/absent; 3) probability: certain/uncertain; 4) legitimacy: deserved/undeserved; and 5) agency: circumstances/other/self. The theory was supported by various tests including one where 120 subjects read stories and rated the intensity of the emotions felt by the protagonists.

Roseman later revisited the structural theory of appraisal with a more 'accurate and comprehensive' theory consisting of two components: 1) motive consistency – when one

evaluates a situation as (in)consistent with one's goals; and 2) evaluation of responsibility or accountability [48].

There are various models of appraisal theory. The majority identify a distinction of associative processing and reasoning (or *mediation*). Furthermore, within the associative processing there is a theme of a *motivation-affordance fit*.

With regards to 'sound quality', similar themes appear. Product sound quality was defined by Blauert as the 'adequacy of a sound in the context of a specific technical goal and/or task' [49]. More recently Jekosch has defined it as the 'result of an assessment of the perceived auditory nature of a sound with respect to its desired nature' [50]. Sound quality can be understood as having three main aspects: stimulus-response compatibility (functionality); pleasantness ('based on an instantaneous overall impression emerging from various sound attributes' [7]); and how identifiable the sounds and their sources [51].

2.4 Overview of Soundscape Evaluation

Objective measures include L_{Aeq}, L_{day}, $L_{evening}$, L_{night}, L_{den} (a time-of-day weighted $L_{Aeq,24hrs}$), L_{A95}, L_{A90}, L_{A50}, Memoli's slope, L_{Amax}, L_{Amin}, L_{A10}, and N(Zwickers loudness) [31]. These are only of use if there is a person-sound relationship that may warrant them (for example, a policy relating L_{Amax} and safety). The backbone of soundscape evaluation is

subjective. Subjective soundscape evaluation scenarios are introduced, followed by evaluation data.

Evaluation scenarios that engage the public in field tests offer the perceptions of the 'experts'[4] – the *users*. It is, however, place-specific and does not take into account how soundscapes are journeyed through and the inherent temporality.

Schafer deemed soundwalking to be the only reliable means of evaluating soundscapes [1]. Soundwalking has the benefit of being journey/experience specific rather than place specific. The downsides are more than practical since the soundwalkers are not actively using a place with typical agency or motivation: they have been taken somewhere and told to listen.

Listener-centred approaches included sound diaries and the online collating of uploaded user recordings. Fiebig et al [52] sourced sound diaries from residents living on a promenade in Berlin, Germany, along with measurements of acoustic parameters. Foale [53] collected sound diaries from residents living in Greater Manchester, UK, with a focus on post-recording interviews. Mydlarz [54] used mobile devices allowing users to upload recordings

[4] To clarify, 'experts' refers here to Schafer's understanding that the earwitnesses using and experiencing a soundscape are the experts on that soundscape

in situ along with soundscape evaluations. The benefit of such listener centred approaches is that the soundscapes being evaluated are those with which the subject has engaged during their daily routines without the influence of the researcher. However, in the case of polyphonic reproduction, it is impractical to equip the subjects with, say, Ambisonic or binaural recording devices.

Soundfield reproduction is used for laboratory listening tests. The benefits are practicality and the ability to control sound material variation from subject to subject as well as present a variety of soundscapes for the purposes of comparison. There is also the potential to represent the temporality of the soundscape, for example Drechsler et al worked daily/seasonal variations into the recordings [55]. As with soundwalking, the subjects lack the agency and motivation of *users* in situ. Furthermore there is the potential influence of the incongruent ecology.

As for forms of subjective measurement, the simplest is perhaps the use of fixed categorical response answers. These could be 'yes' and 'no' or a broad range of descriptors. These have also been used to categorise sounds within a soundscape. Categorical responses are easy and quick for the participant and straight forward to analyse, albeit limited in their scope.

Enquiries consisting of semantic differentials and Likert scales are relatively quick to complete. The quantitative data makes for simple analysis. However, care must be taken in

the selection of the descriptors since their a priori nature will inherently influence the subject's response.

Questionnaires can be structured or semi-structured and conducted in situ, using a soundfield reproduction or retrospectively elsewhere. Open responses provide richer and fuller information and the questioning itself is more realistic [56]. However, open responses – unlike closed responses – are challenging to code or quantify. Interviews have the further potential to be un-structured and thus directed by the interviewee and the themes they deem significant. This suits situations where the interviewee is the expert [57].However, if different participants are asked different questions it is increasingly infeasible to make numerical comparisons [56].

2.5 Semantic Differential Analysis

Semantic differentials are a popular method for assessing the perception of urban soundscapes [52, 54–59], owing to the quantitative nature of the data and hence the ease of comparison with acoustic parameters. The responses can be statistically processed, by means of a principal component analysis (PCA), to identify the underlying components of variance in the responses. These components offer a dimensional model of how the soundscape was perceptually evaluated. The motivation is to provide a useful framework on which to rate and categorise soundscapes.

Kang and Zhang conducted a large scale field survey at two urban soundscapes in Sheffield [59]. A questionnaire consisting of 18 semantic differentials was used to assess subjects' evaluations of the soundscapes. A principal component analysis (PCA) of the data collected produced a four-dimensional space. These dimensions were identified as 1) relaxation/calmness (26%); 2) dynamics/vibrancy (12%); 3) communication (8%); and 4) spatiality (7%).

The study used eighteen indices with seven-point bipolar rating scales [60]. The attributes included connotative meanings of urban environment sounds, such as calming–agitating, interesting–boring and like–dislike, as well as denotative meanings, such as quiet–noisy, sharp–flat and smooth–rough [59]. Therefore subjects responded with regards to the sounds themselves as well as the general impressions and effects. A total of 491 responses were collected. The participants were all members of the public. The study focused on urban green spaces and was therefore limited in the range of soundscapes investigated. It is uncertain whether the dimensions produced are applicable to the broader range of soundscapes experienced by urban dwellers on a daily basis.

In a related study by Kang, the same semantic differential analysis of urban soundscapes was conducted with 223 architectural students at four open public squares in Sheffield. The study found no significant difference between the responses of the students and the public

in the previous study, although the students' responses showed 'slightly more factors and […] more variations between different squares.' [60]

Axelsson et al conducted a similar study that produced the dimensions pleasantness, eventfulness, and familiarity [63]. The key difference from Kang's study is that Axelsson et al used binaural recordings reproduced through headphones in a university listening room. 100 participants were each played 5 – of a total of 50 – 30s excerpts of binaural recordings of (London and Stockholm) urban soundscapes. The participants were architecture university students (with a mean age of 25.6 years) but it is uncertain how well versed the students were in the fields relating to soundscapes, perception and acoustics. 116 semantic scales were used. Rather than consisting of semantic dichotomies, as per Kang's study, Likert scales measured how well a given single attribute was perceived to match the soundscape on a scale of 0-100%. Another key difference from Kang's study was the attribute selection. Descriptive adjectives such as 'loud' or 'sharp' were not included in an effort to avoid encouraging the evaluation of individual sounds [63][64]. Instead a general evaluation was actively encouraged. Therefore there were very few of the 116 scales that could be attributed to the entity of sound. The external validity of the binaural reproduction was not assessed.

Semantic differential analysis provides a model of soundscape perception, albeit specific to a particular set of a priori semantic differentials used. (Throughout this thesis the

‘soundscape dimensions’ procured from a semantic differential analysis are referred to as the *semantic differentials’ soundscape dimensions*.)

2.6 Psycholinguistic Approach

The cross-disciplinary approach of psycholinguistics allows for a more nuanced understanding of sound and sound products as they appear within the soundscape [65]. Psycholinguistics is the study of cognition theoretically grounded in the scientific fields of psychology and linguistics. The work of Dubois (and collaborators) on semantic categories contributed to the ‘understanding of the psychological principles of noise categorization, beyond the physical similarity of sounds’ [66]. The research highlights ‘the need for a shift from physical descriptions to cognitive ones’ [66]. A central finding is ‘that one and the “same” acoustic phenomenon could be classified according either to the source that produces it or to the action generating the noise (this is the case, for instance, with the squeaking of a door, which can be categorized either with “noises of doors” – ‘shutting, slamming, opening of a door’ – or with other instances of grincement ‘squeaking’ – of doors, of windows, or of steps)’ [66]. The former is sound-in-relation-to-object whist the latter is sound-in-relation-to-action (or perhaps sound-as-an-entity).

The same mass study also observed inter-individual variation: ‘acousticians and members of the laboratory sorted noises according to their physical “objective” properties, such as

pitch or temporal evolution, while others sorted the noises according to the sources that produced the acoustic phenomena' [66]. Therefore acoustic-/experiment- related knowledge has an observable influence on subjects' interpretations of individual sounds and soundscapes.

Raimbault conducted in situ surveys in two French cities with a questionnaire that included semantic scales, each followed by an open question relating to the semantic scale response [67]. This enabled a psycholinguistic analysis of the temporal, spatial and activity nature of appraisals and thus a retrospective assessment of the validity of the semantic scales used. Spatiality was found to be judged comparatively thus "underlined the importance to precisely study the referred object or situation in order to assess spatial features" [67]. It was found that the monotonous-varied scale responses related to the listening state adopted by the subject: *monotonous* responses occurred with open responses indicating holistic listening; and *varied* with descriptive listening (i.e. sources/events in relation to discernable objects) [67]. These listening states were shown to be linked to soundscape situation, gender, and subject purpose (or what could be deemed *motivation*).

2.7 Validity of Reproduction

Spatial audio systems are generally evaluated in terms of the localisation of discrete sound sources and the listener's sense of 'envelopment'. The perception of soundscapes out of

context – the very nature of soundfield reproduction – is rarely researched. Two studies are here reviewed, both comparing the differing perceptions of urban soundscapes and ambisonic soundfield reproductions: the first with a semantic differential model and the second with a psycholinguistic model.

Davies et al tested how semantic differentials' soundscape dimensions manifest in the perception of reproduced soundfields by repeating Kang's semantic differential analysis with 15 participants listening to 8-channel ambisonic reproductions of four urban soundfields [59][68]. Whilst Kang et al used Sheffield soundscapes; Davies et al used Manchester and Salford soundscapes. There is yet to be any experimentation that uses the same soundscape to compare soundscape dimensions derived from semantic differential analyses for in situ listening and reproduction listening. The study concluded, by means of general observation, that ambisonic reproduction was able to reproduce the same dimensions as procured from Kang's in situ listening surveys. The a priori nature of the semantic differentials means this method is as much a model of how subjects evaluated along given dimensions as it is a model of cognition.

Guastavino et al offered the categorisation of open responses into three semantic categories as a means of empirically assessing the ecological validity of soundscape reproduction [69]. This approach is 'a linguistic analysis of complex phrasing, rather than a lexical analysis of words [as per the semantic differential analysis], to infer properties of

mental representations of acoustic phenomenon' [70]. A questionnaire of semi-structured questions was conducted at various soundscapes of two French cities. Soundfield recordings were made and reproduced in a university listening room, where (colleague/student) participants completed the same questionnaire. The responses were analysed by categorising words/phrases as subject references, object references or source identification.

This psycholinguistic approach was based on the aforementioned findings of Dubois. The focus of the assessment of ecological validity was of perceived spatial features, hence the analysis differentiated responses to questions regarding the foreground sounds and responses to questions regarding the background noise. Following the categorisation of sound events and background noise, the two were subject to a second process of categorisation:

> *"The verbal data were classified into semantic categories emerging from the answers, namely subject-centred and object-centred descriptions. Subject-centred descriptions refer to the subjects by means of deverbal adjectives (comforting) and complex phrases (has a negative impact on me) constructed on verbs indicating psychological effects. Object-centred descriptions refer to the acoustic phenomenon itself. They include simple adjectives and denominations referring to the physical properties of the*

acoustics signal in terms of frequency (low-pitched), temporal structure (continuous, monotone) or level (loud, intensity).” [70]

It was found that sound events were usually described in relation to ‘the specific sources generating noise’ (76% of occurrences). Background noise was usually described in relation to effects on the subject or the physical properties of the sound itself (named ‘object-centred’[5]) and infrequently in relation to the specific sources generating noise (15% of occurrences). The results of the second stage of categorisation are presented in Figure 2 and Figure 3.

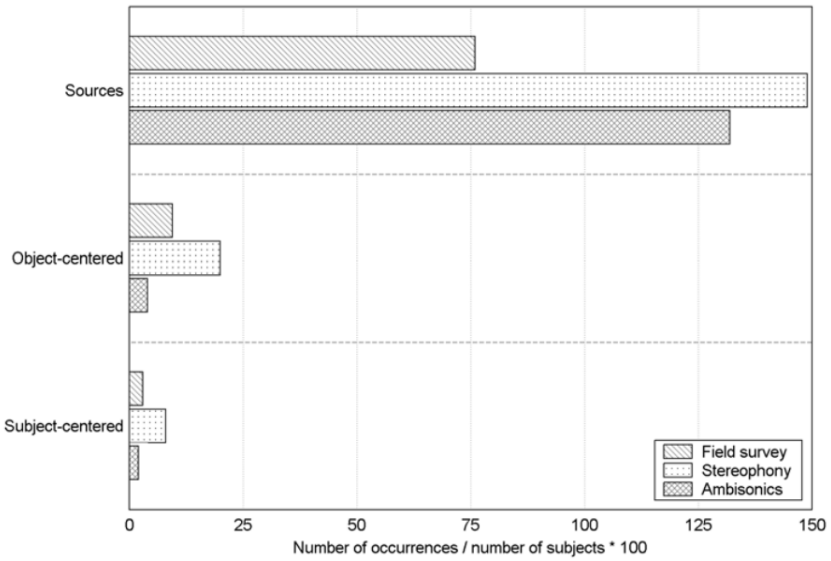


Figure 2 – Comparison of spontaneous descriptions of *source events* classified into semantic categories emerging from the subject responses in the field survey, the stereophonic listening test and the Ambisonics listening test. No significant difference was observed between the three experiments. [70]

[5] ‘Object-centred’ is prescribed a different meaning later in the design section of this thesis.

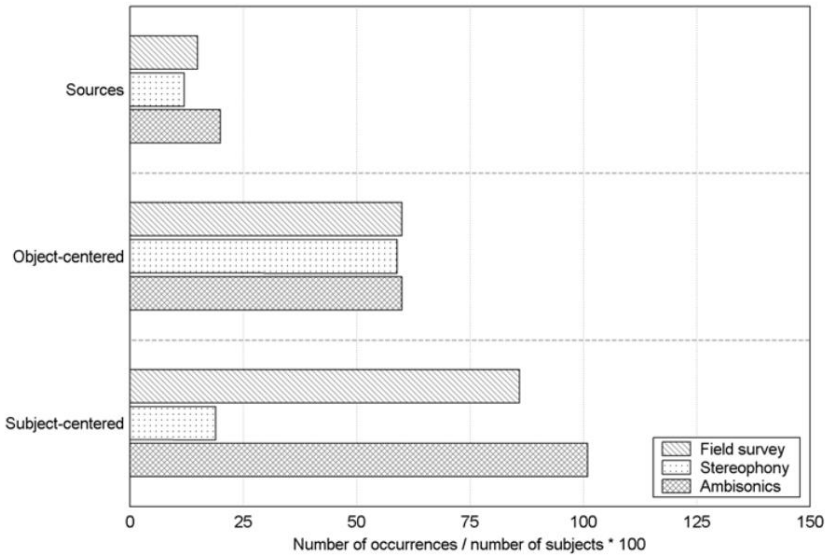


Figure 3 – Comparison of spontaneous descriptions of *background noise* classified into semantic categories emerging from the subject responses in the field survey, the stereophonic listening test and the Ambisonics listening test. A significant difference was observed between the field survey and the stereophonic listening test. No significant difference was observed between the field survey and the Ambisonics listening test. [70]

There was no significant difference in the six datasets (object, subject and sound source counts for foreground and background) between in situ and ambisonic reproduction responses. The study therefore concluded that ambisonic soundfield reproduction is ecologically valid.

The study did not account for the variation in subjects' expertise between the in situ and laboratory tests. In situ subjects were members of the public (most likely naive in relation to acoustics and sound) whilst the reproduction subjects were students and colleagues (most likely 'experts'). Dubois et al conducted a study that found expertise had a significant influence on the linguistic devices – and hence cognitive representations – of soundscapes:

'experts' tended toward descriptive listening whilst 'novices' tended toward holistic listening [71].

The method of analysis itself is limited is several ways. Firstly, the sound event and background noise categories are in fact a priori classifications. The questions themselves determined whether the researcher classified a response in relation to sound events or background noise. There is no other feasible means of distinguishing 'background noise' linguistically. Maffiolo conducted a study that identified the two main categories: event sequences (sequences including specific events) and amorphous sequences (in which no specific event could be isolated) [72]. Amorphous sequences may be common when describing 'background noise' however there is no one-to-one correspondence and such mental representations cannot be identified simply by the occurrence of amorphous sequences.

Simple lexical devices (e.g. 'conversation') used to describe sounds were only treated as sound events (without acknowledging how they relate to the non-sonic). Such lexical devices – or "basic terms" – are limited in number beyond those of human sounds and music [66]. However they do comprise a substantial portion of the sound events described by participants and therefore should be analysed in more depth. For example, the distinction between 'talking' and 'voices' is that of *activity* and *sound entity*. Indeed no such distinction was made between *activity*, *object* and *sound entity* throughout the process of analysis. A

study conducted by Dubois et al stressed that the “same” acoustic phenomenon could be classified according to either the source that produces it or according to the action generating sound [73].

The method has a third – and key – failing. The summation and/or averaging of the semantic category occurrences washes out the information held in the individual cases.

2.8 Summary of Research Motivations

- Soundscape reseearch is a cross-disciplinary field yet much research on the perception of soundscape lacks sufficiently cross-disciplinary methodologies.
- Kang 2010 [74] conducted a large scale field survey using 18 semantic differentials – the results provide a reliable comparision.
- Exisiting studies left scope for investigation of inter-soundscape variation and an indepth comparison of (sound-/acoustic-/experimentation-) expert’s versus user’s soundscape evaluations and perceptions.
- Axelsson did not address the external validity of binaural reproduction (particularly important given the semantic scales were general and hence multimodal rather than sound-specific).
- Existing soundscape studies declare ‘ecological valdity’ of soundfield reproduction whilst electroacoustic reproduction unarguably creates new perceptions. So how

well do exisiting methods model cognitions and how appropriately positioned are the (model-specific) lines between validity and invalidity?

- The psycholinguistic approach provides a compelling means of modelling cognitions of perception.
- Guastavino (2005) employed a process of semantic categorisation of open questionnaire responses to compare the cognitive processes of perception of in situ soundscapes and stereo and ambisonic reproductions for the for the purpose of assessing the 'ecological validity of soundscape reproduction' [59].
- There remains scope to develop this psycholinguistic analysis to focus more clearly on the nature of reproduction.
- All existing study lacks a delineation of ecological validity and external validity, namely the investigation of the role of soundscape, place, and memory (of the subject).

CHAPTER THREE

3 DESIGN

The **Methodology** establishes the **Development of the Research Question**, the **Research Paradigm: Pragmatism** and **Approach: Mixed Model Research** and the conceptualisation of the **Methodology to Method** design process.

The Three Phases of Experimentation are developed. The following sections detail the design considerations: **Reproduction Techniques; Soundscape Selection; Focus Group; Recruiting Participants**; and **Questionnaires**.

3.1 Methodology

The following methodology sections offer the **Development of the Research Question**, **Research Paradigm**, research **Approach**, and how the methodological framework formulated the experimental design.

3.1.1 Development of the Research Question

The central question is how the perception of an urban soundfield reproduced in a room differs from the typical perception of the 'original' soundscape in situ.

It might be argued that such differences in perception are due to nothing more than objective differences between the two soundfields - physical changes to the soundfield incurred during the recording process, for example. However, there are two reasons as to why this is an incomplete and inadequate theory.

Firstly, there is a limit where perceived accuracy plateaus in relation to objective accuracy - in other words, listeners' physical hearing is unable to identify all of the relevant discrepancies between the in situ and reproduced soundfield, and may identify none. Secondly, and most importantly, the soundfield is only one of many variables between the two listening contexts. It is these other variables that form the structure of the research methodology.

A typical reproduction procedure used for the purposes of soundscape research may be as follows: a soundfield recording is made in, say, an urban square; the recording is later reproduced in a room in, for example, a university and presented to a group of participants, possibly colleagues or students familiar with soundscapes, acoustics or subjective test designs. In other words: a soundfield recording is made of a *specific soundscape*; the recording is *auralised* in a new *space* and *time*; presented in a new *place*[6] (a university environment) to a group of people with a new *agency* (motivation for experiencing the new soundscape) and new sound-/acoustic-/experiment-related *memory* engaged in their processes of perception.

Here is a list of the variables identified:

- Soundscape
- Space
- Time
- Auralisation technique
- Place (social/cultural context)

[6] The term 'place' may be commonly understood as geographical, encapsulating 'space' and identity/culture/purpose. Here 'space' is defined as Cartesian and specific; 'place' is defined as a social/cultural environment and its associated (non-specific) spatial context.

- Memory[7]
    - Agency (prior motivation for experiencing soundscape)
    - Expertise (sound-/acoustic-/experiment-related *memory*)

These terms naturally fall into a range of academic disciplines. Therefore a cross-disciplinary approach is required to explore all seven variables.

3.1.2 Research Paradigm: Pragmatism

3.1.2.1 Pragmatism

The term 'pragmatism' was coined by the philosopher Charles Peirce [75]. The origins and developments of pragmatism as a philosophical tradition are still debated today. But to give a sense of the early pragmatist theory of truth, here are C.S. Peirce's and W. James's slogans:

"Truth is the end of inquiry." [76]

"Truth is satisfactory to believe." [75]

From pragmatist philosophy was born pragmatist research. The pragmatist paradigm is now well established as one of the main research paradigms.

[7] The variable *memory* offers completeness but it is too broad and multifaceted a term to be used in relation to the experimentation and was therefore – retrospectively – represented by two variables. `

Pragmatist researchers are not committed to any one system of philosophy or reality. Pragmatism focuses on the 'what' and 'how' of the research problem [77]. Early pragmatists "rejected the scientific notion that social inquiry was able to access the 'truth' about the real world solely by virtue of a single scientific method" [78]. The pragmatist paradigm places "the research problem" as central and applies all approaches to understanding the problem [79]. With the research question central, data collection and analysis methods are chosen as those most likely to provide insights into the question with no philosophical loyalty to any alternative paradigm [78].

Pragmatism is seen as the paradigm that provides the underlying philosophical framework for mixed-methods research [80]. Whilst this consensus may exist, some mixed-methods researchers align themselves philosophically with the transformative paradigm [81], [82]. It may be said, however, that mixed methods could be used with any paradigm [79].

3.1.2.2 Pragmatism within the Philosophy of Science

The pragmatist foundation of this research avows theory generation (e.g. section 5.1) and theoretical corroboration (e.g. section 5.3) outside of *science,* as it is commonly understood. However, a majority of this research can be deemed scientific enquiry and therefore it is helpful to introduce at least one pragmatism-related theory within the *philosophy* of science.

> *"Like the overlapping maps in a Mercator projection, where the ranges of different versions overlap, they predict the same phenomena. But just as there is no flat map that is a good representation of the earth's entire surface, there is no single theory that is a good representation of observations in all situations."* [80]

Stephen Hawking and Leonard Mlodinow coined the term 'model-dependent realism' that rejects absolute certainty and adheres to a multiplicity of meanings and answers. The 'meaningful' element is understood as the 'usefulness' of the model [83], a notion that sits closely to the pragmatist notion of 'truth' as something 'useful to know'.

3.1.2.3 Justification

Pragmatism was chosen for several inter-connected reasons. In the academic field of soundscape research there is a demand for cross-disciplinary understanding, as voiced in section 2.1. Therefore the full cross-discipline complexity of the research question is embraced – rather than limit the enquiry to acoustic parameters of reproduction techniques – and the enquiry has been formulated around seven abstractable and discipline-spanning variables (presented in 3.1.1). Certainly pragmatism provides a platform for the mixed model research to explore these variables. But furthermore, by having no philosophical commitment, the pragmatist researcher has access to an array of potentially useful perspectives.

Soundscape research is a broad and cross-disciplinary field of research. There has long been an acknowledgement of the need for multi-, inter- and, recently, trans-disciplinary methodologies [10]. As might be expected for a relatively young yet broad field there are many fragmentations within soundscape research and under-developed theoretical links between neighbouring disciplines. The pragmatist paradigm of this research aims to help draw together some different perspectives surrounding soundscape perception and soundfield reproduction.

3.1.3 Approach: Mixed Model Research

3.1.3.1 Mixed Models

Mixed research is research that uses qualitative and quantitative techniques. There are two major types: mixed method research and mixed model research. Mixed method research uses a qualitative phase and a quantitative phase. Mixed model research uses qualitative and quantitative approaches mixed within or across the stages of the research process. This research is within-stage mixed model research, exemplified by the use of a questionnaire of open (qualitative) and closed (quantitative) questions [84].

3.1.3.2 Justification

The research problem itself involves perception, place and sound. No single model can fully understand a problem that bridges the subjective and the objective. The inclusion of various

disciplinary approaches could be fruitful: acoustics, psychology, phenomenology, and many more between and beyond. The mixed models employed in this research are by no means exhaustive. (Arguably that would be implausible: based on the assumption that ultimate truth is unattainable, infinite models would be required.) However, given the research problem does not fall into one single model, several models explore the research problem.

A mixed model research approach is required for the purposes of complementarity[8], triangulation[9] and initiation[10] when correlations and irregularities make themselves apparent at the later stages of analysis. The results of one method can be elaborated and clarified by the results of another method; convergence and correspondence can be sought from the results of different methods; and new perspectives can be formed from the discovery of contradiciton.

---

[8] Complementarity seeks elaboration, enhancement, illustration, clarification of the results from different methods [33].

[9] Triangulation seeks convergence, corroboration, correspondence of results from different methods [84].

[10] Initiation seeks the discovery of paradox and contradiction, new perspectives of frameworks, the recasting of questions or results from one method with questions or results from the other method [84].

3.1.4 Methodology to Method

As evidenced in the literature review, the majority of existing soundscape research applies single methods to specific contexts. Less concern is given to how these methods respond to various contexts (for example different soundscapes), or to how various methods' results corroborate or contradict.

The difference between in situ soundscape perception and reproduction soundscape perception is more than merely environmental and in light of this the methodology experiments with various contexts of soundscape perception:

- Electroacoustic reproduction compared to 'real'
- Ambisonic compared to binaural
- Binaural live in situ compared to binaural later elsewhere
- Different soundscapes
- Different subject agencies

The methodology requires several methods and models. Within the traditions of mixed research, the 'mixed research continuum' spans monomethod research and fully mixed research [85]. This design is positioned as 'partially mixed research' on the research continuum: both qualitative and quantitative data were collected - in the form of open responses and semantic differential responses - but the majority of analytical models were statistical (including PCA; ANOVA; G-test). The two existing methods employed are: Kang's

principal component analysis of semantic differentials [84]; and Dubois's semantic categorisation of open responses [59] which was used by Guastavino et al. to assess the 'Ecological Validity of Soundscape Reproduction' [71]. The latter was developed into a new third method. This novel method is referred to in this thesis as the *developed psycholinguistic analysis*. The pragmatist paradigm enabled the modest cross-discipline manoeuvres in Chapter 5: a theory was proposed with regards to 'soundscape dimensions'; and results were discussed in terms of 'embodiment'.

3.1.5 Summary

This research employed a mixed model approach grounded in the pragmatist research paradigm. The methodology is experimental and theoretical, with a strong experimental bias. Within the experimental there is a quantitative bias.

The difference between in situ soundscape perception and reproduction soundscape perception is more than merely acoustic or environmental. The methodology was formulated from the identification of seven variables responsible for the differing perception between in situ and reproduction listening: *auralisation* technique; *soundscape*; *place* affordances; *time*; *space*; and memory – subject *agency* and *expertise*.

3.2 The Three Phases of Experimentation

Phase One compared:

- A focus group's perceptions of five soundscapes 'in situ' and
- The focus group's perceptions of the same five soundscapes reproduced in ambisonics.

A focus group (see 3.5) was used for two reasons. Firstly it remove the confounding factor, subject, between between in situ and reproduction tests. The reproduction test was conducted a sufficient period later than the in situ test, thus ensuring the latter's responses were not significantly influenced by the former. Secondly, the focus group had to be taken to the five soundscape locations therefore they were in 'test conditions' without the agency of the 'users' – the 'agency' variable was thus removed (or at least its effect altered). The use of five varied soundscapes provides comparison between soundscapes.

Phase Two compared:

- the public's perceptions of the in situ MediaCity soundscape,
- further participants' perceptions of an ambisonic reproduction of the MediaCity soundfield, and
- Further participants' perceptions of a binaural reproduction of the MediaCity soundfield.

The public, or ‘users’, were used in situ as a comparison to Phase One’s ‘test conditions’ thus isolating subject agency. Furthermore, in the laboratory listening test practices of soundfield research it is the users’ perspectives that one seeks to gain from the responses of ‘non-user’ participants. Therefore Phase Two’s binaural and ambisonic reproductions’ validity is representative of common practices’ validity – the ability of lab testing with ‘non-users’ to reveal the perspective of in situ ‘users’.

Phase Three focused on:

- The public’s perceptions of a real-time binaural reproduction of the MediaCity soundscape in situ.

Phase Three’s real-time binaural in situ listening sought to isolate the effect of the *electroacoustic ear*. Results bearing closer relation to Phase Two’s binaural listening would suggest it is the reproduction itself that is principally responsible for the differences in perception of in situ and reproduced soundscapes. Results bearing closer relation to Phase Two’s in situ listening would suggest it is the incongruous context of the listening room that is principally responsible for differences in perception.

3.2.1 Phase One: Focus Group

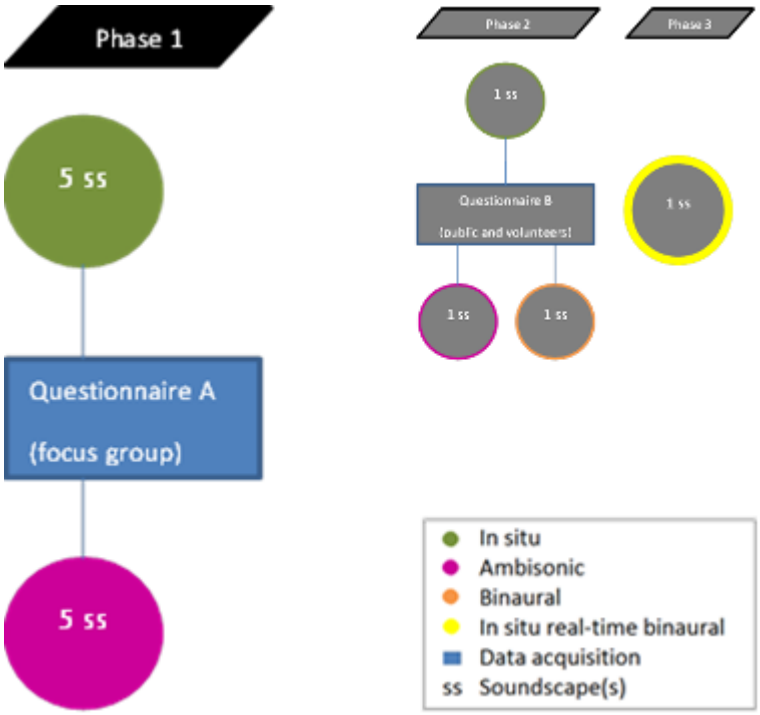


Figure 4 - Phase One of experimentation

3.2.1.1 In Situ Soundscapes

A focus group of five participants were taken to five locations in Salford and Manchester. The locations were chosen such that there was variation in soundscape, activity or purpose, and demographic of users (discussed further in section 3.4).

- the entrance to the **Newton** Building, Peel Park campus, University of Salford, Salford,
- a central foyer of Salford **Precinct**, Salford,
- the public green space at **MediaCity**, Salford,

- the west corner of Manchester's '**Triangle**' and
- the garden of the **Kings** Arms pub, Salford.

The tests took place on an autumn weekday afternoon. The participants were taken, as a group, to the five sites where they individually answered a questionnaire, Questionnaire A, of six open questions. (See section 3.7 and appendix 7.2.1 for the questionnaire design.) Finally, at The Kings Arms they also completed a semantic differentials enquiry of the soundscape. (The semantic differentials enquiry was not used at the first four locations in order to avoid any possible influence on the responses to the following open questions.) At each site, B-format recordings were made with a multi-capsule Soundfield microphone whilst the group completed the questionnaire.

3.2.1.2 Listening Room Reproduction

A few months later, further soundfield recordings (of 10-15 minutes) were made at the same five sites on a weekday afternoon. The ambient and peak sound pressure levels were measured during the period of recording (see section 3.3.1). It is these recordings that were used for the listening room reproduction test. Two-minute clips were extracted from these later recordings. Selection was based on sound quality (since wind noise was an issue) and the typicality of representation. Two minutes has been deemed sufficient for one to perceive and judge a soundscape [70].

The laboratory listening tests took place in the well diffused listening room at the University of Salford (meeting the requirements of BS 6840-13 / IEC 268-13). The recordings were played back through an 8 channel first-order ambisonic loudspeaker configuration (see section 3.3.1). Sound levels were measured and balanced to that of the in situ soundscape using 10 second samples ($L_{Aeq,10s}$). Each participant sat at a desk in the centre of the room. A single ceiling lamp lit the desk such that the loudspeakers, and the rest of the room, were less apparent. The questionnaire (Questionnaire A) was completed with pen and paper, as per the in situ responses. The researcher, illuminated only by a desktop monitor, remained in view in the corner of the room. (Note that the presence of the researcher is consistent with the in situ tests.) The participants were invited one at a time. The clips were played twice with the option of more repeats if desired (although no one chose to do so). The participants were not informed of the origin of the recording - whether the recording was made in their presence during the in situ listening test or not. Research has shown that soundscape assessments can be dependent on the segment of the soundscape that was recorded (when playing back recorded soundscapes in a laboratory situation) [31]. Therefore the playback was of the unheard soundfield recordings so as not to trigger memory of particular (sonic or non-sonic) events, only of the holistic experience.

3.2.2 Phase Two: Ambisonics Versus Binaural

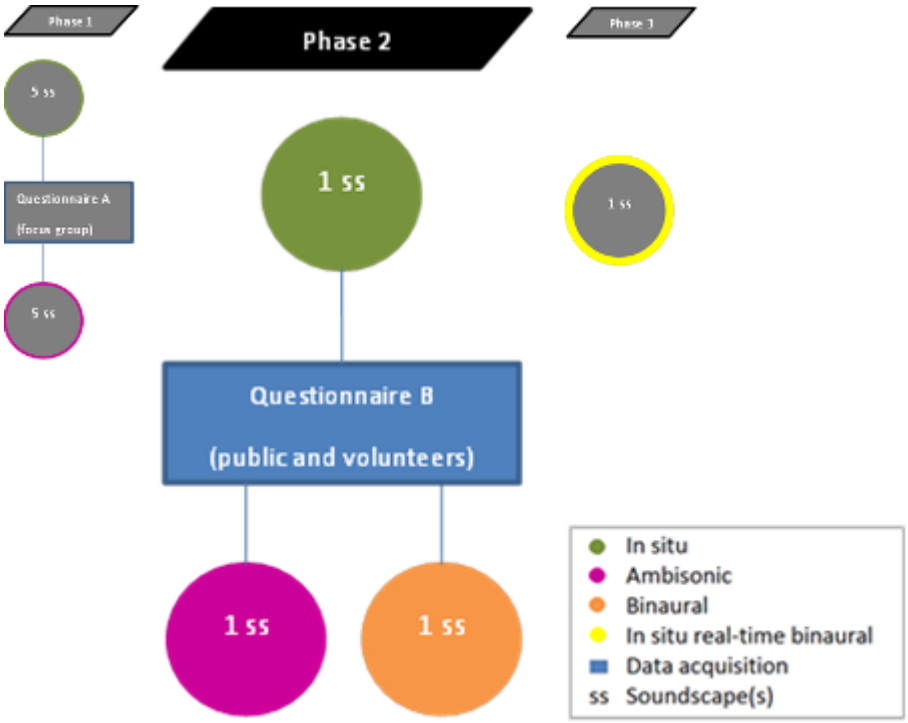


Figure 5 - Phase Two of experimentation

Phase Two consisted firstly of responses from members of the public at MediaCity. These are users of the space who contribute to the soundscape as they actively define the sense of place. Drawing the comparison with Phase One's in situ responses (under 'test conditions') seeks to show the impact of subjects' *expertise* and *motivation*.

Two further tests were performed in the listening room. The first was the Ambisonic reproduction of a MediaCity soundfield recording. The second was the reproduction of a MediaCity binaural recording. The participants were colleagues, staff and students, and thus had a familiar relationship with the acoustic facilities and listening room as per the members

of Phase One's focus group. The participants were invited one at a time to take a seat at a desk in the centre of the listening room. The researcher gave a brief – and selective – summary of the research background. It was important that the participants remained ignorant of the process by which their responses are analysed since such knowledge might influence their wording. The researcher explained the test procedure and what was expected of the subject, and offered the opportunity to decline. For both the ambisonic and binaural tests in the listening room a single ceiling lamp lit the desk such that the soundcard, laptop or loudspeakers and the rest of the room, were less apparent. The questionnaire was completed with pen and paper, as per the in situ responses. The researcher was present in the room thus maintaining continuity with the previous tests.

3.2.2.1 In Situ Soundscape

Eleven participants were surveyed in situ with an extended version of the questionnaire used in the first phase of experiment, Questionnaire B (see 7.2.1). The participants were a random sample of people using the outdoor space at MediaCity (see 7.2.2). Nine out of the eleven participants worked at MediaCity and were taking a lunch break, one was visiting and one was a student living nearby.

3.2.2.2 Listening Room Reproduction: Ambisonics

Twenty participants were surveyed in the listening room with the same questionnaire used in the in situ soundscape test, Questionnaire B. The participants were a random sample of

people working or studying in the Newton Building at the University of Salford. The majority of participants were colleagues working in, or around, the field of acoustics. The reproduction was of the soundfield recording (from Phase One) using the same eight channel Ambisonic configuration but with third order decoding. (See section 3.3.1 for more on the reproduction technique.) The 2min track was on a constant loop, playing as each participant entered and left the listening room.

3.2.2.3 Listening Room Reproduction: Binaural

Fourteen participants were surveyed in the listening room with the same questionnaire, Questionnaire B. The participants were a random sample of people working at, or generally using, the Newton Building at the University of Salford. The majority of participants were colleagues working in, or around, the field of acoustics. There was some crossover with the participants volunteering for the ambisonic reproduction: seven of the fourteen participants had also taken part in the ambisonic test. Given Phase Two's ambisonic and binaural listening tests were conducted several months apart it was assumed that this would have negligible impact on their responses.

Once the test was introduced, each subject fitted a pair of Sennheiser CX300ii in-earphones and 3M ear defenders. An array of new ear buds was made available in various sizes. The 2min track was on a constant loop, playing whilst each participant fitted and removed the earphones.

3.2.3 Phase Three: In Situ Real-Time Reproduction

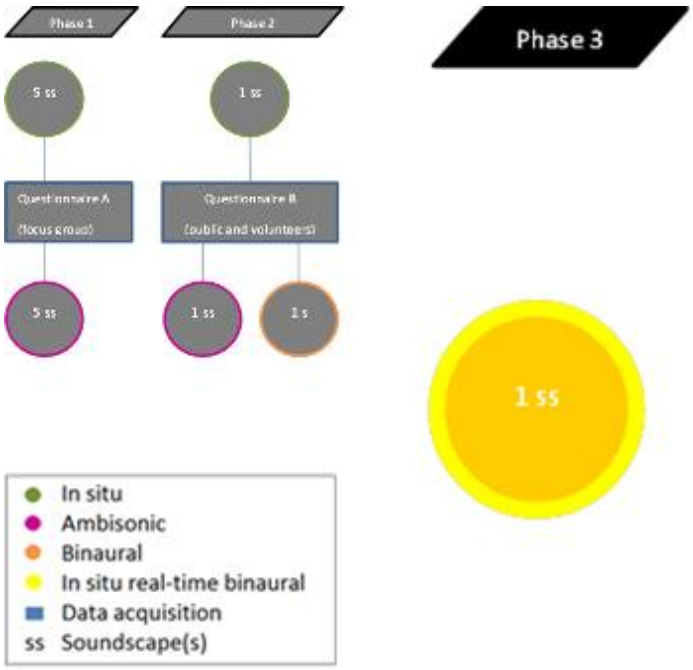


Figure 6 – Phase Three of experimentation

The difference between in situ and reproduction listening experiences was investigated further still here in Phase Three. Binaural technology was used to create a listening experience that is positioned in between that of in situ and lab reproduction listening experiences: real-time binaural reproduction in situ. In other words, the incongruous physical and sociocultural context of the previous listening room reproductions was removed thus isolating the effect of 'electroacoustic ears'.

A real-time binaural reproduction of the MediaCity soundscape was developed. The Brüel & Kjær dummy head rested on a bench, as per the binaural recording procedure of Phase Two (see section 3.3.2). The same pre-amp and 2-channel recorder sent a live stream to the earphones. Participants fitted the earphones (along with the ear defenders) before commencing the questionnaire. The subjects were asked to respond to what they hear, the sound environment.

Seventeen members of the public volunteered to participate. Noteworthy is the tendency for people to assume the subject of the study, and hence the questionnaire, is the reproduction method itself. Indeed much of the participation was owed to the presence of the dummy head and the intrigue it produced. Therefore extra care was taken to ensure the participants understand they are commenting on the 'soundscape heard'. Responses did still occasionally reference the electro-acoustic processing/nature of the experience (as would be expected) but the focus was clearly on the heard soundscape, as desired.

3.3 Reproduction Techniques

Two techniques were chosen for the capture and reproduction of soundfields: Ambisonic and binaural. Various other techniques are available such as mono (loudspeakers or headphones), stereo (loudspeakers or headphones), and wavefield synthesis. For the purposes of this project a spatial reproduction is required to limit the variation in the acoustic signal between in situ and reproduction listening. Wavefield synthesis is still in the stages of development whereas ambisonics may soon meet the mass market. The use of ambisonics allows valid comparisons with the relevant existing research of Kang [86], Guastavino [59]and Davies [70]. Binaural reproduction offers an advanced yet commonly used from of 3D spatial reproduction. It also differs significantly from ambisonics in that one can use headphones thus providing a completely different sonic experience. Binaural reproduction listening test results can be compared to the relevant existing research of Axelsson [69].

3.3.1 Ambisonics

3.3.1.1 Ambisonics: Fundamentals

The ambisonic system was the first technique developed to reproduce an entire soundfield. The simplest form is first-order ambisonic. Extended from this is higher-order ambisonic which splits the soundfield into a series of spherical harmonic functions and in doing so

raises the upper frequency that can be accurately reproduced and expands the space in which the soundfield is accurately produced. The aforementioned space is known as the 'sweet spot' and may be calculated from the system order [63].

Zero order ambisonics only accounts for the pressure component (W). For first order ambisonic recording, the soundfield is stored in 'b-format' consisting of 4 channels, W (omnidirectional information) and X, Y, Z (directional information from figure-of-eight microphones) [87]. These are then derived as:

$$
\begin{aligned}
W &= \sum_{i=1}^{k} s_i \frac{1}{\sqrt{2}} \\
X &= \sum_{i=1}^{k} s_i \cos\theta_i \cos\phi_i \\
Y &= \sum_{i=1}^{k} s_i \sin\theta_i \cos\phi_i \\
Z &= \sum_{i=1}^{k} s_i \sin\phi_i
\end{aligned}
$$

Here *i* is the source, *k* the number of sources, s_i is the *i-th* source mono signal, and θ_i and Φ_i are the horizontal direction (azimuth) and elevation (theta) of the source respectively. The $1/\sqrt{2}$ weighting (on all channels) is an engineering solution required, particularly with synthesised soundfields containing many sound sources, to achieve a better balance of signal levels over the four channels.

Every loudspeaker in the array receives all four channels, each channel weighted according to the position of the loudspeaker:

$$p_j = \frac{1}{L}\left(W\frac{1}{\sqrt{2}} + X\cos\theta_j\cos\phi_j + Y\sin\theta_j\cos\phi_j + Z\sin\phi_j\right)$$

With L, the number of loudspeakers (which must be at least the number of ambisonic channels, four); θ_j and Φ_j, the horizontal direction (azimuth) and elevation (theta) of the *j-th* loudspeaker.

The scalability of ambisonics, with the introduction of higher-orders, allows the soundfield approximation to be optimised for given hardware, ergonomic or packaging constraints. It offers the best accuracy, and hence localisation, without requiring vast numbers of loudspeakers. When recording in ambisonic, manipulation is straightforward and the encoded material is usable on any ambisonic system.

Higher-order ambisonic in 2D takes the form of an equatorial ring. Since humans are more sensitive to localisation within the horizontal plane, 2D higher-order ambisonics is an affective auralisation technique, especially if some elevation is produced by psychoacoustic effects or by vector-based panning with some additional lofted speakers.

3.3.1.2 Choosing a 2D Ambisonic Configuration

A 2D Ambisonic configuration was chosen for two reasons. Primarily it is more practical and portable than a 3D configuration since the eight loudspeakers can be mounted on portable

speaker stands rather than requiring a large and complex custom rig. Portability was deemed important for the later stages of testing that may have required building the Ambisonic rig in various public spaces. (A later change in method gave rise to the use of binaural reproduction for Phase 3, as justified in section 3.3.2.)

Secondly, a study by C. Guastavino reported that 2D reproduction offered a greater level of perceived 'naturalness' for outdoor soundscapes and 3D reproduction was reported to offer a greater 'naturalness' for indoor soundscapes [88]. The study simply asked participants directly to rate the 'naturalness' of various soundfield reproductions. Admittedly this lacks enough evidence to propose only 3D ought to be used for the reproduction of indoor soundscape and only 2D for the reproduction of outdoor soundscapes. However, Guastavino's comparison of perceived naturalness was deemed sufficient to justify the use of a 2D Ambisonic configuration that would, of primary concern, offer a practical and portable solution.

3.3.1.3 Physical Setup

An equatorial ring of eight channels was chosen for the ability to render third order as well as first and second order Ambisonics. Genelec 8030a active loudspeakers were arranged at a (sitting) listener's ear height, starting at 27.5° azimuth and equally spaced at 45° angles at a radius of 1.5m. (Measurements were made to the front surface of the Genelec

loudspeaker casings, laterally central and an inch below the centre of the tweeter.) The loudspeakers' LED displays were covered.

Phase One used first order Ambisonic decoding with MaxMSP. Phase Two used third order Ambisonic decoding with WigWare. An RME HDPSe MADI soundcard was used, feeding to an RME digital to analogue converter. The b-format recordings were 44.1kHz 16bit.

3.3.1.4 Calibration

In Phase One the sound pressure level was measured at MediaCity whilst simultaneously recording. A sound level meter was set to slow integration and A-weighting. The equivalent sound level was measured over ten seconds, LAeq,10s = 55.2dB. In the listening room, all 8 loudspeakers were set to the same volume setting before the ten second b-format recording was reproduced in ambisonics. LAeq,10s was measured with the same sound level meter and the playback amplitude adjusted between repeats until the LAeq,10s at the sweet spot reached 55.2dB.

The process of calibrating was streamlined in Phase Two. A b-format soundfield recording was made of diffuse 80dB(A) white noise in a reverberation room (using the same equipment settings as per field recording). (This was performed at the University of Salford, courtesy of Anugrah Sabdono Sudarsono.) Prior to the listening room tests, the eight channels were levelled one at a time with 70dB(A) of white noise and a sound level meter

at the sweet spot. After balancing, the b-format white noise was reproduced and the master volume set for a meter reading of 80dB(A) at the sweet spot.

3.3.2 Binaural

3.3.2.1 Binaural: Fundamentals

Binaural: “having or relating to two ears”. Binaural recording is a form of stereo recording that takes into account a head-related transfer function thus allowing the listener to localise sound sources. Recordings can be made using a dummy head with two external microphones, internal microphones at the centre of the model outer ears, or at the inner end of a model ear canal. Furthermore, b-format recordings may be decoded to binaural [89]. In short, this complex process involves decoding the b-format signals to virtual loudspeakers whose output is modified by a head-related transfer function (HRTF) relating to the relative postion of the virtual loudspeaker and virtual listener.

Playback is normally through in-ear or head phones. Closed headphones will offer the best insulation. It is possible to use two loudspeakers with crosstalk cancellation but the channel separation cannot reach 100% and unwanted colouration occurs [90]. Ultradirection loudspeakers are theoretically valid for binaural playback but, again, further development is required.

Headtracking can be used with headphone listening. The movement of the listener's head is monitiored and the signal outputs are modified accordingly such that the soundfield appears static relative to the physical listening environment. This is easiest performed with a simulated soundfield of virtual sound sources (within a programming environment such as MaxMSP or Visual Studio). The location of the virtual sources relative to the virtual listener determines the HRTF by which the signal is modified. Moving the relative postion of the virtual sources and virtual listener – cued by the headtracker – utilises the relavant HRTF. B-format recordings can be decoded to binaural in a similar fashion to simulated soundfields (whereby the virtual loudspeakers are the virtual sources).

3.3.2.2 Choosing Binaural

The use of binaural technology in this research stemmed from the need for Phase Three's real-time in situ reproduction. Thus isolating the effect of the 'electroacoustic ear' enables a more detailed understanding of how soundscape perception in situ and reproduction soundscape perception differ. An ambisonic configuration would have several drawbacks for an application such as Phase Three. The captured soundfield and the real-time reproduced soundfield would need to be far apart to avoid feedback. Also there would always be the background noise – something that earphones are able to reduce. Most significantly, the use of binaural provides an interesting comparison to ambisonics.

3.3.2.3 Physical Setup

The Phase Two reproduction was of a two minute clip on continual repeat. The recording was made using a Brüel & Kjær dummy head, Brüel & Kjær pre-amplifier, and Zoom 2-channel digital recorder. The dummy head was positioned on a concrete bench, with foam under its metal supports in an effort towards insulating from any ground vibration. The location of recording was only several meters from where the ambisonic recordings were previously made and thus of a similar soundscape.

The clips were played from a laptop through an M-Audio 2-channel soundcard and Sennheiser CX300ii in-ear headphones. The participants also wore a pair of 3M ear defenders. The same reproduction equipment was used for Phase Two and Three – the only difference being that in Phase Three the binaural capture was relayed live. Acoustically insulating the binaural reproduction for the participants was crucial for Phase Three and whilst the listening room binaural test of Phase Two did not require such insulation, it was important to retain continuity. (The ear defenders affect the frequency response as discussed in section 3.3.1.4.) During Phase Three, a 98Hz high-pass filter was implemented to avoid wind noise.

The combination of in-ear headphones and ear defenders was selected as the best means of blocking out the non-electroacoustic soundscape and hence isolating the binaural reproduction. Whilst open headphones tend to offer better sound quality, closed

headphones offer better isolation. Basic testing with the closed headphones available to the researcher concluded that none had sufficient enough isolation such that the non-electroacoustic soundscape would be inaudible. Furthermore, the soundscape is too varied for active noise cancelling to be a viable option.

3.3.2.4 Calibration

The earphones/ear defenders combination was calibrated at a reference of 1kHz. A calibration signal of 1kHz at 80dB was played directly into the left microphone of the Brüel & Kjær dummy head, with the outer ear removed and the correct adapter to give the calibrator a flush fit. Using the same equipment and setting as per the field recording, the calibration signal was recorded (r1). The same calibration signal of 1kHz at 80dB was played directly into the left ear of the Kevlar dummy head and, again, recorded (r2). This signal, r2, was fed back to the laptop through the M-audio soundcard and analysed with the software WigWare. A note was made of the sound level identified at 1kHz, L_{r2} (which was greater than 80dB). The r1 signal was then played back from a laptop through the M-Audio soundcard and the Sennheiser CX300ii left earphone. The earphone was placed in the intricately modelled ear canal of a Kevlar dummy head, with the 3M ear defenders over the top. The output signal, r3, was sent back to the laptop where the sound level was compared to that of the input signal, r1. For Phase Two, the binaural soundfield recordings were adjusted accordingly, with the software Audacity, by reducing the level by L_N dB, where $L_N = L_{r3} - L_{r2}$. For Phase Three the Zoom 2-channel recorder playback level reduced by L_N dB.

3.3.2.4.1 FREQUENCY RESPONSE

Both the Brüel & Kjær and the Kevlar dummy heads offer flat frequency responses (variation < 1dB). The earphones were expected to peak at 5kHz and dip towards the upper and lower ends of the audible frequency spectrum. However, the effect of the ear defenders was unknown and therefore investigated. White noise at 90dB was sent to the earphone fitted in the left ear of the Kevlar dummy head, returning a signal that was analysed using WigWare's fast Fourier transform analyser, 'FFT Analyzer'. Screenshots showing the frequency response of the earphones in the Kevlar dummy head without ear defender can be found in section 7.1.1.1 of the appendices and with ear defenders in 7.1.1.2. The effect of the ear defenders was to increase the overall levels as well as raise the level of the upper and lower ends of the audible frequency spectrum.

3.3.3 Summary of Reproduction Decisions

Ambisonic reproduction was chosen as it is a commonly used soundfield reproduction technique, and potentially to become more so, therefore it is important to understand the perceptual nature of Ambisonic reproduction. The two dimensional eight channel setup was chosen not only on the grounds of its more common usage but of its relative practicality thus better preserving the practical opportunities for – and comparison with – any potential future Ambisonic tests outside a listening room environment.

Binaural reproduction was chosen for three reasons: it is intrinsically very different a means of reproduction to ambisonics and therefore offers a potentially important contrast; binaural is, again, another common means of soundfield reproduction and is likely to become more so; and binaural reproduction provides adequate practicality to be efficiently used outdoors and in real-time thus opening the opportunity for Phase Three's investigation of the 'electroacoustic ear'.

3.4 Soundscape Selection

Although existing research compares soundscapes, there is a trend for the disproportionate representation of public squares and urban green spaces. Here, comparisons were made with a range of 'lived' soundscapes. Since none of the literature reviewed indicated a greater variation in subject perception than soundscape stimulus (or vice versa), an equal number of participants was used in the focus group as number of soundscapes used in Phase One testing. This gave totals of N=25 in situ and N=25 for ambisonic reproduction. This was deemed sufficient for a meaningful comparison of in situ and ambisonic reproduction and enough to give some insight – albeit with less confidence at N=5 – on the comparison of different soundscapes.

Phase One used five locations. Salford Precinct, a shopping centre, offered an indoor soundscape so that a comparison can be made with outdoor soundscapes in light of the

perceptual models. The Kings Arms pub garden offered an atypical soundscape (given the contrasting time of testing and time of common attendance) and thus highlighted the test conditions. The Triangle and MediaCity offered rather typical examples used in existing soundscape research, both large urban open spaces. The university campus – outside the Newton building entrance – offered a very familiar soundscape for the focus group (consisting of colleagues and hence users of the soundscapes). Essentially, the selection offered a range in terms of dynamics; temporality; familiarity; contextual congruousness; indoor-outdoor; demographics; and proximity to listening room of the later soundfield reproductions. The five locations offer a range of soundscapes that are therefore more representative of 'lived' urban soundscapes than public squares alone.

Phases Two and Three used only the MediaCity soundscape. More specifically, the weekday afternoon soundscape (12pm-3pm) at a particular spot amidst a green open space that connects offices to a waterfront and tram stop. The testing was done over several weeks therefore the weather conditions varied. However, owing to the time of day of testing the social contexts were fairly consistent in that many people taking part were working at MediaCity. For phase three, the weather had to be dry for the sake of the equipment. Furthermore there had to be very little wind in order to avoid wind noise during playback. (Unlike phase two, the recordings cannot be trimmed since it was a live stream.) Testing dates were determined by the Met Office weather forecasts, limited to weekday afternoons with a chance of precipitation less 20% and wind gusts of up to 10mph forecasted. It is worth

noting that the weather conditions were therefore, on median, more pleasant than the Phase Two in situ tests at MediaCity.

3.5 Focus Group

The focus group consisted of five people with whom the researcher was well acquainted. All five were conducting post-graduate study in or around the field of acoustics and may be considered 'experts'. Therefore the focus group was representative of the participants used in the laboratory tests of Phase Two. The use of a focus group in situ facilitates a comparison with the responses from members of the public (collected at MediaCity in Phase Two) – a comparison of 'non-users' (under test conditions) and 'users'.

The incentives were informal. The group and researcher met outside their place of work at the first soundscape location, the front of the Newton Building, University of Salford. The researcher provided transport between sites by means of a hired minibus.

3.6 Recruiting Participants

All participants were asked in person by the researcher whether they wished to participate.

All participants were:

- Asked whether they wish to and have time to participate in a piece of postgraduate research (they are never coerced, asked to help the researcher 'as a favour' etc.)
- Given a full outline of what is required of them before beginning the testing, including expected timings (approximately five minutes) and the fact that questionnaire responses are anonymous
- Given the opportunity to decline after hearing the outline of what is required
- Invited to ask questions relating to the context and nature of the research
- Provided with clean, fresh in-earphone pieces where they are required to wear them, with the researcher unwrapping and attaching them in front of the participant

3.6.1.1 Colleagues

Colleagues were approached one-to-one at work during working hours. They were specifically approached when they do not appear busy (i.e. at a computer, on the phone). They were given the option of participating at a future time/date if more convenient.

3.6.1.2 Members of the Public

Members of the public were selected from passers-by on the day of testing. They were first asked whether they had time (approximately five minutes, ten minutes maximum) to participate anonymously in a postgraduate research project for a local university.

Participants approached were specifically those who did not appear busy (i.e. walking at speed, on the phone). Where possible, participants were those who had already approached the research set-up, and expressed interest in what was taking place.

The context and nature of the research, and exactly what was required of the individual to participate, was explained. Potential participants were then given the opportunity to decline if they didn't have time or were not happy to participate for any reason. Before beginning the testing, participants were reassured that all responses would be anonymous.

3.7 Questionnaires

The questionnaire completed by the focus group in Phase One, Questionnaire A, can be found in appendix 7.2.1. The questionnaire consisted of six open questions (completed individually for all soundscapes) and an additional semantic scales section for the Kings Arms soundscape and respective reproduction. The sixth question was optional and rarely completed and hence omitted in the second questionnaire, Questionnaire B. The semantic scales section was not included for the first four soundscapes so as not to influence the language used in response to the open questions. Raimbault noted that during a study where participants responded to open questions following completion of a semantic differential, the responses almost invariably used at least one word from the differential

[91]. It was particularly important to avoid such an effect with this methodology given the linguistic analysis by which the open responses are analysed.

Phase Two and Three used a modified questionnaire, Questionnaire B, consisting of five open questions (omitting the sixth) and 18 semantic scales. (Questionnaire B is included in appendix 7.2.1.) The 18 semantic scales were those used by Kang and Davies in order to enable comparison [68][59].

3.7.1 Semantic Differentials

In Questionnaire A of Phase One the semantic differential enquiry took place after completion of the open questions at the final soundscape, the beer garden at the Kings Arms pub, Salford. Eight semantic differentials were used:

unpleasant – pleasant
uneventful – eventful
social – unsocial
boring – chaotic
happy – sad
noisy – quiet
exciting – tranquil
unpleasant (environment) – pleasant (environment)

The low number of scales – relative to some previous studies – was intended to ensure the enquiry was not time consuming and therefore could be performed on a mass scale. (That is if the accompanying open responses support the semantic scales responses thus validating the semantic scale enquiry as meaningful.) The polarity of the scales was varied

to encourage responses that focus on each individual scale rather than contribute to the development of left-right polarity for any overarching and simplified expression of, say, good-bad.

Questionnaire B of Phases Two and Three posed 18 semantic scales previously used by Kang [59] and later by Davies [69]. A comparison with Davies et al.'s results could reveal the influence of using the same soundscapes (rather than different locations' soundfields for ambisonic reproduction than in situ) on the reproducibility of the semantic scales' soundscape dimensions. The 18 semantic scales are presented here:

Comfortable - Uncomfortable
Noisy - Quiet
Unpleasant - Pleasant
Interesting - Boring
Natural - Artificial
Like - Dislike
Gentle - Harsh
Hard - Soft
Fast - Slow
Sharp - Flat
Directional - Universal
Varied - Simple
Reverberant - Anechoic
Far - Near
Social - Unsocial
Meaningful - Meaningless
Calming - Agitating
Smooth - Rough

3.7.2 Semantic Categories

The open responses were initially analysed by the categorisation method proposed by Guastavino et al.[70]: 1) sound event related and background noise related; 2) object-centred (i.e. sound entity-centred), subject-centred, and source identification (i.e. an object or objects). Comparison with this existing study isolates the influence of subjects' expertise and agency since Guastavino used the public in situ and students/colleagues in the laboratory.

A second, and new, method was developed using different semantic categories and a statistical analysis to observe the inter-relation of the sematic category occurrences. Simply summating or averaging the category occurrences fails to take into account the individual cases and a lot of information is lost. Therefore following the categorisation process, a principal component analysis was performed. The category occurrences form the variables and all individual question responses form the observations. The resultant components represent the co-variance of the semantic categories, their inter-relation. With this analysis in mind, a developed psycholinguistic analysis was designed with the aim of better representing subjects' perceptions of the sound-ecology relationship (or divide).

The categories themselves are a priori and open to the researcher to choose. Raimbault analysed responses according to the three categories – activities, space, and time patterns – to compare a set of location descriptions and a set of soundscape descriptions [67]. It was

found that spatial attributes mainly characterised locations whilst activities characterised soundscapes. This distinction between (holistic) environment and soundscape as space and activity led to the current study's introduction of the categories 'object' and 'action'. An 'object' category entry was a reference to an object - a noun. An 'action' category entry was a reference to an activity or motion – a verb. 'Activity' entries included silent and sounding actions. "A choir is singing" would have one 'object' entry and one 'action'.

The category 'source identification' was deemed to bridge sound and object/motion. This category fails to observe the differing perceptions of a single phenomenon (see section 2.6) [68]. With this in mind – and the principal component analysis that draws inter-relations – 'sound identification' was replaced by a sound-specific category, 'sound object'. Sound object references are references to the sound itself, the sound as an entity. This would include a general reference such as "a sound" as well as more descriptive references such as "a hum" or "screeching".

The 'sound object' category is only for direct references to a sound or words that describe the sound itself. "The sound of the trees" would have 'sound' as a sound object entry and "a rustling nearby" would have 'rustling' as a sound object entry. The latter example identifies another issue addressed in the categorisation: 'rustling' is an action. Since 'rustling' is rarely used in a literal reference to anything other than sound creation, it is categorised as a 'sound object'. Onomatopoeic verbs are identified as sound object entries.

The ‘subject’ category was retained to provide a dichotomy with ‘object’ and as a means of identifying how the subjects describe themselves as affected by the soundscape. The phrase “I like the sound of the choir” would only have one ‘subject’ entry, ‘I’. However, an effect of the environment on the earwitness was entered as a ‘subject’ entry. “I find the choir very relaxing” would have two ‘subject’ entries, ‘I’ and ‘relaxing’.

A ‘descriptive’ category was introduced with the aim of observing – from the principal component outputs – which of the other four categories was described in more depth. Descriptive entries were descriptions of anything: objects, feelings, verbs, timbres etcetera.

The five categories – subject, object, descriptive, action, and sound object – were used for the analysis of Phase One results. For both in situ and reproduction listening, the ‘descriptive’ occurrences formed their own component. It did not offer a measure of the relative significance of the other four semantics. Instead it merely represented a measure of how much participants wrote and was therefore not included in the developed psycholinguistic analysis.

3.7.3 Potential Drawbacks

The semantic differentials were poorly presented in the questionnaires. A ten point scale ranging from 1 to 10 was used (rather than, say, a nine point scale ranging from -5 to 5).

This goes against best practice for two reasons: the even number of points offers no central value option; and the incremental scaling may give the misleading impression of positive/negative scaling. However, it is deemed safe to assume the data collected was not significantly compromised. Ten is a large number of points relative to comparable existing research where a five point scale is common. The 5^{th} and 6^{th} points of the ten point scale used only deviated 5.6% from the 5.5 mean. Whilst if a subject responding to a 5 point scale (i.e. -2 to 2) deemed a -0.5 rating to be accurate of their evaluation, choosing 0 or -1 would indicate a 10% 'error'. Furthermore, this comparison is based on the easily negated assumption that such variation is significant in relation to subjective variation – an unlikely scenario.

3.8 Design Summary

This research employed a mixed model approach grounded in the pragmatist research paradigm. The methodology was formulated from the identification of seven variables responsible for the differing perception between in situ and reproduction listening: soundscape; space; time; auralisation; place; and memory (understood here as motivation and expertise).

The experimentation accessed these variables with several soundscape perception scenarios: electroacoustic reproduction compared to 'real'; ambisonic compared to binaural; binaural live in situ compared to binaural later elsewhere; different soundscapes; and non-users compared to users.

Qualitative and quantitative data was collected in the form of open question responses and semantic differential responses. Three empirical methods of analysis were employed, two existing and one novel:

1. The extraction of 'soundscape dimensions' from the principal component analysis of 18 semantic differentials;
2. The semantic categorisation of words and phrases from the open responses into 'subject', 'object' and 'source identification'; and
3. The developed psycholinguistic analysis's extraction of 'cognitive dimensions' from the principal component analysis of the occurrences of the four semantic categories 'subject', object', 'action' and 'sound object'.

Phase One compared different soundscapes and the focus group subjects' differing perceptions of in situ and reproduction soundscapes. Phase Two compared the Ambisonic and binaural auralisation techniques and gained 'users'' in situ perceptions to contrast the focus group 'non-users'. Phase Three explored the isolation of 'the electroacoustic ear' – in place, in time, but out of space – with real-time binaural in situ listening.

CHAPTER FOUR

4 DISCUSSION AND RESULTS

The discussion and results chapter is structured by the three phases of experimentation. Each of the three sub-sections includes the within-phase analyses. The findings are later consolidated thematically in the fifth chapter.

Phase One analysed the focus groups' responses at five in situ soundscapes and their ambisonic reproductions in a listening room. The Kings Arms' semantic differential responses were compared and semantic differentials' soundscape dimensions were procured. An existing method of 'ecological validity' assessment (using three semantic categories) was assimilated with the focus groups' open responses across all soundscapes. This was repeated with five semantic categories. The developed psycholinguistic analysis was performed on all soundscapes' responses individually. The developed psycholinguistic analysis was then re-appropriated as a method of validity assessment of soundfield reproduction.

Phase Two analysed public's responses at the MediaCity soundscape in situ and participants' responses in the listening room for an ambisonic reproduction and a binaural reproduction. The semantic differential responses were compared and semantic differentials' soundscape dimensions were produced. The open responses were categorised into four semantic categories and – prior to the principal component analysis of the developed psycholinguistic analyses – a Tukey (HSD) analysis was performed in conjunction with a one-way ANOVA.

Phase Three analysed the public's responses from the real-time binaural reproduction listening test in situ at MediaCity. Semantic differentials' soundscape dimensions were analysed and the developed psycholinguistic analysis was performed and discussed.

4.1 PHASE ONE

The focus group's questionnaire responses were analysed from the five in situ soundscape listening tests and the five ambisonic reproduction listening tests. The section commences with analysis of the Semantic Differentials followed by the **Semantic Categorisation of Open Responses**. The latter includes **A comparison with Existing Research**: Agency and Knowledge; **Five Semantic Categories**: People and Places; and the **Developed Psycholinguistic Analysis**.

4.1.1 Semantic Differentials

4.1.1.1 Soundscape Evaluations

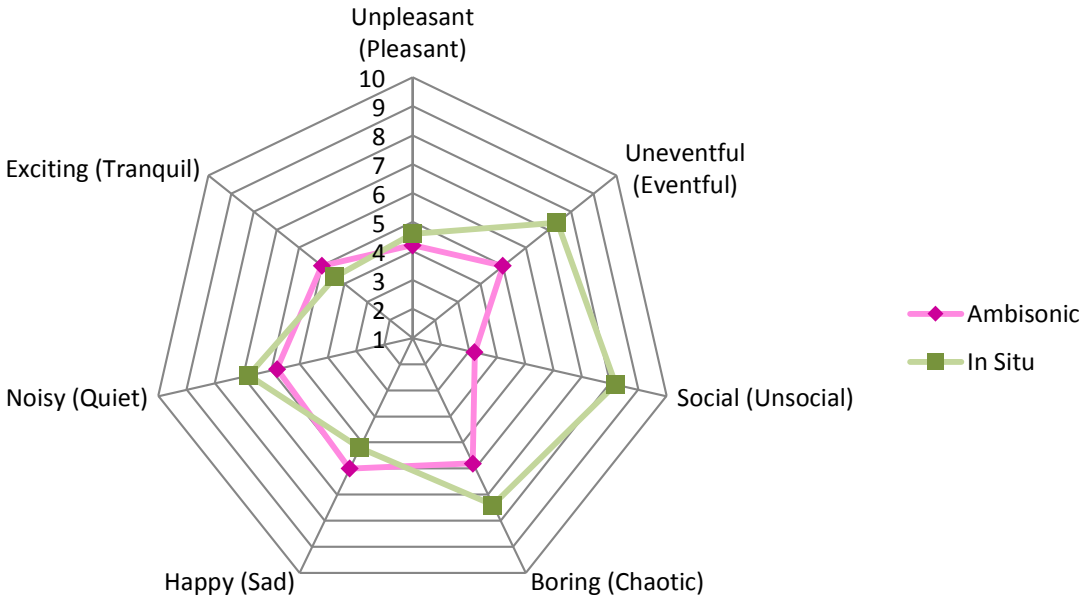


Figure 7 – The focus group's semantic differential responses for the Kings Arms in situ and ambisonic reproduction soundscape listening tests (*N*=5; *N*=5)

The semantic differential enquiry, of the King's Arms pub garden soundscape, showed that most variation between the in situ soundscape and the reproduced soundfield was accounted for by the social descriptor. This was owed to the recording having been made without the participant group present at the site and that the participants are listening to the laboratory reproductions alone in the listening room, with the researcher being the only other person present opposed to four others with whom a couple of hours had been shared.

It is interesting to note that the Ambisonic reproduction soundscape was deemed more eventful and chaotic whilst also quieter. The apparent eventfulness and chaos were potentially enhanced by contrast with the uneventful and dull non-sonic environment of the listening room. These may be deemed descriptors that gesture more significantly towards the other dimensions of the environment, beyond the soundscape, and that one is not accustomed to attaching exclusively to sound. Similarly, the moderate 'isolation' of the soundfield, thus 'alternative perspective', may have facilitated the unveiling of more detail and texture that, in turn, increased the perceived eventfulness and chaos. Furthermore, the in situ soundscape of the Kings Arms was perceived as so uneventful and boring because it did not meet the expectations of the physical environment (owing to the time of day) and therefore contrasted with the participants' hypothetical motivations for being there.

4.1.1.2 Semantic Differentials' Soundscape Dimensions

A principal component analysis produced three components with eigenvalues greater than 1.0 for both in situ soundscape perception and Ambisonic reproduction. Given the semantic differentials could only be included in the questionnaire at the final soundscape at The King's Arms pub garden, five responses were collected. A Kaiser-Mayer-Olkin index (KMO) could not be calculated because the number of observation was lower than the number of variables. Therefore no strong conclusions should be drawn from the results which are included here for completeness and to provide a sense of what this quick, and therefore practical, soundscape enquiry can offer.

	In situ: focus group			Ambisonic: focus group		
	Component			Component		
	1	2	3	1	2	3
% of variance	34.4	33.9	25.3	34.1	31.6	25.7
Unpleasant (Pleasant)	**-0.856**	-0.351	-0.378	0.123	**0.836**	-0.482
Uneventful (Eventful)		-0.422	**0.873**	**-0.864**	0.372	-0.241
Social (Unsocial)		**0.986**		0.288	**-0.889**	-0.347
Boring (Chaotic)	-0.306	0.340	**0.873**	-0.584	**0.756**	-0.292
Happy (Sad)	**0.876**	0.117	-0.122	0.449		**0.768**
Noisy (Quiet)	**0.885**	-0.116	-0.283		-0.113	**0.844**
Exciting (Tranquil)	0.129	**0.979**		**0.999**		

Table 1 – Principal component analysis of semantic scales responses (Kings Arms; *N* = 5; Varimax rotation and Kaiser normalisation)

The in situ responses produced dimensions of appraisal/dynamics, society/activity and eventfulness. The Ambisonic responses produced dimensions of eventfulness, society/pleasantness/activity and dynamics appraisal. Boring-chaotic and exciting-tranquil are both scales of activity – the former negative and the latter positive – and correlation between exciting-tranquil and eventfulness was observable in the second Ambisonic dimension.

The first component in situ represents the subjects' pathetical responses. In contrast, the reproduction's first component represents eventfulness and vibrancy whilst appraisals are relegated to components two and three. This could suggest that the laboratory soundscape interpretations are principally objective. However, this effect may be owed to the nature of the location – a pleasant and familiar pub garden – that influenced the interpretation of an

otherwise unpleasant and unfamiliar soundscape in situ. However there is still the indication that the reproduction shifted objective descriptors ahead of appraisal descriptors.

This theory may extend to the listening of music. The live performance in a concert hall or other venue stimulates a holistic perception beyond the sound itself. The music is shaped by the musicians' identity and presence, the place of performance, and the audience's motivations for being there (for example the acoustics or 'a night out'). However if one wishes to examine the music, an alternative perception might be stimulated listening in a room through loudspeakers. These are differing yet complementary forms of listening.

4.1.2 Semantic Categorisation of Open Responses

4.1.2.1 A comparison with Existing Research: Agency and Knowledge

The method of semantic categorisation, proposed by Guastavino et al. [70], was performed on the Phase One open responses for in situ and Ambisonic reproduction listening. The results are shown in Figure 8 for responses regarding sound events and Figure 9 for responses regarding background noise. In situ descriptions referring to the source generating sound represented 62% of sound event descriptions and 27% of background noise descriptions – comparable to the 76% and 15% identified in the existing study by Guastavino et al (as presented in section 2.7).

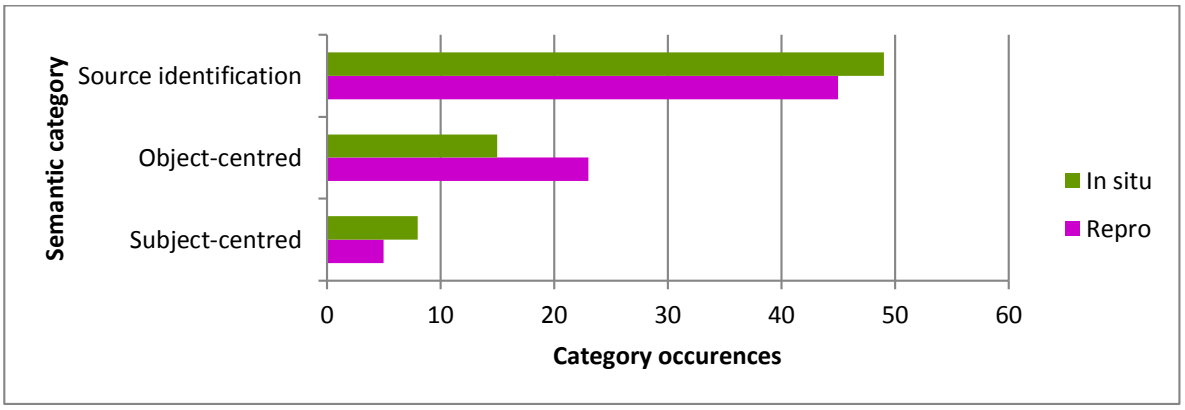


Figure 8 – Open responses referring *sound events* into three semantic categories (as per Guastavino et al)

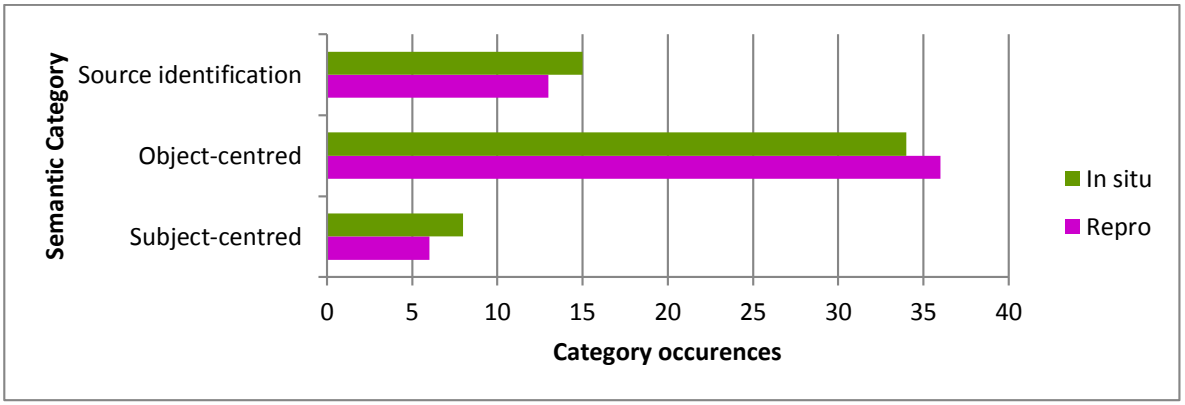


Figure 9 – Open responses referring *background noise* into three semantic categories (as per Guastavino et al)

The distribution of subject-, object-, and source-centred responses will vary due to the construction of the questionnaire. It is the in situ to reproduction variations that are of interest. A chi-squared test of independence was performed for each of the six data pairs (see Table 2). No significant independence (or correlation) was found.

The current results do present an observably stronger correlation between the in situ and ambisonic responses' source identification of sound events (see also Figure 2 and Figure 3

of section 2.7)[11]. This was attributed to the differing expertise of the in situ subjects. The current study used the same (expert) subjects in situ as for the reproduction. Whereas the existing study used members of the public in situ who were most likely not experts and therefore more likely to employ holistic listening and make general descriptions rather than focus on the source of individual sounds. Another potential contributor to general variation between the two studies is the differing agency of the in situ subjects. (The reproduction subjects will have had a similar sense of motivation in both cases.)

The a priori categorisation of sound events and background noise does not adhere to the psycholinguistic approach. Therefore responses referring to sound events and background noise (in the case of the current study, the responses to questions one and two) were categorised collectively to adhere comprehensively to the psycholinguistic approach. The current and existing results are compared in Figure 10 and Figure 11.

[11] Given the lack of data for the existing study a statistical comparison cannot be made.

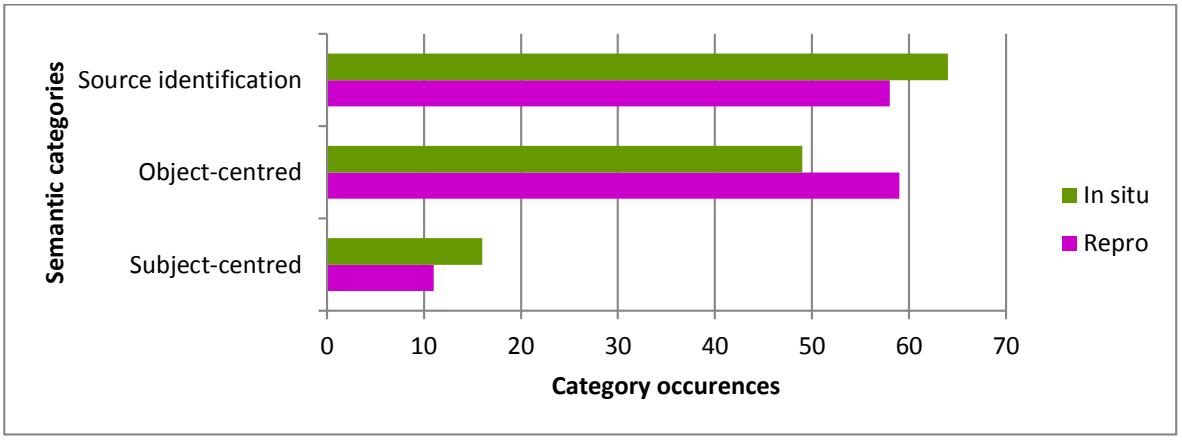


Figure 10 - Open responses referring *sound events* and *background noise* into three semantic categories (as per Guastavino et al)

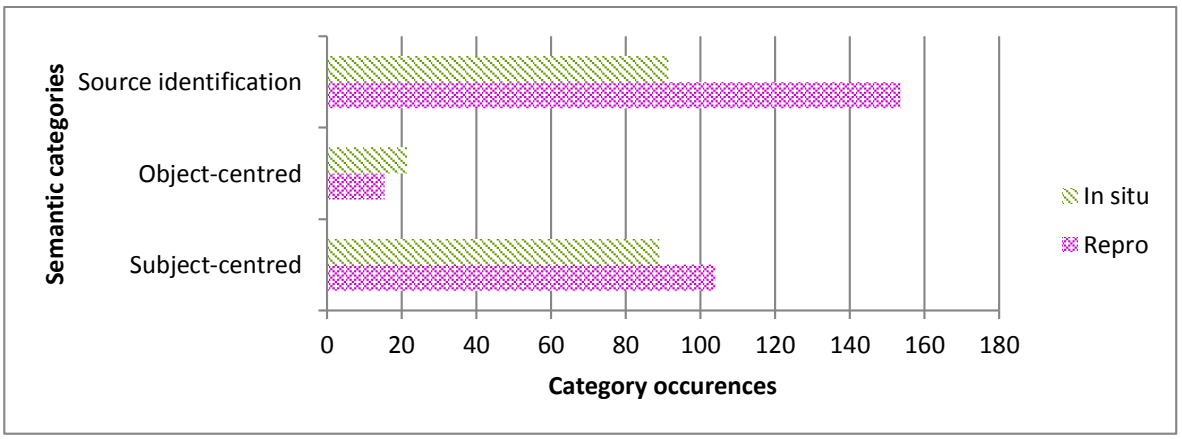


Figure 11 – Guastavino et al's open responses [66] **referring *sound events* and *background noise* into the three semantic categories**

	Sound events			Background noise			Combined		
	source-	subject-	object-	source-	subject-	object-	source-	subject-	object-
Chi-square (Observed value)	15.000	3.333	15.000	10.000	6.250	15.000	60.000	6.778	60.000
Chi-square (Critical value)	21.026	9.488	21.026	15.507	12.592	21.026	74.468	15.507	65.171
p-value	0.241	0.504	0.241	0.265	0.396	0.241	0.333	0.561	0.115

Table 2 – Chi-squared test of independence between in situ and ambisonic reproduction responses' three semantic category counts

The current results showed no significant variation from in situ to ambisonic reproduction. However the ambisonic reproductions do appear to have stimulated more reference to object-centred descriptions. This could be a result of the ecology – shifting more attention to the sound itself and less to the sound in relation to the non-sonic environment.

Overall there appears to be a good correlation between the responses to the perception of the in situ and listening room tests (Figure 8; Figure 9; Figure 10). The number of occurrences in each category, for in situ and reproduction, was compared with the statistical X^2 test. No significant difference was found. Therefore Guastavino's method of ecological validity assessment suggests there is no significant difference in the cognitive process of perceiving in situ and ambisonic reproduction soundscapes. However, this does not necessarily mean that no difference exists. The fact that there is interaction between the senses implies that a reproduction will – at some level – be perceived differently to a soundscape in situ. However, the point at which a reproduction technique is classified as ecologically valid – or invalid – is model-dependant.

4.1.2.1.1 SUMMARY

The method of semantic categorisation, proposed by Guastavino et al. [70], was performed on the Phase One open responses for in situ and ambisonic reproduction listening. The 'subject' and 'object' references were consistent for in situ and ambisonic reproduction listening – as per Guastavino et al.'s previous study. The source identification was also

consistent for in situ and ambisonic reproduction listening – whilst it was less so for the previous study. This was attributed to the current study's alignment of subject expertise and motivation across in situ and reproduction tests.

In conclusion this particular method did not reveal any significant difference between the cognitive processes of perceiving in situ soundscapes and perceiving ambisonic soundfield reproductions.

4.1.2.2 Five Semantic Categories: People and Places

Different soundscapes are inevitably evaluated differently. But are different soundscapes listened to in different ways? A comparison of the focus groups in situ responses at the five locations of Phase One indeed showed variation.

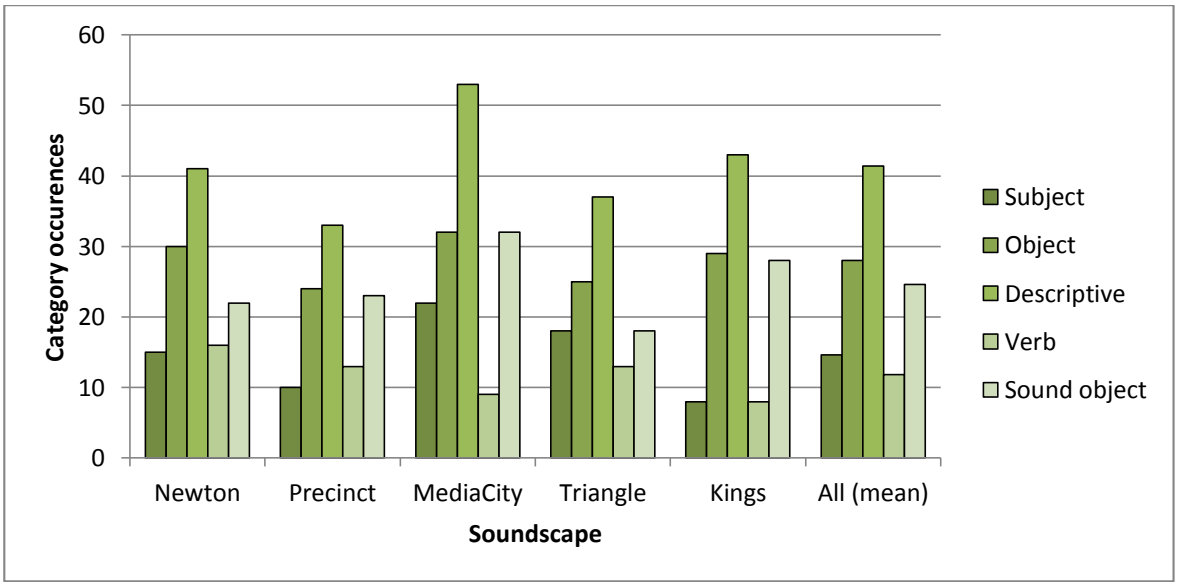


Figure 12 – Responses' semantic category occurrences for the five soundscapes in situ

Figure 12 shows the total semantic category counts from all five participants' responses to the first five questions. Responses to the sixth, and optional, question were excluded due to too few completions.

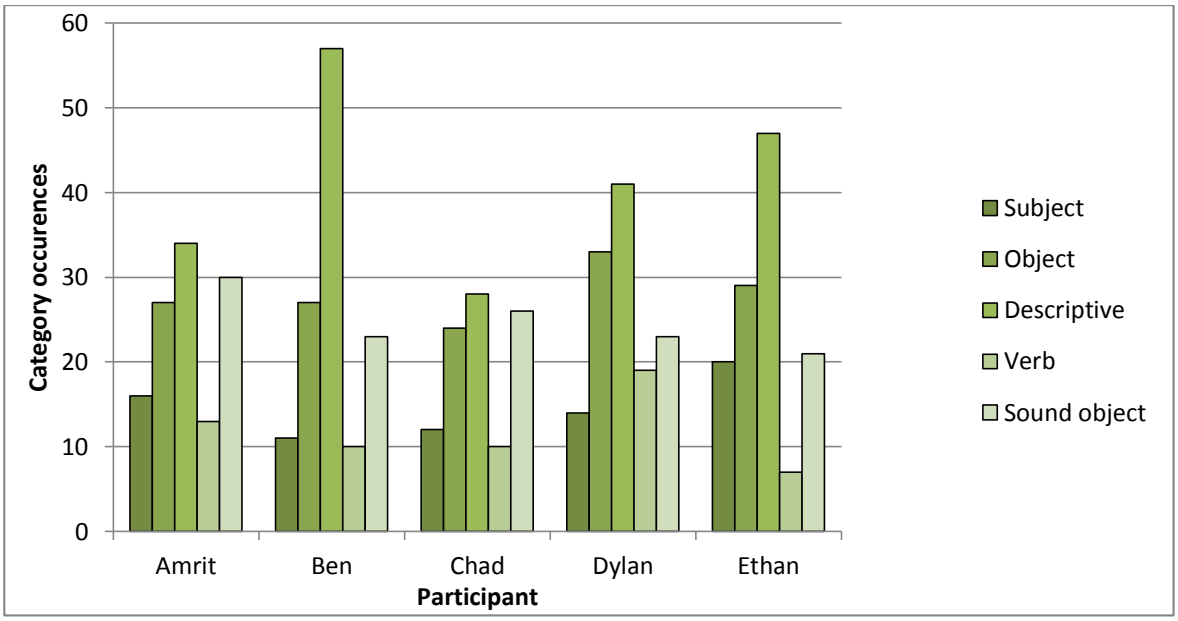


Figure 13 – Responses' semantic category occurrences for the five participants in situ

Figure 13 shows the total semantic counts from all five locations and five questions. (The names have been fictionalised.) Both Figure 12 and Figure 13 show the relative frequency of the five semantics with 'descriptive' being the most frequent category, followed by 'object', 'sound object', 'subject' and 'verb'.

A Wilks' G-square test was performed to examine the relation between the participants and the places on the occurrence of each semantic category[12]. Regarding 'subject' references, there was a significant link between participant and place, G^2 (2, N = 25) = 26.30, p = 0.036. [Chi-square test: X^2 (2, N = 25) = 26.30, p = 0.082.] From this one can infer that the occurrence of 'subject' references depends upon a listener-place relationship. Regarding 'descriptive' references, there was no significant link between participant and place, G^2 (2, N = 25) = 26.30, p = 0.950. [Chi-square test: X^2 (2, N = 25) = 26.30, p = 0.963.] From this one can infer that the occurrence of descriptive words and clauses is independent of a listener-place relationship.

It was hypothesised that subject references may relate to the intimacy of the participant/location place relationship. The participants' responses to how often they frequent the test locations was analysed and compared to 'subject' references across all five locations. The multiple choice responses – 'daily', 'weekly', 'monthly', 'rarely', and 'never' – were converted to estimates of visits per year – 188, 40, 12, 2, and 0 – for each participant and location. A Wilks' G-square test was performed to examine the relation between numbers of 'subject' references and estimated numbers of visits per year. The relationship between these variables was significant, G^2 (2, N = 25) = 9.49, p < 0.0001. [Chi-

[12] An alpha level of 0.05 was used for all statistical tests.

square test: X^2 (2, N = 25) = 9.49, p < 0.0001.] The more familiar the subject with the place the more 'subject' related the response.

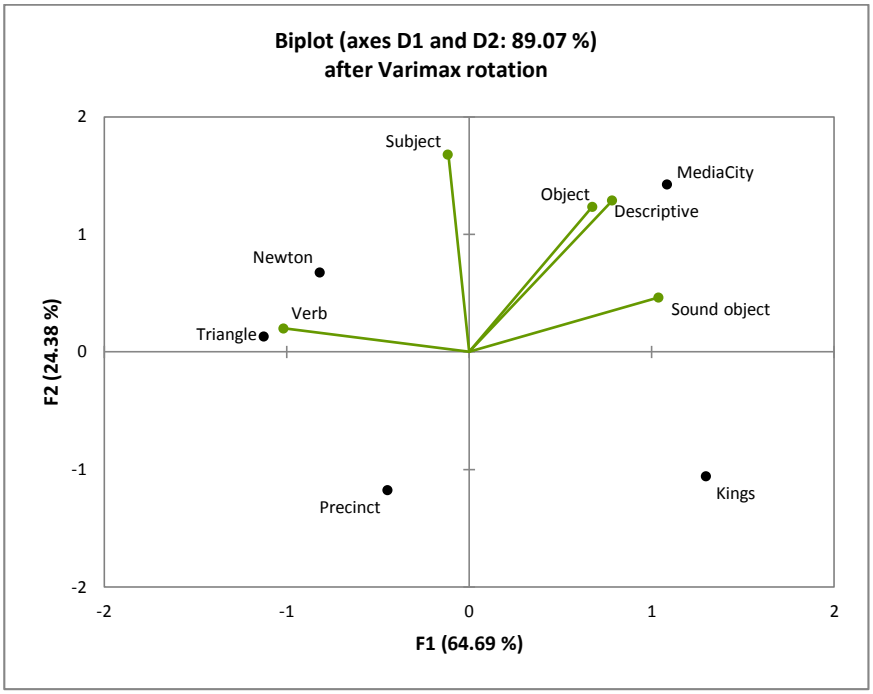


Figure 14 – Semantic categories' and locations' contributions to the first two components (Varimax rotation and Kaiser normalisation)

Figure 14 offers a visual representation of the semantic categories' contributions to the first two factors following a principal component analysis. The category counts were summed for each location (giving n=5) and a Varimax rotation and Kaiser normalisation were performed on the first two factors. An inverse relationship between 'verb' and 'sound object' can be seen.

The observation points represent the centroids of each soundscape's contributions. The locations' contributions on the first factor represent the (visual) eventfulness with

Manchester's Triangle and the Newton building location appearing the most eventful and the Kings Arms pub and MediaCity appearing the least eventful. (The researcher's observations during the listening tests reported a similar trend, albeit with the Triangle and Precinct as the most eventful.)

4.1.2.2.1 SUMMARY

There was a significant link between participant and place manifest in the occurrence of 'subject' references in the open responses. Therefore the person-place relationship played a role in the written evaluation of the soundscapes. There was a significant independence between participant and place manifest in the occurrence of 'descriptive' references in the open responses. The relation between numbers of 'subject' references and the estimated numbers of visits per year to each of the five locations was significant, thus the more familiar the subject is with the soundscape the more 'subject' related the response.

4.1.2.3 Developed Psycholinguistic Analysis

All locations were analysed collectively and then individually in order to compare the differing cognitive dimensions for each soundscape. The 'descriptive' semantic category was omitted in the developed psycholinguistic analyses – here and in Phases Two and Three. The use of four semantic categories was found to be sufficient to reveal contrasting cognitive process of soundscape perception since the 'descriptive' semantic category varied

independently of other categories and did not show any variation between in situ and Ambisonic listening.

Two empirical indices were used to assess significance: Bartlett's test of sphericity and the Kaiser Meyer Olkin measure of sampling adequacy. The Triangle Ambisonic data showed no significance: Bartlett's $X^2(25)=4.3$, $p=0.632$; and KMO=0.432. The MediaCity ambisonic data showed only low significance: Bartlett's: $X^2(25)=10.4$, $p=0.108$ (and KMO=0.630). The Newton in situ data showed only low significance: Bartlett's $X^2(25)=11.6$, $p=0.070$; and KMO=0.474. The nine other datasets showed significant factors: $p<0.05$; KMO>0.5.

	All soundscapes (In situ)			All soundscapes (Ambi)		
	Component			Component		
	1	2	3	1	2	3
% of variance	41.0	26.5	25.0	37.1	25.2	25.0
Subject			**0.997**			**1.000**
Object	**0.859**	0.319		**0.813**	0.242	
Verb	0.145	**0.976**		**0.878**	0.101	
Sound object	**0.938**			0.228	**0.970**	

Table 3 – Principal component analysis of in situ responses' (N=125; Bartlett's: $X^2=104.9$, $p<0.0001$; KMO=0.511) and ambisonic reproduction responses' (N=125; Bartlett's: $X^2=60.5$, $p<0.0001$; KMO=0.644) semantic category counts (Varimax rotation and Kaiser normalisation)

	Newton (In situ)			Newton (Ambi)		
	Component			Component		
	1	2	3	1	2	3
% of variance	37.7	25.2	27.5	39.8	27.5	25.7
Subject			**0.971**	-0.102	**0.963**	0.165
Object	**0.809**		-0.372	0.270	0.169	**0.942**
Verb	**0.921**		0.110	**0.787**	-0.373	0.289
Sound object		**0.995**		**0.943**		0.172

Table 4 - Principal component analysis of in situ responses (N=25; Bartlett's: X^2=11.6, p=0.070; KMO=0.474) and ambisonic reproduction responses' (N=25; Bartlett's: X^2=25.4, p=0.0003; KMO=0.543) semantic category counts (Varimax rotation and Kaiser normalisation)

	Precinct (In situ)			Precinct (Ambi)		
	Component			Component		
	1	2	3	1	2	3
% of variance	38.3	25.1	25.8	45.4	25.4	25.1
Subject		**0.994**			**0.986**	0.137
Object	**0.821**		0.291	**0.949**		0.143
Verb	0.195		**0.968**	**0.936**	-0.150	0.131
Sound object	**0.903**			0.179	0.142	**0.974**

Table 5 - Principal component analysis of in situ responses (N=25; Bartlett's: X^2=12.4, p=0.054; KMO=0.625) and ambisonic reproduction responses' (N=25; Bartlett's: X^2=30.9, p<0.0001; KMO=0.526) semantic category counts (Varimax rotation and Kaiser normalisation)

	MediaCity (In situ)			MediaCity (Ambi)		
	Component			Component		
	1	2	3	1	2	3
% of variance	44.6	25.2	26.6	34.5	25.4	27.5
Subject		**0.988**	-0.152		**0.991**	
Object	**0.893**		0.350	**0.697**	-0.167	0.427
Verb	0.219	-0.174	**0.956**	**0.925**		
Sound object	**0.969**			0.193		**0.955**

Table 6 - Principal component analysis of in situ responses (N=25; Bartlett's: X^2=34.8, p<0.0001; KMO=0.512) and ambisonic reproduction responses' (N=25; Bartlett's: X^2=10.4, p=0.108; KMO=0.630) semantic category counts (Varimax rotation and Kaiser normalisation)

	Triangle (In situ)			Triangle (Ambi)		
	Component			Component		
	1	2	3	1	2	3
% of variance	42.1	25.1	25.4	30.3	29.4	26.0
Subject		**0.990**	-0.125	0.121	**0.919**	
Object	**0.890**		0.225	**0.902**	0.174	
Verb	0.170	-0.132	**0.974**			**0.988**
Sound object	**0.927**			**0.619**	-0.550	0.244

Table 7 - Principal component analysis of in situ responses (N=25; Bartlett's: X^2=18.1, p=0.006; KMO=0.544) and ambisonic reproduction responses' (N=25; Bartlett's: X^2=4.3, p=0.632; KMO=0.432) semantic category counts (Varimax rotation and Kaiser normalisation)

	Kings (In situ)			Kings (Ambi)		
	Component			Component		
	1	2	3	1	2	3
% of variance	41.5	26.1	25.7	39.3	26.8	28.0
Subject	0.244		**0.968**		**0.971**	0.173
Object	**0.888**	0.261	0.127	**0.766**	-0.264	0.467
Verb	0.110	**0.988**		0.222	0.238	**0.928**
Sound object	**0.893**		0.272	**0.966**		0.101

Table 8 - Principal component analysis of in situ responses (N=25; Bartlett's: X^2=22.5, p=0.001; KMO=0.585) and ambisonic reproduction responses' (N=25; Bartlett's: X^2=26.8, p<0.0002; KMO=0.493) semantic category counts (Varimax rotation and Kaiser normalisation)

The sound object and object co-variance indicate sounds were described not only as sounds (sound entities) but very much in relation to the physical (visual) environment. The object and verb co-variance – with sound object independent – indicates sounds were described as sound entities but also, independently, in terms of things in motion. (The sound object and verb co-variance – from the ambisonic reproduction of the soundfield outside the Newton Building – indicates sounds were described as sound entities, but consistently in

relation to motion.) The reproducibility – from in situ to reproduction – of these semantic dimensions is the proposed assessment of external validity.

This measure of ecological validity can be observed from the variable-component loading in bold on the principal component analysis tables. (The variable-component loadings in **bold** indicate the component to which the variable has the greatest loading. The variable-component loadings underlined indicate the variable from which the component has the greatest loading.) The developed psycholinguistic assessment of all soundscapes collectively deemed the ambisonic reproduction ecologically invalid.

Assessing the soundscapes individually, only the ambisonic reproductions of the Triangle soundfield and the Kings soundfield held ecological validity. It is worth noting that the Triangle and, in particular, the Kings soundscapes were the most unusual of the five soundscape in the fact that – given the tests were conducted on a weekday afternoon – they were far less populated (by people) than one would encounter at the typical visiting hours of weekends and evenings. Therefore the current conclusion is that ambisonic reproduction is generally ecologically invalid but ecological validity is also to some extent soundscape-specific.

The analysis of Phase One's in situ open responses at MediaCity (see Table 6) will enable a comparison with Phase Two's in situ open responses (see section 5.2.1.1 of the

Consolidation chapter). The key difference between the two tests is that Phase One's earwitnesses ('non-users') were recruited participants driven to the soundscape's location whilst Phase Two's participants were active users of the soundscape's place ('users'). The second difference is that four of the five participants (comprising the focus group) of Phase One were studying in the field of acoustics at the time of testing, including one participant in soundscape-related research. The focus group could therefore be considered 'expert listeners' as well as 'non-users'.

4.1.3 Phase One Summary

The Kings Arms ambisonic reproduction was evaluated as less social – due to the lack of the accompanying focus group – and more eventful. The semantic differential analysis showed that the ambisonic reproduction was evaluated less in terms of appraisal and more in terms of objective descriptors.

The assimilation of an existing method of ecological validity assessment with the focus group indicated 'ecological validity' with observably greater significance. Therefore subject memory – agency and expertise – has a significant influence on the external validity of soundfield reproduction.

The different soundscapes were evaluated in different ways. ‘subject’ references were significantly linked to person-place relationship as well as estimated number of visits annually. The developed psycholinguistic analysis showed differing variations in the perception of in situ and ambisonic soundfields across the five soundfields. The developed psycholinguistic analysis was introduced as an assessment of ecological validity. Only the Triangle and Kings Arms soundscapes held ecological validity in the reproductions. A collective analysis indicated ambisonic soundfield reproduction did not hold external validity.

4.2 PHASE TWO

Phase Two compared in situ, ambisonic and binaural listening. The MediaCity questionnaire responses are analysed. The **Semantic Differentials** are analysed first, followed by the **Semantic Categorisation of Open Responses.** Each section includes the in situ soundscape listening tests with members of the public followed by the ambisonic and binaural reproduction laboratory listening tests with colleague/student participants.

4.2.1 Semantic Differentials

4.2.1.1 Soundscape Evaluations: MediaCity

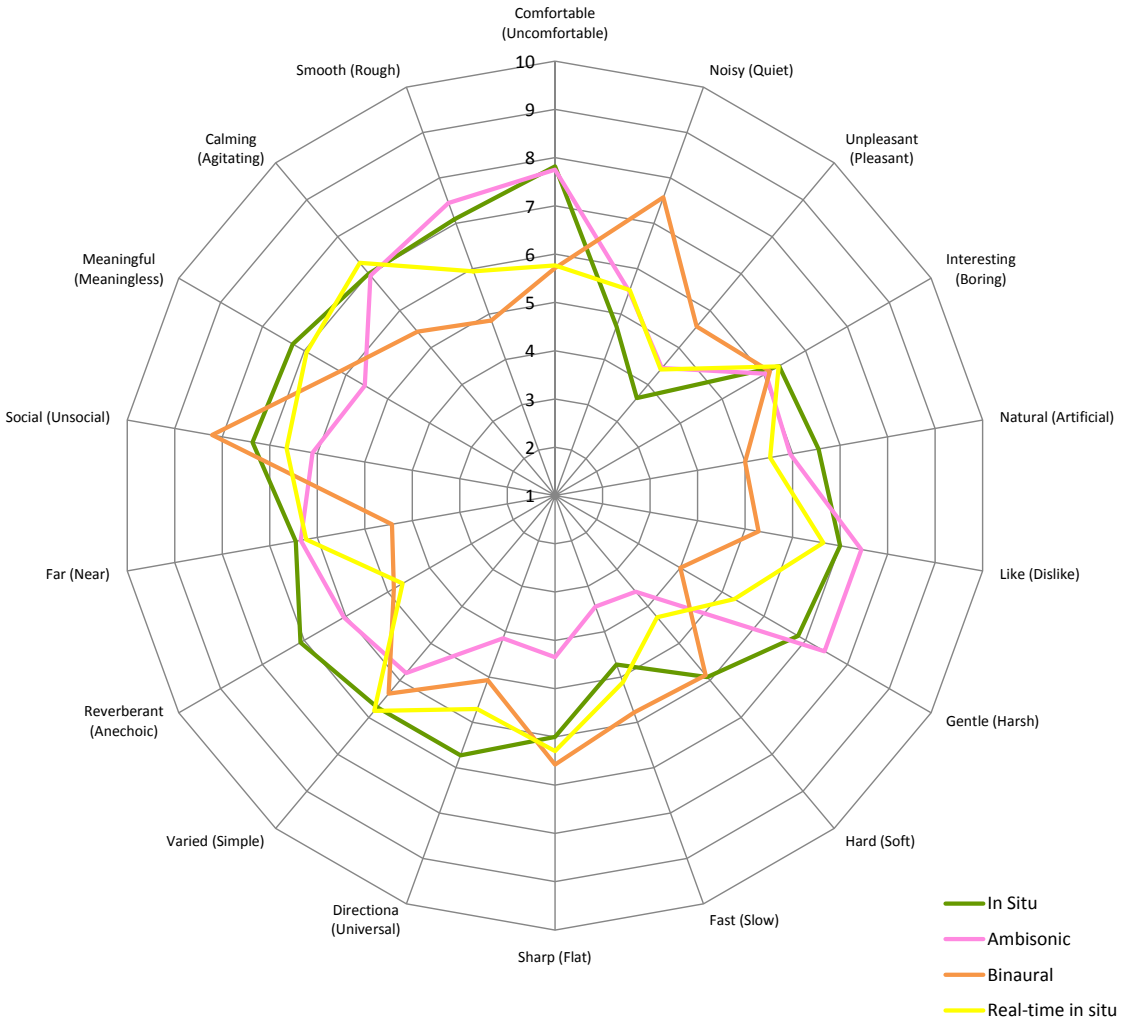


Figure 15 - Semantic differentials responses (means) from Phases Two and Three

The in situ and binaural responses were compared. The best continuities were with hard-soft, interesting-boring, varied-simple, and sharp-flat. The largest discrepancies were with noisy-quiet, gentle-harsh, reverberant-anechoic, and smooth-rough.

The in situ and ambisonic responses were compared. The best continuities were with calming-agitating; comfortable-uncomfortable; smooth-rough; and far-near. The largest discrepancies were with directional-universal; hard-soft; meaningful-meaningless; and sharp-flat.

The ambisonic and binaural were compared. The largest discrepancies manifested in the ambisonic reproduction soundscape being evaluated as smoother, slower, softer and flatter.

Smooth-rough was reproduced well with ambisonics but not with binaural. Sharp-flat was reproduced well with binaural but not with ambisonics. Phase One's social-unsocial discrepancy was not reproduced, most likely due to Phase Two's subjects not being in a group, but participating individually.

4.2.1.2 Semantic Differentials' Soundscape Dimensions: In Situ

The semantic scales responses for in situ listening at MediaCity are analysed and compared to the existing findings of Kang [59]. Table 9 shows the factor loadings from each of the eighteen scales following a principal component analysis of the in situ listening responses. The underlined figures represent the scale with the greatest contribution to each component. The bold figures represent the component with the greatest contribution from each scale. Loadings less than 0.1 have been omitted.

	In situ			
	Component			
	1	2	3	4
% of Variance	22.2	21.3	13.9	12.7
Unpleasant Pleasant	**-0.847**	-0.305		0.394
Reverberant Anechoic	**-0.796**	0.245		
Fast Slow	**0.781**		-0.220	0.238
Comfortable Uncomfortable	**0.728**	0.318		-0.218
Directional Universal	**0.588**	-0.285		-0.522
Like Dislike	0.542	**0.572**		
Sharp Flat	0.541	-0.235	**0.625**	
Meaningful Meaningless	0.509	0.241	-0.277	**0.672**
Natural Artificial	0.332	0.296	**-0.612**	
Gentle Harsh	0.265		**-0.571**	
Interesting Boring	-0.200	**0.815**		
Calming Agitating		**0.898**		
Social Unsocial		**0.607**		0.324
Hard Soft			**0.857**	
Smooth Rough		**0.907**		
Noisy Quiet		0.474	**0.676**	
Far Near				**0.513**
Varied Simple			0.201	**0.887**

Table 9 - Principal component analysis of semantic scale responses for in situ listening (MediaCity; N=11; Varimax rotation and Kaiser normalisation)

Table 10 shows the factor loadings from each of the eighteen scales following a principal component analysis of Kang's in situ listening responses at two urban locations in Sheffield. Note that Kang used a 'meaningful-insignificant' scale whereas this research used 'meaningful-meaningless'. Kang interpreted dimensions of relaxation (26%), communication (12%), spatiality (8%) and dynamics (7%).

	Kang* (in situ)			
	Component			
	1	2	3	4
% of Variance	26	12	8	7
Comfort Discomfort	**0.701**	0.164	0.138	
Quiet Noisy	**0.774**			
Pleasant Unpleasant	**0.784**	0.258	0.157	
Boring Interesting	**0.435**	0.272	0.274	0.103
Natural Artificial	**0.532**	0.102	0.240	
Like Dislike	0.519	**0.575**	0.247	0.151
Gentle Harsh	0.502	**0.531**	0.123	
Hard Soft				**0.812**
Fast Slow				**0.827**
Sharp Flat	0.220		0.345	**0.488**
Directional Universal	0.234		**0.441**	0.267
Varied Simple	0.115		**0.674**	0.167
Reverberant Anechoic	0.204		**0.531**	
Far Near			**0.550**	
Social Unsocial		**0.672**	0.462	
Meaningful Insignificant	0.126	**0.585**	0.469	
Calming Agitating	-0.143	**0.708**	0.286	
Smooth Rough		**0.683**	0.396	

Table 10 - Kang's in situ responses [59] **(* Two urban locations in Sheffield; *N*=491; KMO=0.798; PCA with Varimax rotation and Kaiser normalisation.)**

	In situ							
	Dimension							
	1		2		3		4	
% of Variance	22.2		21.3		13.9		12.7	
Semantic scale & Factor loading	**Unpleasant Pleasant**	**-0.847**	**Smooth Rough**	**0.907**	**Hard Soft**	**0.857**	**Varied Simple**	**0.887**
	Reverberant Anechoic	**-0.796**	**Calming Agitating**	**0.898**	**Noisy Quiet**	**0.676**	**Meaningful Meaningless**	**0.672**
	Fast Slow	**0.782**	**Interesting Boring**	**0.815**	**Sharp Flat**	**0.625**	Directional Universal	-0.522
	Comfortable Uncomfortable	**0.729**	**Social Unsocial**	**0.607**	**Natural Artificial**	**-0.612**	**Far Near**	**0.513**
	Directional Universal	**0.588**	**Like Dislike**	**0.572**	**Gentle Harsh**	**-0.571**		
	Like Dislike	0.542						
	Sharp Flat	0.541						
	Meaningful Meaningless	0.509						

Table 11 – Semantic scale loadings on four principal components from in situ listening responses (MediaCity; *N*=11; PCA with Varimax rotation and Kaiser normalisation)

	Kang* (in situ)							
	Dimension							
	1		2		3		4	
% of Variance	26		12		8		7	
Semantic scale & Factor loading	**Unpleasant Pleasant**	**-0.784**	**Calming Agitating**	**0.708**	**Varied Simple**	**0.674**	**Fast Slow**	**0.827**
	Noisy Quiet	**-0.774**	**Smooth Rough**	**0.683**	**Far Near**	**0.550**	**Hard Soft**	**0.812**
	Comfort Discomfort	**0.701**	**Social Unsocial**	**0.672**	**Reverberant Anechoic**	**0.531**	**Sharp Flat**	**0.488**
	Natural Artificial	**0.532**	**Meaningful Insignificant**	**0.585**	**Directional Universal**	**0.441**		
	Like Dislike	0.519	**Like Dislike**	**0.575**				
	Gentle Harsh	0.502	**Gentle Harsh**	**0.531**				
	Boring Interesting	**0.435**						

Table 12 – Semantic scale loadings on four principal components from Kang's in situ responses [59] **(* Two urban locations in Sheffield; *N*=491; KMO=0.798; PCA with Varimax rotation and Kaiser normalization)**

Table 11 shows the semantic scale loadings on all four rotated and normalised principal components from Phase Two's in situ tests at MediaCity. Table 12 shows the loadings from Kang's in situ tests in Sheffield (with pleasant/unpleasant and quiet/noisy scales reversed

for ease of comparison with Table 11). Scales in bold indicate that the scale loading on the associated component was greater than on any other component. Further loadings with magnitude greater than 0.5 are also included.

Interestingly, the first component (22.%) from in situ responses draws together appreciation - 'pleasant' and 'comfortable' – with spatiality - 'anechoic' and 'directional' – and dynamics – 'fast'[13]. This does not correlate with existing findings such as Kang's [59] whereby the first component can be defined as 'relaxation'. Instead it may be interpreted as an evaluation of the spatiality and pace of the soundscape (and probably the location). In section 4.2.1.2 this first component is further discussed and contextualised in order to understand its definition.

The second component (21.3%) from in situ responses accounts for a very similar portion of the overall variability as the first component (22.2%). The 'smooth', 'calming', 'interesting', 'social' and 'like' contributions are consonant with Kang's second dimension, communication (12%), of 'calming', 'smooth', 'social', 'meaningful', 'like' and 'gentle' (see Table 11).

[13] The semantic scales are here referred to by their relative polarity at one (arbitrary) end of the respective component dimension.

The third component (13.9%) represents a dimension of ‘quiet’, ‘soft’, ‘flat’, ‘natural’, and ‘gentle’ and is thus closely tied to the signal itself. It equates to Kang’s dynamics dimension but may be better represented as ‘timbre and dynamics’.

The fourth component (12.7%) represents a dimension of ‘varied’, ‘meaningful’, ‘far’ and ‘universal’. This is thematically similar to Kang’s third (spatiality) dimension of ‘varied’, ‘far’, ‘reverberant’, and ‘directional’. (‘meaningful’ also contributes to Kang’s third dimension, almost as much as it does to the second dimension, communication.)

4.2.1.2.1 SUMMARY

Eleven participants completed the questionnaire including 18 semantic differentials. Nine of the eleven participants worked at MediaCity and were taking a lunch break. A principal component analysis of the results produced four dimensions. These dimensions were identified as:

1. Spatiality/pace evaluation
2. Content evaluation
3. Timbre and dynamics
4. Spatiality

4.2.1.3 Semantic Differentials' Soundscape Dimensions: Ambisonic versus Binaural

The ambisonic and binaural listening tests produced similarly weighted dimensions, as can be seen in Table 13 and Table 14 where the semantic scales' loadings are ordered on each component. The number of observations was greater than the number of variables for the Ambisonic data but not for the binaural data. Therefore it was only possible to conduct the Bartlett's test of sphericity and the KMO measure of sampling adequacy for the Ambisonic data: $X^2(20)=238.5$, $p<0.0001$; KMO=0.350. The low sampling adequacy means that any conclusions made from the PCA cannot be held with high confidence. Further investigation would be required using a greater number of participants. However, given the somewhat abstracted nature of the top level analysis, the results are presented. Scales in bold indicate that the scale loading on the associated component is greater than on any other component. Further loadings with magnitude greater than 0.5 are also included.

	Ambisonic							
	Dimension							
	1		2		3		4	
% of Variance	23.3		15.9		15.7		12.2	
Semantic scale & Factor loading	**Gentle Harsh**	**0.903**	**Meaningful Meaningless**	**0.876**	**Sharp Flat**	**0.851**	**Far Near**	**-0.800**
	Calming Agitating	**0.848**	**Unpleasant Pleasant**	**-0.699**	**Directional Universal**	**0.838**	**Natural Artificial**	**0.678**
	Hard Soft	**-0.830**	**Like Dislike**	**0.588**	**Smooth Rough**	**-0.651**	**Social Unsocial**	**0.511**
	Comfortable Uncomfortable	**0.726**	**Reverberant Anechoic**	**0.539**	**Varied Simple**	**0.538**		
	Smooth Rough	0.590	**Fast Slow**	**0.496**	**Noisy Quiet**	**0.495**		
	Reverberant Anechoic	0.504	**Interesting Boring**	**0.437**				

Table 13 - Semantic scale loadings on four principal components from Ambisonic reproduction responses (MediaCity; *N*=20; PCA with Varimax rotation and Kaiser normalisation)

	Binaural							
	Dimension							
	1		2		3		4	
% of Variance	25.8		18.3		16.0		12.6	
Semantic scale & Factor loading	**Sharp Flat**	**0.962**	**Comfortable Uncomfortable**	**0.834**	**Interesting Boring**	**0.888**	**Far Near**	**-0.804**
	Hard Soft	**0.777**	**Unpleasant Pleasant**	**-0.742**	**Like Dislike**	**0.831**	**Directional Universal**	**0.600**
	Smooth Rough	**-0.763**	**Social Unsocial**	**0.704**	**Gentle Harsh**	**0.781**	**Reverberant Anechoic**	**0.395**
	Fast Slow	**0.727**	**Meaningful Meaningless**	**0.684**	**Varied Simple**	**0.600**		
	Calming Agitating	**-0.681**	**Natural Artificial**	**0.518**				
	Noisy Quiet	**0.639**						
	Social Unsocial	0.622						
	Varied Simple	0.536						

Table 14 - Semantic scale loadings on four principal components from binaural reproduction responses (MediaCity; *N*=14; PCA with Varimax rotation and Kaiser normalisation)

Gentle-harsh and hard-soft moved from the third dimension, *timbre and dynamics*, in situ to the first dimension, entity and effect, for ambisonics. This trend was augmented in the binaural responses where sharp-flat, hard-soft and fast-slow moved from in situ's *timbre*

and dynamics to the first binaural dimension. The ambisonic reproduction saw pleasant-unpleasant move from the first (in situ) to the second dimension making appraisal no longer the primary means of evaluation.

The first component (23.3%) from the Ambisonic listening test responses represents a dimension of variance of 'gentle', 'calming', 'soft' and 'comfortable'[14]. The 'comfortable' and 'calming' scales relate to listener acceptance (and the motivation-affordance fit nature of the in situ first dimension). However, the 'gentle' and 'soft' (and 'smooth') relate to timbre and dynamics (as per the in situ third dimension). The word 'gentle'[15] sits somewhere between 'calming' and 'soft/smooth', in the sense that it is understood as a description of an effect. Soft/smooth describes the sound as an entity; comfortable describes the physical effect of the entity; gentle describes the effect of the entity in action; and calming describes the effect on a subject – whether the cause is embodied in an entity or not. Collectively these semantics define a dimension describing the sound as an entity and its effect on the subject.

[14] The semantic scales are here referred to by their relative polarity at one (arbitrary) end of the respective component dimension.

[15] Gentle: "moderate in action, effect, or degree; not strong or violent".

The first component (25.3%) from the binaural listening test responses represents a dimension of variance of ‘flat’, ‘soft’, ‘smooth’, ‘slow’, ‘calming’, and ‘quiet’. This relates to the Ambisonic first dimension of ‘entity and effect’ but there is a strong focus on the sound entity and only ‘calming’ to describe effect.

The second component (15.9%) from the Ambisonic listening responses represents a dimension of ‘meaningful’, ‘pleasant’, ‘like’, ‘reverberant’, ‘fast’ and ‘interesting’. ‘meaningful’ and ‘interesting’ relate primarily to the sound content whilst ‘like’ and ‘pleasant’ relate primarily to the subject’s evaluation. (This is not to say that meaning and interest are purely objective, they are indeed subjective.) This content and evaluation is also linked to spatial characteristics and, to a lesser extent, temporal characteristics.

The second component (18.3%) from the binaural listening responses represents a dimension of ‘comfortable’, ‘pleasant’, ‘social’, ‘meaningful’ and ‘natural’. There are no spatial or temporal elements as per the Ambisonic second dimension. Although natural-artificial could be a description of the sound content and/or sound entity, principally this is an appraisal of the sound content.

The third component (15.7%) from the Ambisonic responses represents variation of ‘flat’, ‘universal’, ‘smooth’, ‘simple’ and ‘quiet’. This is a dimension of spatiality, timbre and dynamics. It could also be generalised as a description of the sound entity and space. The

third component (16.0%) from the binaural responses represents a dimension of ‘interesting’, ‘like’, ‘gentle’ and ‘varied’. It is an evaluation of the sound content and sound entity.

The fourth dimension (12.2%) from the Ambisonic responses represents a dimension of ‘near’, ‘natural’ and ‘social’. This descriptor of content and space bares a strong connect to how one might describe an environment holistically. ‘near – far’ is the strongest contribution and the other variables may follow from the subjects’ choice of whether to put accent on the near or the far. For example, if the distant soundscape is social, and the near unsocial, the choice (a U shaped variability – shown in Appendices 7.3) of unsocial – social will depend upon the near – far selection. The fourth component (12.6%) from the binaural responses is space-centric. It represents a dimension of ‘far’, ‘universal’ and ‘anechoic’.

It should be noted that interpretations have been based on singular dictionary definitions of the semantics: plurality has been neglected. The reality is that people attribute words with different meanings, and with differing levels of understanding. In the public domain there is indeed a lack of common syntax surrounding sound. The potentially unfamiliar application of (otherwise familiar) words will incur an additional level of subjective variation.

4.2.1.4 Summary

The ambisonic listening responses produced four dimensions interpreted as:

1. Sound entity and effect (23%)
2. Sound content evaluation (16%)
3. Spatiality, timbre and dynamics (16%)
4. Near-far phenomenon (12%)

The binaural listening responses produced four dimensions interpreted as:

1. Sound entity (26%)
2. Sound content evaluation (18%)
3. Sound entity/content evaluation (16%)
4. Spatiality (13%)

Twenty participants completed the questionnaire for the ambisonic reproduction and fourteen for the binaural reproduction. The sampling adequacy was very low for the PCA of ambisonic listening responses and therefore the analysis can be held with only limited confidence.

4.2.2 Semantic Categorisation of Open Responses

4.2.2.1 Four Semantic Categories

Twenty participants completed Phase Two's Ambisonic soundscape questionnaire and fourteen completed the binaural soundscape questionnaire. Semantic category occurrences were compared for in situ, Ambisonic and binaural listening[16]. Figure 16 shows the average occurrence (per participant) of four semantic categories for three listening scenarios.

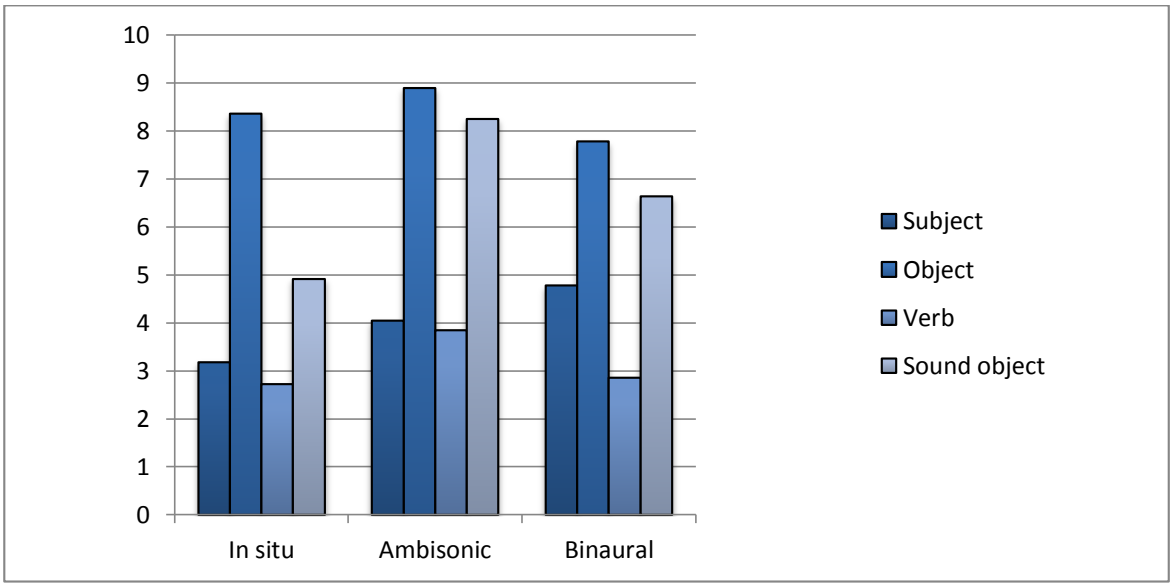


Figure 16 – Mean average semantic category occurrences for in situ, Ambisonic and binaural listening responses

[16] Guastavino's method of ecological validity assessment is not employed with the MediaCity related data collected in Phase Two because – owing to the place-specific nature of Phase Two – a comparison may not necessarily be valid.

A one-way MANOVA was performed on all 225 question responses for the three listening scenarios and four semantic category variables. There was a significant difference in 'sound object' occurrences between in situ and ambisonic listening, $F(2, 205) = 4.836$, $p = 0.009$. A Tukey (HSD) post-hoc analysis was performed on the differences between listening scenarios, as shown in Table 15.

Semantic Category	Contrast	Difference	Standardized difference	Critical value	Pr > Diff	Significant
Subject	Binaural vs In situ	0.321	1.668	2.360	0.220	No
	Binaural vs Ambisonic	0.147	0.884	2.360	0.651	No
	Ambisonic vs In situ	0.174	0.969	2.360	0.597	No
Object	Ambisonic vs Binaural	0.223	0.861	2.360	0.665	No
	Ambisonic vs In situ	0.107	0.385	2.360	0.922	No
	In situ vs Binaural	0.116	0.386	2.360	0.921	No
Verb	Ambisonic vs In situ	0.225	1.367	2.360	0.360	No
	Ambisonic vs Binaural	0.199	1.302	2.360	0.395	No
	Binaural vs In situ	0.026	0.147	2.360	0.988	No
Sound object	**Ambisonic vs In situ**	**0.668**	**3.062**	**2.360**	**0.007**	**Yes**
	Ambisonic vs Binaural	0.321	1.587	2.360	0.254	No
	Binaural vs In situ	0.347	1.480	2.360	0.302	No

Table 15 – Tukey analysis of the differences between in situ, Ambisonic and binaural listening tests' semantic category counts (95% confidence interval; *N*=225)

The number of 'sound object' references in the subjects' open responses is significantly greater for ambisonic listening than in situ listening. This suggests there is a difference in how in situ and ambisonic soundscapes are perceives (contrary to the findings of Guastavino et al.). The greater reference to sound objects within a reproduction soundscape could be explained by the lack of other stimuli - the subject focuses on the sound itself. The reproduction aids they cognitive separation of the sound from the otherwise holistic

environment. However, this analysis does not give any insight as to how or why the ambisonic and binaural results differed.

4.2.2.2 The Developed Psycholinguistic Analysis: In Situ

Eleven participants completed Phase Two's soundscape questionnaire at MediaCity. Here the novel method of analysis was used. The responses were categorised by the semantic categories 'subject', 'object', 'action' and 'sound object'. A principal component analysis took into account the individual cases and provided the inter-relation between the semantic category occurrences with the aim of representing the cognitive process of perception. Each individual question response was processed as an observation therefore eleven participants and the questionnaire of five questions gave N=55.

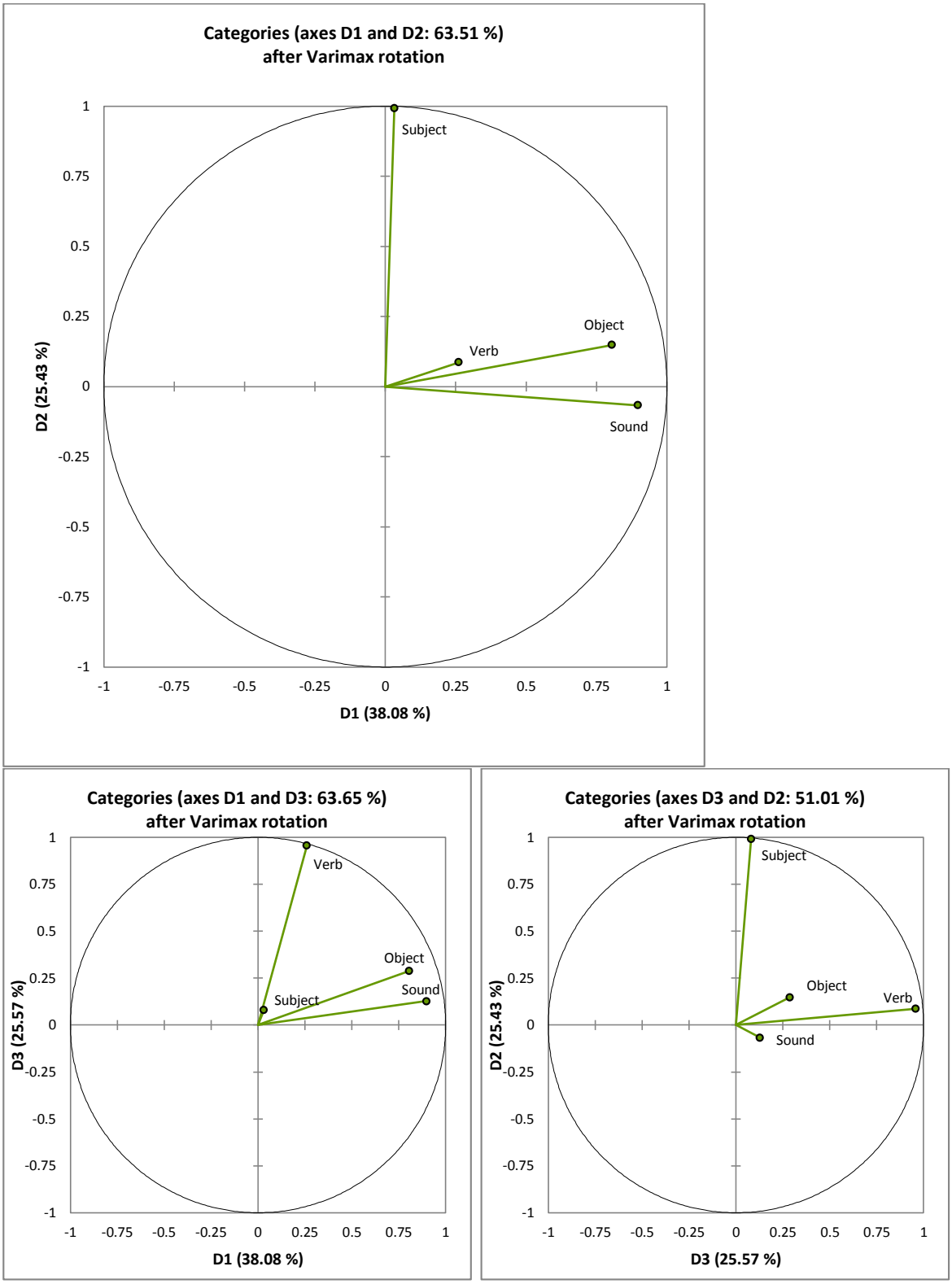


Figure 17 - Semantic categories' contributions to first three factors from in situ listening responses (with Varimax rotation and Kaiser normalisation)

	In situ (*N*=55)		
	Component		
	1	2	3
% of variance	38.1	25.4	25.6
Subject		**0.992**	
Object	**0.805**	0.148	0.288
Action (verb)	0.261		**0.958**
Sound object	**0.897**		0.127

Table 16 – Principal component analysis of Phase Two's MediaCity in situ responses' semantic category counts (Bartlett's X2=34.6, p<0.0001; KMO = 0.651) with Varimax rotation of first three factors and Kaiser normalisation

The first component links object and sound object references. This indicates that sounds were identified in relation to the physical and visual. The observations' factors scores show how each individual response contributed to the components. The two observations with largest factor scores on the first component, 2.26 and 2.18 (both greater than one standard deviation above the mean), exemplify the nature of the first component:

> *"Birds, wind rustling the trees, bell ringing. Engine noise. Feet walking / run for tram."*
>
> *"Workmen / clanging. Faint laughter chatter. Bike wheels moving faintly."*

The second component isolates the variability of subject references. The reference to subject(s) is independent of object, verb and sound object references. It instead varies with the individual and the question. The two observations with largest factor scores on the second component, 3.35 and 3.50 (both greater than one standard deviation above the mean), exemplify the nature of the second component:

"It makes me feel welcome and lets my mind wonder, to go to another place."

"Distant, unintrusive, if you're consciously listening towards it all you're aware of people doing work."

The third component isolates the variability of verb references. The reference to verbs/motion is independent of subject and sound object references, whist weakly linked to object references. These are very dynamic responses that relate more to the visual than the sonic. The observation with largest factor score on the third component, 5.53 (greater than one standard deviation above the mean), exemplifies the nature of the third component:

"People chatting – light, casual, social. Wind – gentle waves of white noise in my ear and faint tree leaves rustling. Heavy vehicles moving and clanking – invisible but implies construction, expansion. Wheeling office chairs across paving – bit comical. Segway going past. Helicopter."

The same dimensions – sound object / object; subject; verb – were produced here with the public as produced in Phase One (in Table 16) with the focus group. However, there is a notable difference in the variance explained by the first component – 'sound object / object' – between Phase One, 53%, and Phase Two, 38%. The users responded with a relatively holistic interpretation of the MediaCity environment whilst the focus group demonstrated a strong focus on source identification. This can be explained by the differing motivation

and expertise between the two groups and is discussed further in the Consolidation chapter, section 5.2.

4.2.2.3 Summary

The Tukey (HSD) analysis, in conjunction with a one-way ANOVA, showed that the number of 'sound object' references in the subjects' open responses was significantly greater for Ambisonic listening than in situ listening.

The developed psycholinguistic analysis gave the in situ dimensions:

1. Sound object and object joint references
2. Subject references
3. Action (verb) references

The entity of sound was described in relation to the physical surrounding, the environmental entity.

4.2.2.4 Developed Psycholinguistic Analysis: Ambisonic versus Binaural

The principal component analyses provided the inter-relation between the semantic category occurrences as a means of representing the cognitive process of perception (Figure 18 and Figure 19). Each individual question response was processed as an observation giving *N*=100 for the ambisonic test and *N*=70 for the binaural test. Figure 18 and Figure 19 show the first three components from principal component analyses of the ambisonic and

binaural listening responses' semantic category occurrences. Table 17 and Table 18 give the factor loadings.

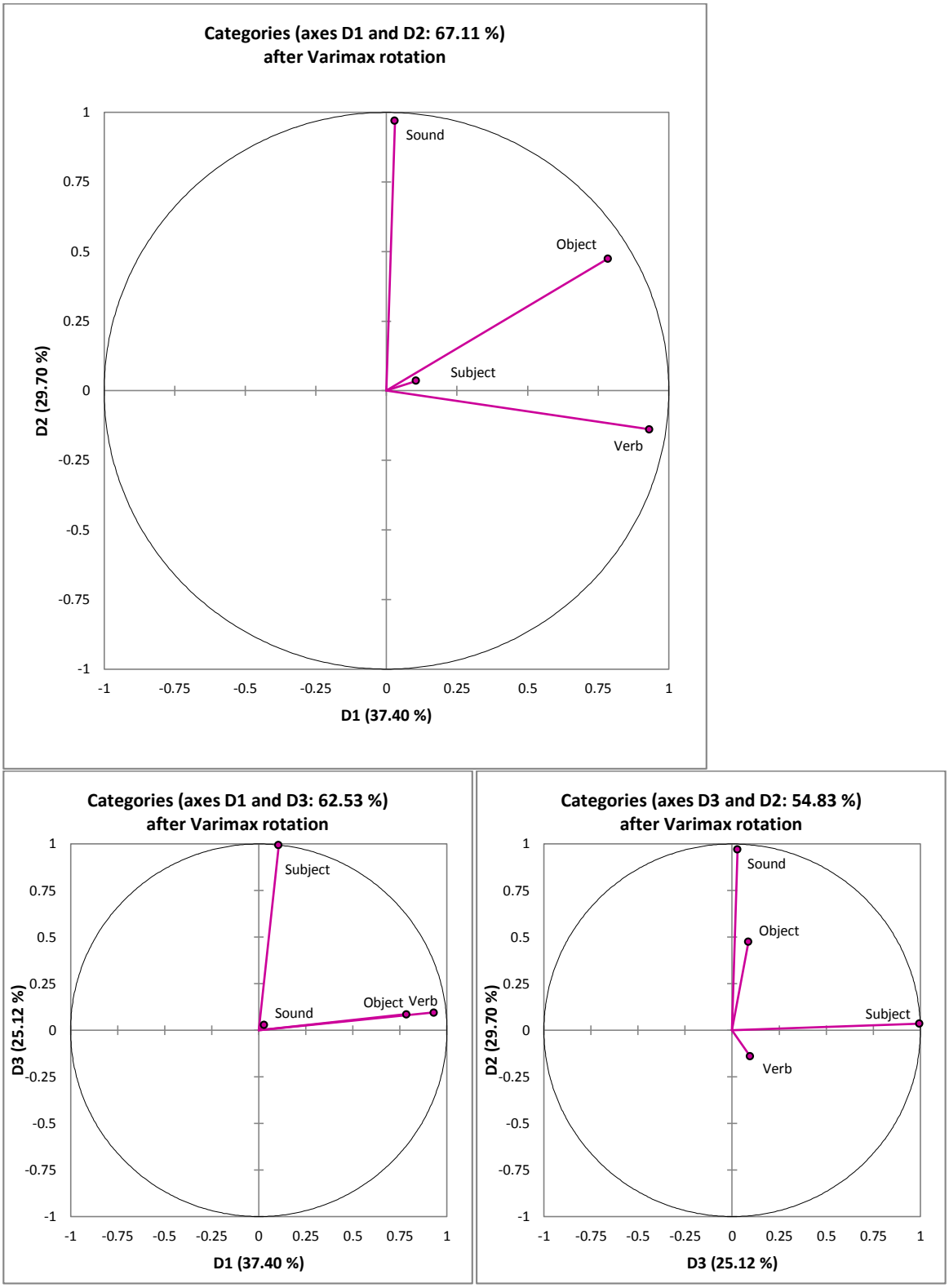
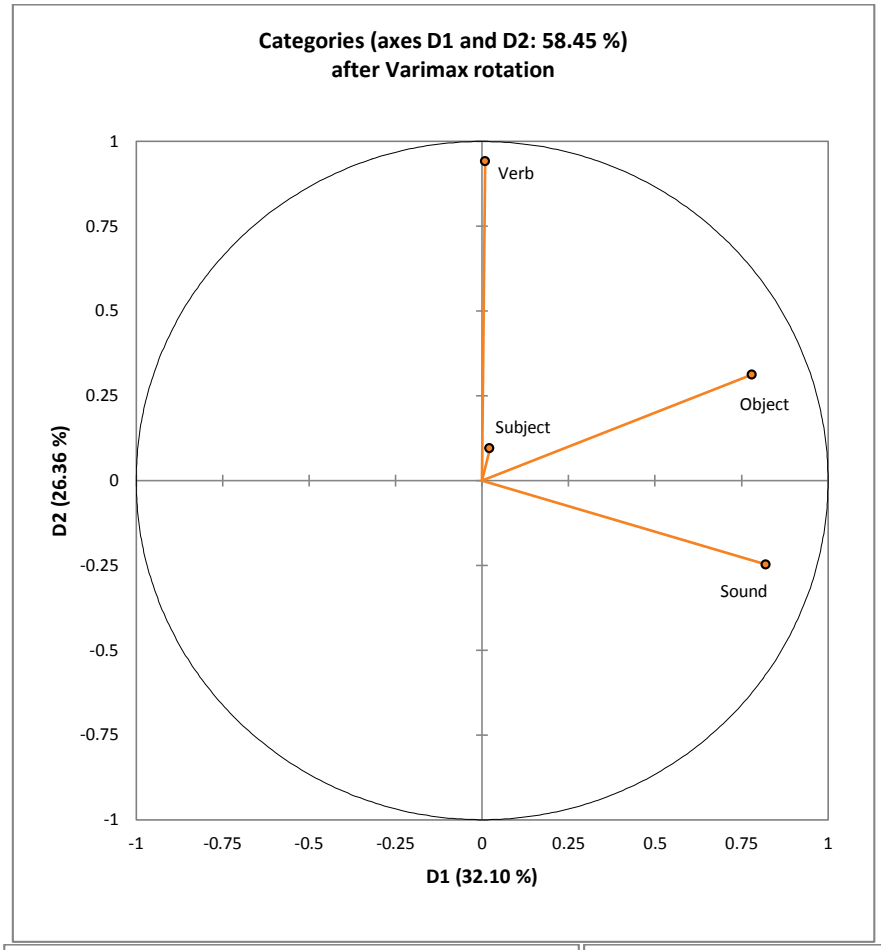
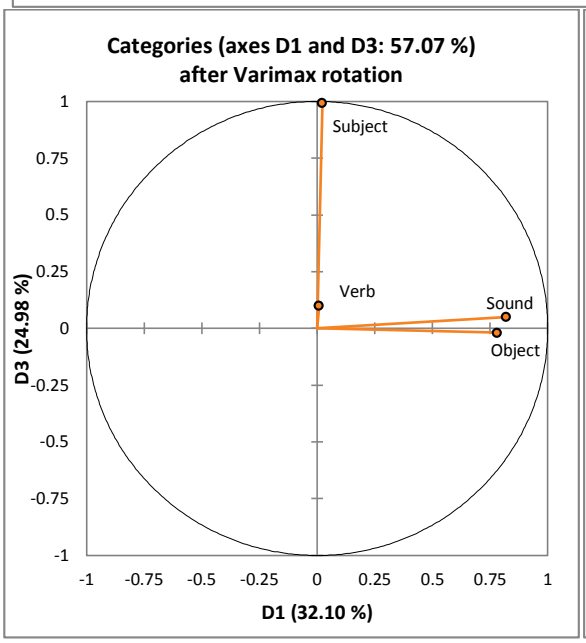


Figure 18 - Semantic categories' contributions to the first three factors from Ambisonic listening responses (Varimax rotation and Kaiser normalisation)





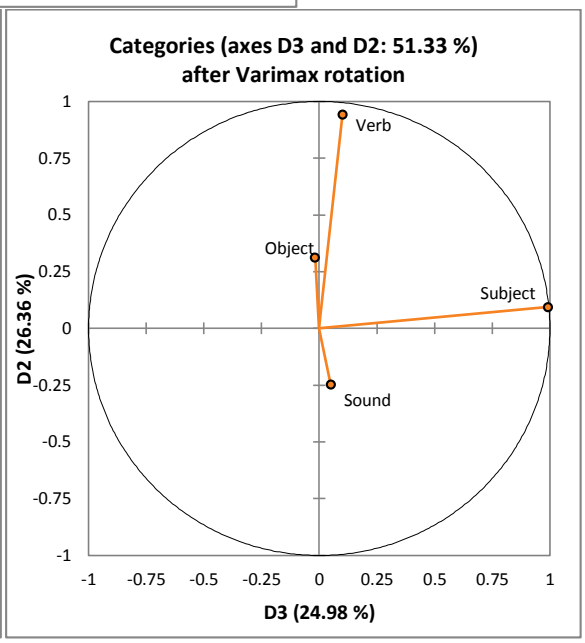


Figure 19 – Semantic categories' contributions to the first three factors from binaural listening responses (Varimax rotation and Kaiser normalization)

	Ambisonic (*N*=100)		
	Component		
	1	2	3
% of variance	37.4	29.7	25.1
Subject	0.106		**0.994**
Object	**0.785**	0.475	
Verb	**0.932**	-0.139	
Sound object		**0.971**	

Table 17 – Semantic categories' contributions from the principal component analysis of MediaCity Ambisonic listening responses (Bartlett's X^2=65.9, p<0.0001; KMO = 0.427; Varimax rotation and Kaiser normalisation)

	Binaural (*N*=70)		
	Component		
	1	2	3
% of variance	32.1	26.4	25.0
Subject			**0.993**
Object	**0.781**	0.312	
Verb		**0.942**	0.102
Sound object	**0.820**	-0.248	

Table 18 – Semantic categories' contributions from the principal component analysis of MediaCity binaural listening responses (Bartlett's X^2=12.6, p=0.02; KMO = 0.482; Varimax rotation and Kaiser normalisation)

Two Ambisonic responses with squared cosines greater than one standard deviation above the mean (0.940 and 0.777) exemplify the nature of the first component, 'action and objects':

> *"Traffic, footsteps, talking, metro."*
>
> *"Conversation (very quiet). Steps. Tram approaching, sound of wheels rolling along rails, squeaking (high pitch noise). Coughs occasionally. Construction work. I feel neutral about everything apart from tram squeak which I did not like."*

Two Ambisonic responses with squared cosines greater than one standard deviation above the mean (0.941 and 0.849) exemplify the nature of the second component, 'sound objects':

"Remove the oscillation in the background sound – it is distracting – the least pleasant part of the soundscape."

"Calm. Variation, but not too much. No loud 'crash' and 'bangs'."

Two binaural responses with squared cosines greater than one standard deviation above the mean (0.876 and 0.843) exemplify the nature of the first component, 'sound objects and objects':

"It is noisy, but I can recognise some events."

"The background noise is a mixture of voices and a general hiss. Quite harsh sounding as if the floor and walls are hard, maybe tiled. Lots of high frequency content."

Two binaural responses with squared cosines greater than one standard deviation above the mean (0.941 and 0.904) exemplify the nature of the second component, 'action':

"It feels like recording was made near a shopping centre. People are related and enjoying themselves."

"Bustling, lots of layers, seems filled with people but non-threatening. Bell of the shop = sociable."

There is a clear similarity between the in situ and binaural dimensions. The relation between sound and the physical/visual is retained with binaural soundfield reproduction. The Ambisonic soundscape of the listening room stimulated a perception whereby sounds were described independently of physical objects. Here object and verb references co-varied in the sense that participants commented on (ideas of) things happening, objects in motion, around them yet sound was identified and described independently.

The Ambisonic and binaural recordings used, although similar, were captured on different days (rather than the very same 2mins) and therefore were not of identical content. However, the soundscapes experienced by each in situ subject differed to at least the same degree since the eleven members of the public did not complete the questionnaire simultaneously. It is recommended that the proposed greater ecological validity of binaural reproduction is reaffirmed with further testing using the same soundfield. This could be achieved by either recording with a soundfield microphone and dummy head simultaneously or repeating the binaural listening test with a binaural clip synthesised from the b-format recordings used in the current ambisonic listening test.

The focus group responses from Phase One's ambisonic reproduction of the MediaCity soundfield (Table 6) produced the same semantic dimensions. Therefore one can rule out there having been any significant influence, in this case, of the focus group's (memory of the) prior in situ listening test on their responses to the later ambisonic reproduction test.

4.2.2.4.1 SUMMARY

The developed psycholinguistic analysis gave the ambisonic dimensions:

1. Verb and object joint references
2. Sound object references
3. Subject references

The developed psycholinguistic analysis gave the binaural dimensions:

1. Sound object and object joint references
2. Verb references
3. Subject references

The semantic categorisation of open responses was analysed with two methods. The overall semantic occurrences for the in situ, Ambisonic and binaural listening scenarios revealed a significant difference in 'sound object' references between in situ and Ambisonic. The developed psycholinguistic analysis gave the same dimensions from in situ and binaural listening tests. In conclusion, the binaural urban soundscape reproduction held external validity over the Ambisonic reproduction.

4.2.3 Phase Two Summary

The semantic differential analysis gave the in situ dimensions: spatiality/tempo evaluation (22%); content evaluation (21%); timbre and dynamics (14%); spatiality (13%). The

ambisonic dimensions were: sound entity evaluation (23%); sound content evaluation (16%); spatiality, timbre and dynamics (16%); near-far phenomenon (12%). The binaural dimensions were: sound entity (26%); sound content evaluation (18%); entity/content evaluation (16%); spatiality (13%). The ambisonic reproduction, not the binaural reproduction, of the MediaCity soundfield stimulated similar dimensions to that of the in situ soundscape. This was comparison was made on the basis that the first two ambisonic dimensions retained the more subjective differentials (e.g. pleasant-unpleasant) whilst the binaural first dimension became mostly objective.

Open responses were analysed by two methods. The Tukey (HSD) analysis, in conjunction with a one-way ANOVA, showed that the number of 'sound object' references in the subjects' open responses was significantly greater for Ambisonic listening than in situ listening. The developed psycholinguistic analysis gave the ambisonic dimensions: verb and object joint references; sound object references; subject references. The binaural dimensions were: sound object and object joint references; verb references; subject references. Binaural reproduction held external validity over ambisonic reproduction.

4.3 PHASE THREE

The in situ real-time binaural listening – or 'electroacoustic ear' – test results are analysed with the production of the **Semantic Differentials**' 'soundscape dimensions' and the ***developed psycholinguistic analysis*** from the **Semantic Categorisation of Open Responses**.

4.3.1 Semantic Differentials

4.3.1.1 Semantic Differentials' Soundscape Dimensions: An Electroacoustic Ear

The principal component analysis of the semantic differential responses from in situ real-time binaural listening presented relatively even loading on the first two factors shown in Table 19. This is similar to that of the in situ results and contrasting that of the Ambisonic and binaural results.

	Real-time binaural in situ							
	Dimension							
	1		2		3		4	
% of Variance	21.9		20.8		15.3		11.0	
Semantic scale & Factor loading	**Smooth Rough**	**-0.920**	**Varied Simple**	**0.884**	**Noisy Quiet**	**0.756**	**Far Near**	**-0.721**
	Sharp Flat	**0.837**	**Like Dislike**	**0.843**	**Social Unsocial**	**0.686**	**Reverberent Anechoic**	**-0.622**
	Directional Universal	**0.785**	**Interesting Boring**	**0.775**	**Unpleasant Pleasant**	**0.647**		
	Calming Agitating	**-0.761**	**Comfortable Uncomfortable**	**0.735**	**Gentle Harsh**	**-0.629**		
	Hard Soft	**0.664**	**Meaningful Meaningless**	**0.596**	**Natural Artificial**	**-0.530**		
			Fast Slow	**-0.488**				

Table 19 - Semantic scale loadings on four principal components from in situ real-time binaural reproduction responses (MediaCity; *N*=17; PCA with Varimax rotation and Kaiser normalisation)

The first component (21.9%) from the in situ real-time binaural listening responses represents a dimension of variance of 'smooth', 'flat', 'universal', 'calming' and 'soft'. It may be deemed a description of the sound entity ('smooth', 'flat' and 'soft'). The fact that directionality/universality varies with such descriptors suggests that this dimension is, more

specifically, a description of the (binaural) signal. That is, the subjects evaluated the sound of the reproduction itself. ‘calming’ has previously been understood as the effect on the subject, be it of tactility or non-physical. It is plausible that the meaning of calming/agitating was (mis-)understood, and used, as a simplistic adjective, a description of a noun (or ‘entity’). In this case, the alternative hypothesis would be that ‘calming’ was used as a description of the sound entity, which in turn was an element of the (spatial) signal. If one is to assume this is not the case - and that ‘calming’ and ‘agitating’ were understood as subjective effects – then this first dimension is not only a description of the binaural signal but of the experience itself.

The second component (20.8%) from the in situ real-time binaural listening responses represents a dimension of variance of ‘varied’, ‘like’, ‘interesting’, ‘comfortable’, ‘meaningful’ and ‘slow’. ‘interesting’ and ‘meaningful’ relate to the sound content, that communicated by sound. ‘varied’ is most likely also in reference to the sound content and the nature of the sound sources. ‘like’ and ‘comfortable’ are appraisals. The greater contribution from ‘like’ (0.843) over ‘comfortable’ (0.735) is concordant with the dimension’s proposed sound content focus since ‘comfortable’ has tactile connotations and ‘like’ does not. The temporal characteristic, ‘slow’, co-varies but given the small contribution (0.488) need not be of great concern. Therefore the second dimension can be understood as sound content appraisal.

The third component (15.3%) represents a dimension of ‘quiet’, ‘unsocial’, ‘pleasant’, ‘gentle’ and ‘natural’. This is an unexpected mix of dynamics, communication, appraisal and timbre. The focus is on the sound entity (‘quiet’, ‘gentle’ and ‘natural’) but in relation to sociability and appraisal. An interpretation assisted by a process of elimination would suggest that this dimension represents an evaluation of the embodied soundscape.

The fourth component (11.0%) represents a dimension of ‘near’ and ‘anechoic’. This is a spatial dimension. Directionality/universality has previously been seen to co-vary with the spatial dimension but in this case of in situ real-time binaural reproduction the directionality/universality scale has been adopted by the first dimension to describe the ‘electroacoustic ear’.

4.3.1.1.1 SUMMARY

The real-time binaural in situ listening responses produced four dimensions interpreted as:

1. Experience evaluation (the binaural signal)
2. Sound content evaluation
3. Soundscape entity evaluation (embodied)
4. Spatiality

4.3.2 Semantic Categorisation of Open Responses

4.3.2.1 Developed Psycholinguistic Analysis: An Electroacoustic Ear

It can be seen from Figure 20 that the principal component analysis revealed semantic relationships different from that of Phase Two's binaural and in situ results. Table 20 shows the factor loadings.

The observation with the largest squared cosine on the first component (0.891) relates to the response that best exemplifies the first dimension, 'action and object':

> *"Building works, people chatting, glasses clinking with ice in it, people walking on the grass, paper rustling."*

The observation with the largest squared cosine on the second component (0.946) relates to the response that best exemplifies the second dimension, 'subject':

> *"Being able to hear something from a distance but seems like its next to you."*

The observation with the largest squared cosine on the third component (0.828) relates to the response that best exemplifies the third dimension, 'sound object':

> *"Slightly harsh/noisy with more prominent high frequency noise. Lots of higher frequencies? Leaves / voices more prominent."*

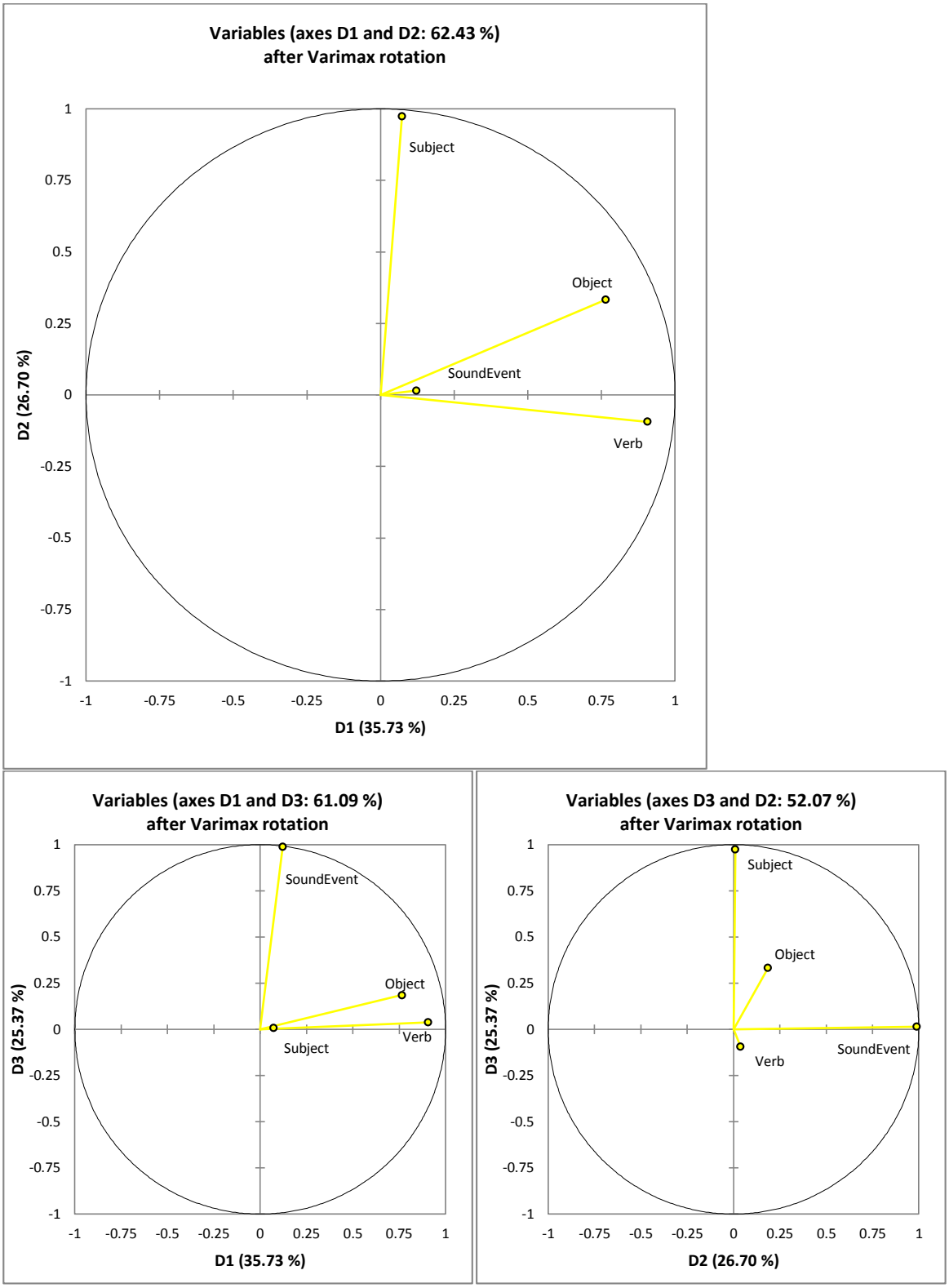


Figure 20 - Semantic categories' contributions to the first three factors from real-time binaural in situ listening responses (Varimax rotation and Kaiser normalization)

	Binaural in situ (n=80)		
	Component		
	1	2	3
% of variance	34.9	31.2	25.3
Subject		**0.974**	
Object	**0.766**	0.333	0.186
Verb	**0.907**		
Sound object	0.122		**0.989**

Table 20 - Semantic categories' contributions to the first three factors from real-time binaural in situ listening responses (Bartlett's X^2=63.17, p<0.0001; KMO=0.592; Varimax rotation and Kaiser normalisation)

The in situ binaural responses, contrary to those of the Ambisonic, showed more variability with subject references than sound object references. This may be explained by the (electroacoustic ear) experience itself, an experience many found "interesting", "weird", "fascinating" and "like an outer-body experience". Dramatised or not, these comments suggest it was a novel personal experience that inherently influenced their written responses with references to self in their descriptions of their sonic experience.

Real-time binaural in situ reproduction and Ambisonic reproduction responses produced the same first dimension of action and object. As per the Ambisonic responses, sound object references were made quite independently of object references. In other words, the sound was detached from its physical source. This may be explained by a shift of participants' focus to the reproduction itself, as opposed to the soundscape as embodied in the environment.

The aforementioned quote that commented on “high frequency noise”[17] is indicative of how some evaluations were framed around ‘an acoustic signal’ rather than a soundscape.

The first and third dimensions from the analysis of semantic differentials represented the binaural experience and the sound environment. Both relate to sound due to the soundscape focus of the questionnaire and the inherent restrictions of semantic differentials. Therefore one might not necessarily expect a subject-sound or object-sound relationship to arise from the analysis of the free-form responses.

4.3.2.1.1 SUMMARY

The developed psycholinguistic analysis of real-time in situ binaural responses gave three dimensions very similar to that of the Ambisonic responses. The sound object references were independent of both object and verb references. Object and verb references co-varied.

[17] The binaural reproduction produced a relatively flat frequency response (see 3.3.2.4.1). However, occasional ill-fitment of the earpiece - resulting in a loss of lower frequencies – may have occurred despite the researcher’s checks.

4.3.3 Phase Three Summary

The real-time binaural in situ listening responses produced four dimensions interpreted as: experience evaluation [the binaural signal] (22%); sound content evaluation (21%); [embodied] soundscape entity evaluation (15%); spatiality (11%). The developed psycholinguistic analysis of real-time in situ binaural responses gave three dimensions very similar to that of the Ambisonic responses. The sound object references were independent of both object and verb references. Object and verb references co-varied. The findings are further discussed in section 5.3.2.

4.4 Discussion and Results Overview

Phase One analysed the focus groups' responses at five in situ soundscapes and their ambisonic reproductions in a listening room. The Kings Arms' semantic differential responses were compared and 'soundscape dimensions' were procured. An existing method of ecological validity assessment (using three semantic categories) was assimilated with the focus groups' open responses across all soundscapes. This was repeated with five semantic categories. The developed psycholinguistic analysis was performed on the all soundscapes' responses individually. The developed psycholinguistic analysis was then re-appropriated as a method of validity assessment of soundfield reproduction.

Phase Two analysed members of the public's responses at the MediaCity soundscape in situ and participants' responses in the listening room for an ambisonic reproduction and a binaural reproduction. The semantic differential responses were compared and 'soundscape dimensions' were produced. The open responses were categorised into four semantic categories and – prior to the principal component analysis of the developed psycholinguistic analyses – a Tukey (HSD) analysis was performed in conjunction with a one-way ANOVA.

Phase Three analysed the public's responses from the real-time binaural reproduction listening test in situ at MediaCity. Semantic differentials' soundscape dimensions were analysed and the developed psycholinguistic analysis was performed and discussed.

A large quantity of rich data was collected. A total of 111 questionnaires were completed, 50 from the focus group and 61 from other participants, providing 71 sets of semantic differential responses and 555 data of open responses comprising a total of 6706 words.

In Phase One the Kings Arms ambisonic reproduction was evaluated as less social – due to the lack of the accompanying focus group – and more eventful. The semantic differential analysis showed that the ambisonic reproduction was evaluated less in terms of appraisal and more in terms of objective semantics.

The assimilation of an existing method of ecological validity assessment with the focus group indicated 'ecological validity' with greater significance. Therefore subject memory – agency and expertise – had an influence on the external validity of soundfield reproduction.

The five different soundscapes were evaluated in different ways. 'subject' references were significantly linked to person-place relationship as well as estimated number of visits annually. The developed psycholinguistic analysis showed differing variations in the perception of in situ and ambisonic soundfields across the five soundfields. The developed

psycholinguistic analysis was introduced as an assessment of ecological validity. Only the Triangle and Kings Arms soundscapes held ecological validity in the ambisonic reproductions. A collective analysis of all five soundscapes indicated an ecological invalidity of ambisonic soundfield reproduction. This can be deemed the general case, whilst cautious of the soundscape-specific nature of ecological validity.

In Phase Two the semantic differential analysis gave the in situ dimensions: spatiality/tempo evaluation (22%); content evaluation (21%); timbre and dynamics (14%); spatiality (13%). The ambisonic dimensions were: sound entity evaluation (23%); sound content evaluation (16%); spatiality, timbre and dynamics (16%); near-far phenomenon (12%). The binaural dimensions were: sound entity (26%); sound content evaluation (18%); entity/content evaluation (16%); spatiality (13%). The ambisonic reproduction, not the binaural reproduction, of the MediaCity soundfield stimulated similar dimensions to that of the in situ soundscape. This was comparison was made on the basis that the first two ambisonic dimensions retained the more subjective differentials (e.g. pleasant-unpleasant) whilst the binaural first dimension became mostly objective.

Open responses were analysed by two methods. The Tukey (HSD) analysis, in conjunction with a one-way ANOVA, showed that the number of 'sound object' references in the subjects' open responses was significantly greater for Ambisonic listening than in situ listening. The developed psycholinguistic analysis gave the ambisonic dimensions: verb and

object joint references; sound object references; subject references. The binaural dimensions were: sound object and object joint references; verb references; subject references. Binaural reproduction held external validity over ambisonic reproduction.

In Phase Three the real-time binaural in situ listening responses produced four dimensions interpreted as: experience evaluation [the binaural signal] (22%); sound content evaluation (21%); [embodied] soundscape entity evaluation (15%); spatiality (11%). The developed psycholinguistic analysis of real-time in situ binaural responses gave three dimensions very similar to that of the Ambisonic responses. The sound object references were independent of both object and verb references. Object and verb references co-varied.

CHAPTER FIVE

5 CONSOLIDATION

The semantic differentials' soundscape dimensions are re-interpreted by relating **Semantic Differentials and Appraisal Theory**. The **External Validity of Soundfield Reproduction** is discussed in relation to both semantic differentials' soundscape dimensions and the developed psycholinguistic analysis before comparing ambisonic and binaural reproduction. Phases Two and Three's results are discussed in terms of **Embodiment and Ventriloquism**. The roles and influences of the variables around which the methodology was inspired/designed – Error! Reference source not found. – are discussed in light of the esearch findings.

5.1 Semantic Differentials and Appraisal Theory

A soundscape dimension model is proposed. The proposed model is not soundscape specific. It has been used to explain – and account for – the variation represented by the first two principal components arising from Phase Two's in situ listening and from Kang's in situ listening tests' findings. Appraisal theory was introduced in section 2.3.3 and it is here that it comes into play in the interpretation of semantic differentials' soundscape dimensions.

The first component (22.%) from in situ responses drew together appreciation - 'pleasant' and 'comfortable' – with spatiality - 'anechoic' and 'directional' – and dynamics – 'fast'. This does not correlate with existing findings such as Kang's [59] whereby the first component can be defined as 'relaxation'. This first factor must be contextualised in order to further understand its definition.

The majority of people using the open outdoor space at MediaCity between the hours of 12:00 and 15:00, and hence the majority of participants, were those taking a lunch break from work in the offices of the immediate vicinity. In fact nine out of the eleven subjects tested were doing just this. They had actively chosen to spend their rest/leisure time outdoors. The choice implies there was *motivation*. The outdoor soundscape contrasted with the office soundscape in spatial terms: the lack of echo and (partly consequentially) its directionality. It is also probably faster in its temporal variation or least one is more aware

of the pace as one's mind slows, resting from the focuses of work. Given this influence of the participants' place-relationship with the soundscape, the first dimension can be understood as how appropriate the soundscape *affordances* are for the earwitness' requirements or desires. It is proposed that the first dimension should be understood as 'motivation-affordance fit'.

A comparison with Kang's soundscape dimensions showed variation that can be explained by the differing affordances (or purposes) of the soundscape locations and the differing motivations for the earwitnesses' presence there. In the case of Kang's soundscapes, those of urban parks, the motivation for public attendance is clear: to relax in a peaceful environment. Indeed Kang's first dimension included 'quiet', 'boring' and 'natural'. There is quite a strong consensus on the role of urban parks and hence the 'non-user' participants used in Kang's study may well have partly evaluated the parks' soundscapes – and the motivation-affordance fit – in terms of the obvious hypothetical motivation, to relax.

The second dimensions both contained elements of sound content/communication and appraisal. The naming of a communication dimension somewhat neglects the contributions of 'calming', 'smooth' and 'like' in both instances. It is proposed that the second dimension should be understood as 'mediation'. It represents the emotions extracted from the motivation-affordance fit alongside the reasoning that relates sound to place.

These two dimensions, 'motivation-affordance fit' and 'mediation' are of particular significance not only because they explain the greatest variance but because they align sonic descriptors (such as 'artificial') with appraisals (such as 'like'). The third and fourth dimensions do not do this and may be deemed the 'left-overs' – the sound entity/spatiality semantics that did not play a role in the subjects' judgements of motivation-affordance fit in the first dimension. Their grouping merely reflects their lexical meaning. The second dimensions generally contained all the sound content/communication semantics for the subjects' mediations. It is hypothesised that if more content/communication differentials had been included, there would have been another dimension of sound content/communication 'left-overs' too. The fewer number of content differentials than entity differentials meant this was not the case with the 18 semantic differentials used.

The model still holds for the reproduction listening semantic differentials' soundscape dimensions. The difference with the binaural was that the *motivation-affordance fit* and *mediation* dimensions were manifest in the second and third dimensions. Instead of the primary dimension of evaluation being a part of the subject's appraisal, it was purely objective (with the exception of calming-agitating): the 'timbre and dynamics' dimension from the in situ listening was promoted from third to first. Gentle-harsh and hard-soft moved from the third dimension, *timbre and dynamics*, in situ to the first dimension, entity and effect, for ambisonics. For both ambisonic and binaural reproductions, pleasant-unpleasant moved from the first (in situ) to the second dimensions.

Unfortunately a detailed comparison with Phase One's semantic differential results is not possible due to the differing semantic scales used. However it can still be noted that – as per the Phase One results – the reproductions were evaluated more objectively than the in situ soundscapes. Further investigation is advised to (dis)prove the alternative hypothesis that the objective focus of the reproduction listening tests was not due to the reproduction itself but the agency and expertise of the participants.

5.1.1.1 Expectation and Motivation-affordance fit

A good match between the subject's motivations and the soundscape's affordances gives the subject a positive appraisal and thus a positive emotion is extracted. This relates somewhat to the theory of expectation: the idea that a positive appraisal follows the matching of a soundscape and the subject's expectation [59]. However, the expectation theory does not take into account appreciation of the unexpected. For example, during the Phase One testing with the focus group at MediaCity a distant choir could be heard: it was both unexpected and incongruous with the immediate surroundings but all participants reviewed the sound as positive. It is motivation that may (or may not) summon expectation. One may have a pre-conceived idea of how a motivation-affordance fitting environment may sound based on prior experience but that is not to say that an unexpectedly sounding environment will not also be motivation-affordance fitting and positively appraised.

The notion of ‘motivation-affordance fit’ encompasses the particular scenarios where expectation plays a role but also scenarios where the unexpected meets the earwitness’s motivations and is positively appraised. It also avoids diminishing the role of the sonic as merely following the visual. In the fields of urban and architectural theory cities are understood as multifaceted, temporal and there to be explored. The idea of dwellers purely seeking to journey through the city uninhibited by the unexpected would be an alien phenomenon. It would point to an ideal that manifests as architectural homogeneity.

Expectation theory is not reliable and, most importantly, the implicit doctrine that a good soundscape follows its visual counterpart is not constructive in the promotion of soundscapes. Therefore it is proposed that the theory of expectation be transcended by the theory of motivation-affordance fit.

Beyond the purpose of replacing expectation theory, the theory of expectation-affordance fit is broadly applicable due to its relation to the process of appraisal. The interpretations of the semantic differentials’ soundscape dimensions also incorporated the appraisal process of mediation. This was due to the test nature of the evaluations: the soundscape questionnaire itself encouraged descriptive listening – and hence mediation – as well as holistic listening. However, in ‘normal’ listening conditions holistic listening is a prominent listening state with regards to the soundscape. Therefore it is the motivation-affordance fit that should be central to the (professional) evaluation of soundscapes.

5.1.1.2 Application in Urban Planning and Design

Place affordance is a familiar notion in the field of architecture whereby spaces are designed around people and purpose. It is therefore useful to extend the notion of place affordances to the sonic. The concept of 'zoning' seeks to break the tendency for homogeneity across a city (and its soundscapes) by defining spaces with clear place identities. This could be as simple as emphasising a space as a place where one can relax, or work, or play. It is the geographically-specific design of a motivation-affordance fit.

The occurrence of a motivation-affordance fit dimension within the semantic differentials' soundscape dimensions poses a method of assessing the motivation-affordance fit of a soundscape. It also provides the ability to identify which features/characteristics of a soundscape are perceived as key affordances in relation to the subject's motivations, and in turn their appraisal of the soundscape.

5.1.1.3 Summary

It is proposed that a PCA of semantic differential responses for any in situ listening tests will produce two dimensions that represent the motivation-affordance fit and mediation identified in the psychological theories of appraisal. When using Kang's 18 semantic differentials in situ the following dimensions were interpreted: 1) **Motivation-affordance fit;** 2) **Mediation;** 3) Sound entity; 4) Spatiality. Furthermore, it is proposed that the theory of 'expectation' is transcended with the theory of 'motivation-affordance fit'.

5.2 External Validity of Soundfield Reproduction

5.2.1 Ecological Validity and External Validity

The ecological validity of a study is how well the materials and setting of the study approximate the real-world that is being examined. External validity is the degree to which the results can be *generalised* to different participants, places and time periods [92]. The external validity of a reproduction soundscape listening test is how well the findings can be generalised to not only other situations (the original soundscape in situ) but other people (e.g. non-expert users).

It was found in Phase One, section 4.1.2.1, that Guastavino's method of assessing 'ecological validity' [93] extracted greater validity when comparing the in situ and Ambisonic responses of the focus group opposed to members of the public in situ and volunteer colleagues and students in a listening room (as per the existing study). The participants themselves played a significant role in the differing perceptions of the in situ soundscape perception tests and the Ambisonic reproduction soundscape perception tests. The perceptions (extracted by this particular method) were not due to ecology alone but the subjects' *agency* and *expertise*. Therefore Phase One assessed ecological validity whereas Guastavino, unknowingly, actually assessed external validity.

5.2.1.1 MediaCity Soundscape: Users and Non-users

There is an observable difference in the variance explained by the first component, 'sound object / object', between Phase One, 53%, and Phase Two, 38% (as shown in Table 21).

	Focus group			Public		
	Component			Component		
	1	2	3	1	2	3
% of variance	44.6	25.2	26.6	38.1	25.4	25.6
Subject		**0.988**	-0.152		**0.992**	
Object	**0.893**		0.350	**0.805**	0.148	0.288
Verb	0.219	-0.174	**0.956**	0.261		**0.958**
Sound object	**0.969**			**0.897**		0.127

Table 21 – Principal component analyses of Phase One's (focus group) and Phase Two's (public) in situ open responses at MediaCity (KMO = 0.512 and 0.651; Varimax rotation and Kaiser normalisation)

The users (Phase Two) responded with a relatively holistic interpretation of the MediaCity environment whilst the focus group (Phase One) demonstrated a strong focus on source identification. This is counter to the trend for 'experts' to focus on the properties of the sound itself (i.e. the second component) but may be explained by the differing motivations between the two groups.

The appraisal of the soundscape encompasses a motivation-affordance. In the case of the users, their motivation (for being there) was generally to take a lunch break away from the office and the affordance was the open outdoor space (4.2.1). This affordance manifested both sonically and non-sonically. Therefore the users' responses are based not only on the

soundscape but the place as a whole. However, in the case of the non-users, there is a lack of agency and motivation comparable to that of the users. Their primary motivation was to complete the questionnaire as best they could (and perhaps as quickly as they could)[18]. Therefore the affordance was primarily sonic: a good motivation-affordance fit would be extracted from a soundscape consisting of easily-identifiable sound and plenty of it.

It is not possible to isolate the influence of user/non-user or expert/non-expert from the results. Therefore further research is recommended: to repeat the questionnaire with 'non-experts' *taken* to MediaCity.

5.2.2 Methods of Assessment

External validity and ecological validity are both method- and data- specific. For example if open responses to an urban soundscape questionnaire are to be collected from subjects experiencing soundfield reproductions, the reproducibility of a set of semantic differentials' soundscape-specific dimensions will not necessarily infer the validity of the open responses collected – and vice versa.

[18] It is noteworthy that participants deemed it difficult and somewhat unnatural responding to the soundscape questionnaire.

The experimentation presented here collected two data types: the quantitative semantic differentials and the qualitative open responses. The three methods constituted the in situ/reproduction comparisons of: semantic differential analyses; the frequencies of semantic category occurrences in open responses; and the developed psycholinguistic analyses. Strictly speaking, laboratory experiments should be validated with 'real-world' collection of the *same* data type (be it conducted or referenced from existing research). Essentially one seeks external validity – although this may also manifest as ecological validity (as explained in section 5.2.1).

Laboratory semantic differential responses should be validated with in situ semantic differential responses. This means at least observing the continuity of the differentials' evaluations (as, for example, performed in section 4.2.1.1) and at best observing the continuity of the differentials' PCA dimensions (as, for example, performed in sections 4.2.1.2 and **Error! Reference source not found.**). The latter method offers the additional erification that the in situ and laboratory test subjects have engaged the descriptors similarly in their cognitive processes of auditory perception (evidenced from the PCA's interpretation of the differentials' inter-relational variance). The principal component analysis delves somewhat below the surface of a priori specificity, towards perception – that one (wishfully) seeks to measure. Explicitly, subjectivity is considered not only within the confines of a priori dichotomies but the perceived inter-relation between differentials.

As for a measure of validity, the external validity (or ecological validity) can be assessed by observational comparison of the dimensions as per Davies et al [70], or by statistical comparison as per Sudarsono et al. [69] where ANOVAs and post-hoc Tukey analyses were performed (individually) for each of the (many) variable-factor loadings (or 'differential-dimension' loadings). The former provides a holistic assessment and the latter provides an empirical assessment.

It is proposed that the reproducibility of semantic differentials' soundscape dimensions should be a question of whether the reproduction's two dimensions of appraisal (as discussed in section 5.1) occur on the same dimensions as on the in situ dimensions. That is to say: if the in situ soundscape was evaluated most significantly by appraisal – motivation-affordance fit and mediation on the first and second dimensions (as per the MediaCity responses) – then the reproduction should also be evaluated most significantly by appraisal in order to hold validity.

As for open response data types, the developed psycholinguistic analysis presented in this research provides a more accurate means of validity assessment. This method has, so far, been used primarily to illuminate the differing perceptions in various listening scenarios. Since there is no ultimate model or understanding of soundscape perception or audition, only relative merits and cautions can be discussed.

To summarise, the developed psycholinguistic assessment of validity has three stages: the semantic categorisation of open responses; a principal component analysis of semantic category counts; and a comparison of the dimensions in terms of the semantics (variables) maintaining the greatest loading on the same dimension (component). The proposed method adheres comprehensively to the psycholinguistic approach and is not reliant upon the a priori categorisation based on the directive focus of questions in the survey questionnaire. The existing method for the assessment of ecological validity [94] observes only the total semantic occurrences and hence fails to acknowledge the individual cases. The developed psycholinguistic assessment proposed takes into account each individual question response for each individual subject. It does so by observing the inter-semantic relationships thus giving a more detailed representation of the complex cognitive process by which the soundscape is perceived. Furthermore, the categorisation of 'action (verb)' and 'sound object' – in place of 'source identification' – focuses on the key issue, the new ecology. It is the relationships between visual (object), self (subject), action and sound that a soundfield reproduction re-establishes physically. Therefore it is the relationships between visual, self, action and sound that may re-establish cognitively.

5.2.3 Ambisonic Versus Binaural

Existing soundscape research has more frequently employed ambisonic reproduction techniques for laboratory listening tests. The two reproduction techniques were compared based on the research findings.

The semantic differentials' soundscape dimensions showed variation from in situ to ambisonic and binaural. Phase Two's ambisonic semantic differentials' soundscape dimensions better related to the in situ results, whilst the binaural dimensions differed somewhat further.

The binaural differed in that the *motivation-affordance fit* and *mediation* dimensions were manifest in the second and third dimensions. Instead of the primary dimension of evaluation being a part of the subject's appraisal, it was purely objective (with the exception of calming-agitating): the 'timbre and dynamics' dimension from the in situ listening was promoted from third to first dimension. Gentle-harsh and hard-soft moved from the third dimension, *timbre and dynamics*, in situ to the first dimension, entity and effect, for ambisonics. For both ambisonic and binaural reproductions, pleasant-unpleasant moved from the first (in situ) to the second dimensions. However, the ambisonic reproduction responses retained appraisal on the first two – and most significant – dimensions.

The developed psycholinguist assessment of external validity indicated that binaural soundfield reproduction of the MediaCity soundscape held external validity and ambisonic soundfield reproduction of the MediaCity soundscape did not hold external validity (Table 24). It did also reveal variation between the five soundscapes tested (4.1.2.3) suggesting the reproducibility of these semantic dimensions – and hence cognitive process of perception

– is, in part, soundscape-specific. However, the MediaCity soundscape (Table 23) procured the same dimensions as the collective analysis of all five soundscapes (Table 22) therefore the MediaCity soundscape can be deemed typical. To conclude, binaural reproduction generally holds external validity over ambisonic reproduction but in some instances of atypical soundscapes this may not be the case. In the lack of any equivalent counter-argument, it is proposed that binaural should be the default soundfield reproduction technique for the purposes of any laboratory-based subjective soundscape research.

	Five varied soundscapes: Focus group			Five varied soundscapes: Focus group		
	Component			Component		
	1	2	3	1	2	3
% of variance	41.0	26.5	25.0	37.1	25.2	25.0
Subject			**0.997**			**1.000**
Object	**0.859**	0.319		**0.813**	0.242	
Verb	0.145	**0.976**		**0.878**	0.101	
Sound object	**0.938**			0.228	**0.970**	

Table 22 – Developed psycholinguistic assessment of the focus group responses from in situ and ambisonic listening of all five soundscapes (see 4.1.2.3)

	MediaCity: Focus group			MediaCity: Focus group		
	Component			Component		
	1	2	3	1	2	3
% of variance	44.6	26.6	25.2	34.5	25.4	27.5
Subject		-0.152	**0.988**		**0.991**	
Object	**0.893**	0.350		**0.697**	-0.167	0.427
Verb	0.219	**0.956**	-0.174	**0.925**		
Sound object	**0.969**			0.193		**0.955**

Table 23 - Developed psycholinguistic assessment of the focus group responses from in situ and ambisonic listening of the MediaCity soundscape (see 4.1.2.3)

	MediaCity: Public			MediaCity: Participants			MediaCity: Participants		
	Component			Component			Component		
	1	2	3	1	2	3	1	2	3
% of variance	38.1	25.6	25.4	37.4	29.7	25.1	32.1	26.4	25.0
Subject		**0.992**		0.106		**0.994**			**0.993**
Object	**0.805**	0.148	0.288	**0.785**	0.475		**0.781**	0.312	
Verb	0.261		**0.958**	**0.932**	-0.139			**0.942**	0.102
Sound object	**0.897**		0.127		**0.971**		**0.820**	-0.248	

Table 24 - Developed psycholinguistic assessment of the public participant responses from in situ, ambisonic and binaural listening of the MediaCity soundscape (see 4.2.2.2)

5.3 Embodiment and Ventriloquism

5.3.1 Embodiment

It has been established that the listening room binaural reproduction held external validity over the ambisonic reproduction. But why was this so? Explanations are offered around the notion of *embodiment*.

Prior to data collection, the theoretical assumption was made that both listening scenarios – Ambisonic and binaural – are of a soundscape. That is to say the soundfield is embodied as a soundscape. Therefore the Ambisonic listening test was of the listening room soundscape – a soundscape modified by the Ambisonic auralisation of a soundfield recording, rather than simply 'The MediaCity Soundscape'. This assumption did not affect

the methods or results, but merely the language used – a bi-product of the care taken to respect the two distinctly different definitions of ‘soundscape’ and ‘soundfield’.

Binaural listening with earphones is not as straight forward. It is not embodied – or re-embodied – in the same way as the Ambisonic modified, or indeed the in situ, soundscapes. The differing levels of embodiment were mediated by three senses: auditory, tactile, and visual. Firstly, the binaural soundfield does not interact with the new acoustic space; it is not reflected and moulded by the surfaces of the listening room (or ‘original’ space as per Phase Three). Secondly, the binaural soundfield lacks tactility. During in situ and ambisonic listening the earwitness can move their head without sounds following them – sound remains fixed in the space. Listening is an active process – sense and action are integrated with sensory-motor coupling – so without such tactility one does not hear space to the same degree. One hears less. Thirdly, sight plays a significant role in the embodied-ness of the in situ soundscape, indeed open responses referred to sounds in relation to the visual environment (4.2.2.2). The loudspeakers used for the ambisonic reproductions were still visible during the listening tests, despite the lighting having been designed to diminish their prominence (3.2.1.2; 0). Whereas the earphones (and ear defenders) where not visually present during the test.

The more (re-)embodied the reproduction soundscape, the greater its perception differed from that of the in situ soundscape. The binaural reproduction in the listening room

remained dis-embodied: the new place, space-time, and motivations did not fully attach themselves to the sonic.

The phenomenologist Maurice Merleau-Ponty wrote:

> *"If a phenomenon – for example a reflection or a breath of light wind – is only presented to one of my senses, it is a phantom, and it only gets close to a real existence if by chance, it became capable of speaking to my other senses, like the wind when it is made visible in the turbulence of the countryside."* [70]

The ambisonic's sonic phenomenon approaches *real existence* whilst the binaural's sonic phenomenon is a *phantom*. It is the phantom, the figment of imagination, that further engages the earwitness's imagination, and past experience of places heard. This sonic phenomenon is embodied not in the room but the mind's ear.

It may be possible to tentatively extrapolate this theory from soundscape listening to music listening. 'Musical listening' – as defined by Gaver – is concerned with the sound as an entity [30]. However another element of listening to music is how (intangible) elements relate to oneself. If earphones tend to engage the listener's imagination then this element may be enhanced. If loudspeakers tend towards the objectification of sound then they encourage a more strictly 'musical listening'. Listening to music through earphones may improve the

personal holistic experience whilst listening to music through loudspeakers may aid the critical study of a composition.

5.3.2 Ventriloquism

Why did the perception stimulated by the electroacoustic ear correlate – in terms of the developed psycholinguistic analysis results – with that of the Ambisonic reproduction in the listening room?

The 'electroacoustic ear' was explored in Phase Three with the real-time binaural reproduction in situ listening test. The motivation was that analysis of the subjective responses may locate the electroacoustic ear soundscape perception relative to that of in situ soundscape perception and that of binaural (listening room reproduction) soundscape perception. It was hypothesised that the electroacoustic ear would sit somewhere between in situ and binaural reproduction. This would: a) isolate *auralisation* technique and *space* variables – and thus eliminate the *time* and *place* variables – when drawing comparison with in situ listening; b) isolate *time, place, agency* and *expertise* – and thus eliminate *auralisation* technique and *space* – when drawing comparison with binaural reproduction in the listening room; and thus c) weigh the relative significance of the differing soundfield and differing ecology. However, the results suggested a rather different sensory experience altogether.

The electroacoustic ear in situ soundscape perception test related to the study of 'ventriloquism' [95] [96]. Ventriloquism is understood as an audio-visual spatial conflict. The studies of which are generally presented as informing on the processes of intermodal coordination (the *pre*-cognition interaction between the two senses). In the case of an audio-visual spatial conflict (without any sensory degradation) the "apparent location of sounds is shifted toward the simultaneous visual inputs, in spite of [any additional] instructions to ignore the latter. This effect is called the 'visual bias of apparent auditory location'" [35].

This visual bias of apparent auditory location will have been limited in this study due to two factors. Firstly the effect is understood as angle-limited (circa 30°) whereas here the conflict is essentially translational (circa 2m) and therefore subject to a proximity lower-limit. Secondly, and most significantly, the dummy head is fixed and the subject free to move their head, thus – when the subject does so – the conflict is revived. With regards to the developed psycholinguistic analysis results, subjects' responses indicated a separation of action/object semantics and sound object semantics. Given the audio-visual conflicts/resolutions occur prior to the semantic level of cognition; one can conclude that the visual- auditory- space conflict was never fully resolved prior to cognition.

The Ambisonic reproduction responses exhibited the action/object–soundobject separation due to the incongruent ecology within which the soundscape was embodied whilst the electroacoustic ear responses exhibited the separation due to the visual- auditory- space conflict of the congruent ecology.

The real-time binaural in situ semantic differential responses corroborate with the psycholinguistic analysis interpretations. The first dimension was very similar to the binaural first dimension but with the addition of directional-universal. This suggests that it was actually the signal/experience being evaluated rather than the soundscape itself. The fourth simply represented spatiality. The second and third dimensions contained the majority of the subjective semantics and therefore were understood as the two dimensions of subjects' appraisals: the second an appraisal of the sound content/communication and the third an appraisal of the soundscape entity. This soundscape entity appraisal also included social-unsocial and therefore was tied to the sense of place. It was an appraisal of the *embodied* soundscape. So here one observes a splitting of the soundscape entity: on one dimension the binaural experience; and on the other the embodied soundscape entity.

The audio-visual conflict was evidenced not only in the object and soundobject semantic separation of open responses but the splitting of the soundscape entity in the semantic differential responses – sound heard and sound seen.

5.4 The Seven Variables: Revisited

The research methodology was designed following the identification of abstractable variables that may bridge/divide in situ urban soundscape perception and reproduced urban soundfield perception. Here, the role of each variable in the differing perceptions of in situ and reproduction listening is discussed in light of the research findings.

5.4.1 Soundscape

The five soundscapes, in general, stimulated the same cognition of perception: sounds were interpreted as sound objects/sound entities in relation to physical/visual sources, with independent references to activities. The ambisonic auralisation tests showed that the soundscapes themselves influenced the reproducibility of their in situ perceptions. The general trend of the ambisonic reproduction perceptions – as modelled by the developed psycholinguistic analysis – was the interpretation of sounds as either objects in motion or, independently, direct references to sound objects themselves – sounds as entities. However, ambisonic reproductions of two of the five soundfields held ecological validity. In conclusion the ecological validity of ambisonic reproduction was to some extent soundscape-specific.

5.4.2 Place

The in situ real-time binaural reproduction test was unable to control – and thus unable to isolate the effect of – the *place* variable due to the occurrence of other phenomena. The significance of place manifested in subject-place relationships. The focus group's responses, specifically the 'subject' references, were influenced by the subjects' familiarity with place (p110). The semantic differentials' soundscape dimensions from the members of the public at MediaCity – i.e. the 'users' – utilised openness-related descriptors ('anechoic' and 'directional') in the appraisal of the soundscape (p165). It was proposed that such descriptors represented the place-affordances playing a role in the appraisal process of the 'users'. Although the supporting results from this research were case-specific (i.e. people working at MediaCity), place-affordances were also observable in the semantic differentials' soundscape dimensions of an existing study on urban parks.

5.4.3 Motivation and Expertise

Motivations and expertise were not delineated in the experimentation. However, the role of motivation was apparent on the motivation-affordance fit dimension of the generalised soundscape dimensions model. Participants from members of the public at MediaCity were predominantly office workers of the immediate vicinity taking lunch outdoors. Their motivation to lunch outdoors identified the aforementioned place-affordances in their

motivation-affordance fit appraisals, observable on the first of the semantic differentials' soundscape dimensions.

Motivations for one's presence separated the 'users' and 'non-users'. The inherent test conditions of the focus group in situ test at MediaCity meant the focus group subjects lacked the agency of the 'users'. During the reproduction tests the subjects were unable to adopt such nuanced motivations (opposed to, say, the more universal motivations for visiting an urban park).

5.4.4 Space and Time

Space played a role with the real-time binaural reproduction in situ – where soundscape, place, memory, and time were kept unchanged from that of the public in situ responses. Given the small degree of spatial displacement an audio-visual spatial conflict arose. This would not have been the case if the reproduction had been indoors in, say, a foyer lacking visual cues. The result was a cognitive process of perception comparable to that of the ambisonic reproduction – perception of a disembodied soundscape. Furthermore the semantic differential analysis revealed that the sonic experience itself was the primary evaluation, and independent of evaluations of the soundscape in relation to place.

The influence of time was not isolated in the experimentation – a real-time laboratory reproduction would be required.

5.4.5 Auralisation

Ambisonic and binaural reproductions were compared in a laboratory scenario. According to the developed psycholinguistic analysis of binaural reproduction of the MediaCity soundfield stimulated a cognitive process of perception closer to that of in situ listening than that stimulated by the ambisonic reproduction. Binaural reproduction of the MediaCity soundfield held external validity over the ambisonic reproduction. The external (in)validity of ambisonic reproduction was found to be somewhat soundscape-specific and therefore one must be cautious that this may also be the case with binaural reproduction. However, perceptions of the MediaCity soundscape were representative of all five soundscape-specific perceptions (when analysed collectively). Therefore it was concluded that in general binaural reproduction holds external validity over ambisonic reproduction, but the soundscape-specific nature of the reproducibility of soundscape perceptions suggests that certain soundscapes may provide exceptions to the rule.

5.5 Methodological Review

The key novelties of this research have been identified. The research methodology is the first to:

1. Assess the external validity of binaural soundfield reproduction;
2. Analyse the perception of real-time binaural reproduction in situ;

3. Employ a semantic differential analysis and semantic categorisation analysis in parallel;
4. Isolate the effect of ‘test conditions’ (agency and expertise); and hence first to…
5. Develop a method for assessing ecological validity; (that also does…)
6. Improve upon the accuracy of the external validity assessment proposed in ‘Ecological Validity of Soundscape Reproduction’ [70] with the development of a psycholinguistic analysis that takes account of individual cases;

CHAPTER SIX

6 CONCLUSIONS

6.1 Overview

Soundfield reproduction has a vast array of applications, yet the ecological validity – and external validity – has been little researched beyond the narrow limits of the physical approach.

This research employed a **mixed model approach** grounded in the pragmatist research paradigm. The design was positioned as 'partially mixed research' on the research continuum owing to a post-positivist dominant paradigm emphasis decision. The methodology is experimental and theoretical, with a strong experimental bias. Within the experimental there was a quantitative bias.

The **focus** was the difference between in situ soundscape perception and reproduction soundscape perception – beyond the acoustic. The methodology was formulated from the identification of seven variables responsible for the differing perception between in situ and

reproduction listening: *auralisation* technique; *soundscape*; *place* affordances; *time*; *space*; and memory – subject *agency* and *expertise*.

The experimentation assessed these variables with several soundscape perception scenarios: electroacoustic reproduction compared to 'real'; ambisonic compared to binaural; binaural live in situ compared to binaural later elsewhere; different soundscapes; and a pair subject types, namely the public and colleagues/students.

Qualitative and quantitative data was collected in the form of open question responses and semantic differential responses. Three empirical methods of analysis were employed – two existing and one novel:

1. The extraction of 'soundscape dimensions' from the principal component analysis of 18 semantic differentials (as per Kang et al. [59]);
2. The semantic categorisation of words and phrases from the open responses into 'subject', 'object' and 'source identification' (as per Guastavino et al. [70]); and
3. The *developed psycholinguistic analysis* (novel).

The ***developed psycholinguistic analysis*** extracts 'cognitive dimensions' from the principal component analysis of the occurrences of the semantic categories 'subject', object', 'action' and 'sound object'. The analysis adheres comprehensively to the psycholinguistic approach and is not reliant upon the a priori categorisation based on the focus of directive questions

of the survey questionnaire; it takes into account each individual case (every response for every subject); and the additional categorisation of 'action (verb)' and 'sound object' focuses on the key issue – the new ecology. It is the relationships between object, self, action and sound that soundfield reproduction physically re-establishes. Therefore it is the relationships between visual, self, action and sound that may re-establish cognitively. As an assessment of the external/ecological **validity of soundfield reproduction**, the in situ and reproduction cognitive dimensions produced by the developed psycholinguistic approach can be compared in terms of whether the semantics maintain the greatest loading on the same dimension.

There were **three phases of experimentation**. Phase One compared different soundscapes and the focus group subjects' differing perceptions of in situ and reproduction soundscapes. Phase Two compared the Ambisonic and binaural auralisation techniques and gained users' in situ perceptions to contrast the focus group's agency. Phase Three explored the isolation of 'the electroacoustic ear' – in place, in time, but out of space – with real-time binaural in situ listening.

In **Phase One** the Kings Arms ambisonic reproduction was evaluated by the focus group as less social on the semantic differentials than the in situ soundscape – due to the absence of the accompanying focus group members in the reproduction soundscape. The semantic

differential analysis showed that the ambisonic reproduction was evaluated less in terms of appraisal and more in terms of objective descriptors.

The semantic categorisation of the open responses revealed that the five soundscapes were evaluated in different ways. 'Subject' references were significantly linked to person-place relationship as well as estimated number of visits annually.

The assimilation of an existing method of ecological validity assessment (as per Guastavino et al. [70]) indicated 'ecological validity' with greater significance than indicated in the existing study. Therefore subject memory – *agency* and *expertise* – had an influence on the external validity of soundfield reproduction.

The developed psycholinguistic analysis showed differing perception of in situ and ambisonic soundfields. Only the Triangle and Kings Arms soundscapes held ecological validity in the ambisonic reproductions. A collective analysis of all five soundscapes indicated an ecological invalidity of ambisonic soundfield reproduction. This can be deemed the general case, whilst cautious of the soundscape-specific nature of ecological validity.

In **Phase Two** the semantic differential analysis gave the in situ dimensions: spatiality/tempo evaluation (22%); content evaluation (21%); timbre and dynamics (14%); spatiality (13%). The ambisonic dimensions were: sound entity evaluation (23%); sound

content evaluation (16%); spatiality, timbre and dynamics (16%); near-far phenomenon (12%). The binaural dimensions were: sound entity (26%); sound content evaluation (18%); entity/content evaluation (16%); spatiality (13%). It was proposed that a PCA of semantic differential responses for any in situ listening tests will produce two dimensions that represent the ***motivation-affordance fit*** and ***mediation*** identified in the psychological theories of appraisal. The notion motivation-affordance fit not only transcends the theory of expectation, but should be considered central in the evaluation of soundscapes.

The ambisonic reproduction, not the binaural reproduction, of the MediaCity soundfield stimulated similar dimensions to that of the in situ soundscape. The first two ambisonic dimensions retained the appraisal differentials (e.g. pleasant-unpleasant) whilst the binaural *motivation-affordance fit* and *mediation* dimensions were manifest in the second and third dimensions. Instead of the primary dimension of binaural soundscape evaluation being a part of the subject's appraisal, it was predominantly objective.

The in situ, ambisonic and binaural open responses, once categorised by 'subject', object', 'action' and 'sound object' references, were compared: a Tukey (HSD) analysis, in conjunction with a one-way ANOVA, showed that the number of 'sound object' references in the subjects' open responses was significantly greater for Ambisonic listening than in situ listening.

The developed psycholinguistic analysis gave the ambisonic dimensions: verb and object joint references; sound object references; subject references. The binaural dimensions were: sound object and object joint references; verb references; subject references. The category counts and developed psycholinguistic method both indicated in agreement that binaural reproduction held external validity over ambisonic reproduction.

However, the Phase One results suggested the reproducibility of these semantic dimensions – and hence cognitive process of perception – is, in part, soundscape-specific. Since the MediaCity soundscape (Table 23) procured the same dimensions as the collective analysis of all five soundscapes (Table 22), the MediaCity soundscape can be deemed typical. Therefore it was concluded that binaural reproduction generally holds external validity over ambisonic reproduction yet one should be cautious of the soundscape-specific nature of the reproducibility of cognitive dimensions. It is proposed that – in the absence of a study specific assessment of external validity – binaural soundfield reproduction should be used for the purposes of any laboratory-based subjective soundscape research.

Why binaural reproduction – and not ambisonic reproduction – stimulated a perception closer to that of 'reality' was discussed from a phenomenological perspective. The more (re-)embodied the reproduction soundscape, the greater its perception differes from that of an in situ soundscape. The binaural reproduction in the listening room remained dis-embodied:

the new place, space-time, and motivations did not fully attach themselves to the sonic. The sonic phenomenon was re-embodied not in the room but the mind's ear.

In **Phase Three** the real-time binaural in situ listening responses produced four dimensions interpreted as: experience evaluation [of the binaural signal] (22%); sound content evaluation (21%); [embodied] soundscape entity evaluation (15%); spatiality (11%). The developed psycholinguistic analysis of real-time in situ binaural responses gave three dimensions very similar to that of the Ambisonic responses. The sound object references were independent of both object and verb references. Object and verb references co-varied. The Ambisonic reproduction responses exhibited the action/object–soundobject separation due to the incongruent ecology within which the soundscape was embodied whilst the electroacoustic ear responses exhibited the separation due to the visual-auditory- space conflict of the congruent ecology.

6.2 Summary of Findings

6.2.1 Corroborative

- Semantic differentials' soundscape dimensions from in situ soundscape perception can be adequately reproduced with a soundfield recording reproduced in a listening

room with an eight channel ambisonic auralisation. This is in agreement with a study by Davies et al. [70].

6.2.2 Novel

- Two methods of analysing soundscape perception were compared.
    - The semantic differentials' soundscape dimensions implied subtle differences between in situ and ambisonic listening.
    - The semantic category counts from open responses implied no difference in the cognitive processes of perception.
- The developed psycholinguistic method is not influenced by a priori criteria, it observes the distinction of object-orientated and action-orientated sound source identification, and acknowledges individual cases of response by not being reliant upon a generalised summation. The method observed nuances (motivation, expertise and place) that existing methods did not.
- Semantic differentials' soundscape dimensions and the developed psycholinguistic approach were repurposed as methods for assessing the ecological/external validity of urban soundfield reproduction.
    - The semantic differentials' first two soundscape dimensions related to – and can be interpreted with – appraisal theory's 'motivation-affordance fit' and 'mediation'. A reproductions replication of a 'motivation-affordance fit' first

dimension and ‘mediation’ second dimension infers an assessment of validity.

    - The developed psycholinguistic approach produces three dimensions of four semantic categories – ‘subject’, ‘object’, ‘action’ and ‘sound-object’. A reproductions replication of the in situ dimensions – ‘object/sound-object’, ‘action’ and ‘subject’ infers an assessment of validity.

- The methodology was able to identify the influence of various factors on the differing perceptions of in situ and reproduction soundscapes.
    - The nature of a soundscape has an influence on how its reproduction will be perceived therefore the validity of soundfield reproduction is, to an extent, soundscape-specific.
    - Subject motivation and expertise can play a significant role in the differing perceptions of in situ and reproduction soundscapes.
    - In general binaural reproduction holds external validity over ambisonic reproduction.
- The isolation of the ‘electroacoustic ear’ with real-time binaural reproduction in situ stimulated an audio-visual conflict. The subjects’ process of perception was – according to the developed psycholinguistic analysis – closer to that of ambisonic reproductions than binaural. It was therefore a less natural process of perception than laboratory-based binaural reproduction.

6.2.3 Proposals

- Motivation-affordance fit is a key part of soundscape appraisal and therefore should be a key feature of the assessment and design of urban soundscapes. The dimension of motivation-affordance fit within semantic differentials' soundscape dimensions is means of identifying specific soundscape characteristics and features that play a prominent role in the soundscape perception and appraisal for a given space. This approach can in turn inform the choice of sonic affordances in the processes of urban design and planning.
- The perceived embodiment of a reproduction soundscape can be deemed to differentiate ambisonic and binaural auralisation techniques. The use of loudspeakers re-embodies a recorded soundscape whilst the use of earphones maintains the disembodiment of a soundscape recording and thus the soundscape is re-embodied in the mind's ear by engaging memory of past experience.
- Binaural reproduction should be the default auralisation technique when evaluating soundscape perceptions in laboratory conditions.

6.3 Future Work and Recommendations

- The influences of the earwitness's expertise and motivation could be delineated with a small amount of further investigation. Conducting the same survey questionnaire with a non-expert focus group escorted to MediaCity – i.e. a group of

non-expert non-users – would isolate subjects' motivation in comparison with the results of the group of expert non-users and the non-expert users. This would provide a more detailed dissection of subjective 'test-conditions'. Furthermore, the use of the 18 semantic differentials would enable the proposed soundscape appraisal's dimension theory to be further validated.

- Raimbault investigated the appropriateness of the descriptors constituting semantic scales by identifying polarised ('U') response frequencies – indicative of differing interpretations [67].The near-far phenomenon (whereby near-far responses strongly co-varied with otherwise non-spatial attributes such as social-unsocial in a polarised fashion) could be used to separate the responses into two data sets to analysed individually. The same separated analysis could be performed around the varied-simple responses. Previous research found that 'varied' responses were indicative of a subject employing descriptive listening and 'simple' of holistic listening. Separating the responses by listening state would offer understanding of how such listening states manifest in the psycholinguistic interpretations of the respective open responses. This could in turn begin to shed light on if and how people can change their auditory strategy. For example, placing the subject 'in role' may offer a means of prescribing an auditory strategy and thus reducing subjective variation.

- Since the validity of reproduction was found to be soundscape-specific, it is more important not to assume binaural reproduction will always offer external validity for laboratory-based soundscape evaluation tests. Therefore a quicker and simpler method of assessment would be of use. One potential avenue for investigation is a categorical questionnaire. Soundscape orientated questions would have various options made available – each set of options would offer words or phrases with similar (more obvious) meanings but conceptually different – for example “lots of conversation” or “lots of people talking”. The options would offer a range of object-, action-, subject- and sound entity- centred response options (potentially drawn from the array of pre-categorised sound descriptors offered by Torben Holm Pedersen [69]). The sematic preferences would then be compared for in situ and reproduction listening. Conducting the developed psycholinguistic analysis in parallel would enable validation.

- Composers and artists working with sound have at their disposal handbooks on acoustic effects and processing techniques. However there is no equivalent for *perceptual* effects. This is a study that would benefit from a broad review of soundscape literature but the greatest contributions from this thesis would be the cataloguing of the perceptual effects of differing reproduction techniques. The descriptions could work with themes such as embodiment, ventriloquism and memory-sound-place.

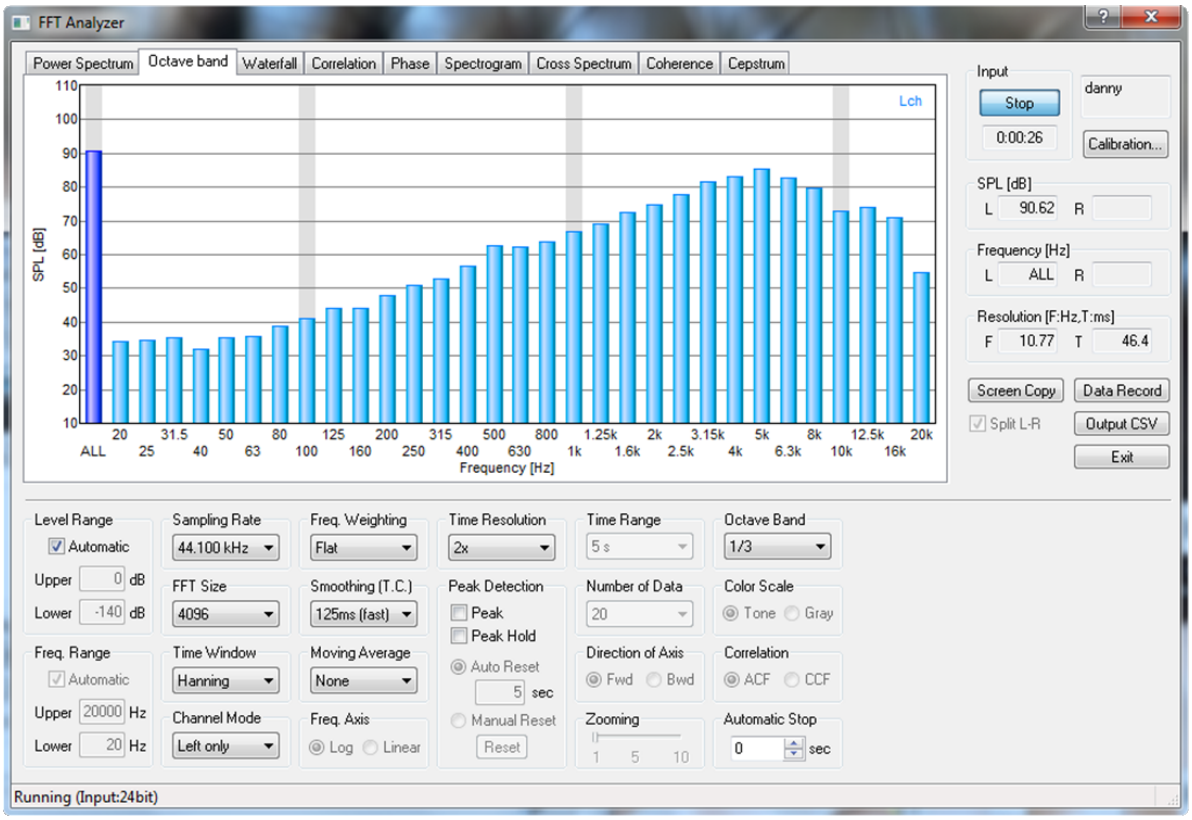
END MATTER

7 APPENDICES

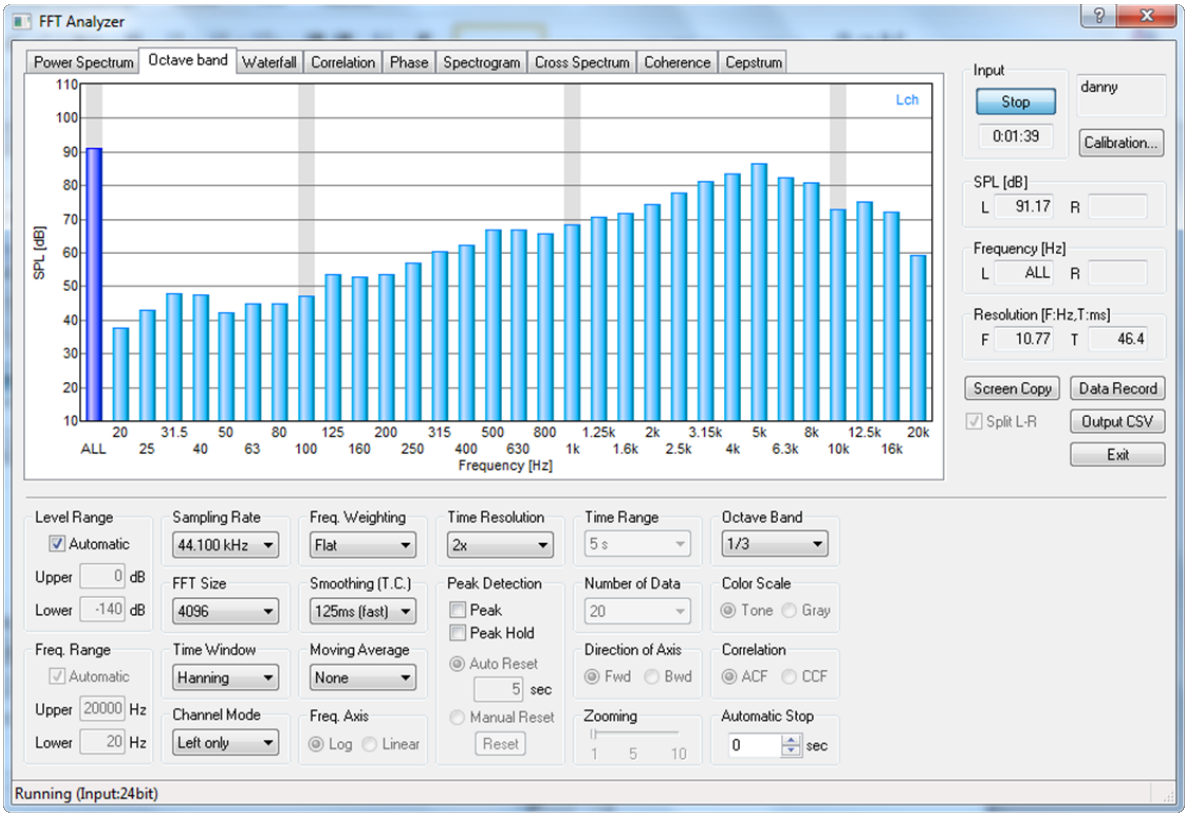
7.1 Binaural Reproduction

7.1.1 Frequency Response

7.1.1.1 WigWare software screenshot showing the frequency response of the earphones in the Kevlar dummy head without ear defenders.



7.1.1.2 WigWare software screenshot showing the frequency response of the earphones in the Kevlar dummy head with ear defenders.



7.2 Questionnaires

7.2.1 Questionnaire A

Questionnaire: Location 5

Regarding soundscape…

Describe the character and feel of the elements that appear more prominent.

Describe the character and feel of the background noise.

What do you like about this soundscape? ..and why?

What would you change about this soundscape? …and why?

What element has most effect on you? Describe. *(Is it exciting, for example?)*

Please rate the soundscape (by circling the relevant number on these scales):

Unpleasant	1	2	3	4	5	6	7	8	9	10	Pleasant
Uneventful	1	2	3	4	5	6	7	8	9	10	Eventful
Social	1	2	3	4	5	6	7	8	9	10	Unsocial
Boring	1	2	3	4	5	6	7	8	9	10	Chaotic
Happy	1	2	3	4	5	6	7	8	9	10	Sad
Noisy	1	2	3	4	5	6	7	8	9	10	Quiet
Exciting	1	2	3	4	5	6	7	8	9	10	Tranquil

That's it! Thank you for taking part.

7.2.2 Questionnaire B

A Soundscape Questionnaire

Regarding soundscape…

Describe the character and feel of the elements that appear more prominent.

Describe the character and feel of the background noise.

What do you like about this soundscape? ..and why?

What would you change about this soundscape? …and why?

What element has most effect on you? Describe. *(Is it exciting, for example?)*

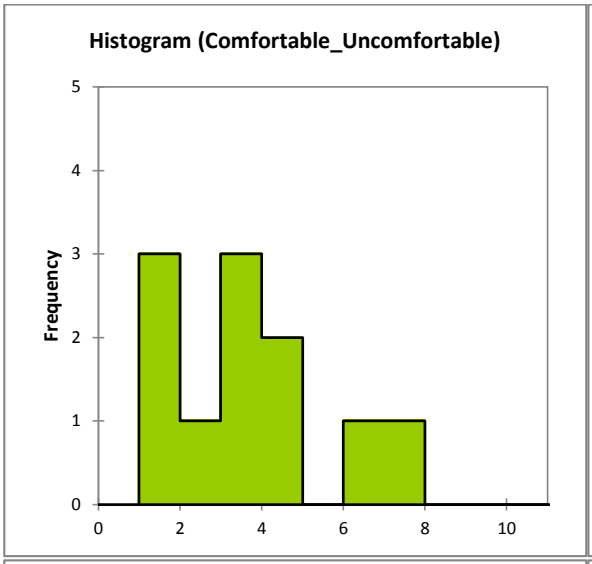
Please rate the soundscape (by circling the relevant number on these scales):

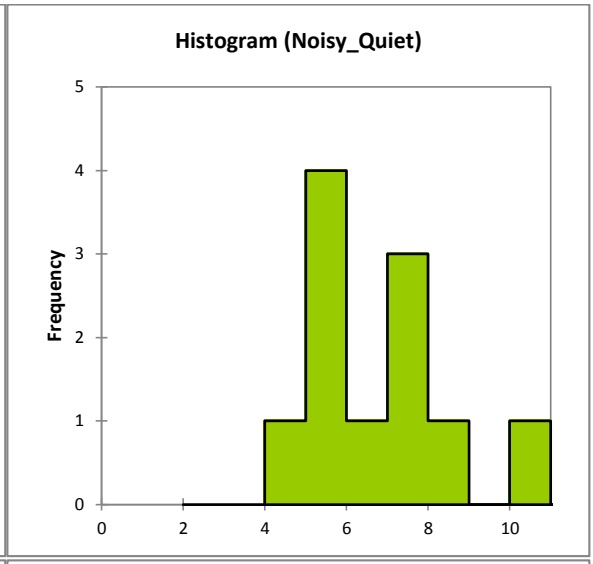
Comfortable	1	2	3	4	5	6	7	8	9	10	Uncomfortable
Noisy	1	2	3	4	5	6	7	8	9	10	Quiet
Unpleasant	1	2	3	4	5	6	7	8	9	10	Pleasant
Interesting	1	2	3	4	5	6	7	8	9	10	Boring
Natural	1	2	3	4	5	6	7	8	9	10	Artificial
Like	1	2	3	4	5	6	7	8	9	10	Dislike
Gentle	1	2	3	4	5	6	7	8	9	10	Harsh
Hard	1	2	3	4	5	6	7	8	9	10	Soft
Fast	1	2	3	4	5	6	7	8	9	10	Slow
Sharp	1	2	3	4	5	6	7	8	9	10	Flat
Directional	1	2	3	4	5	6	7	8	9	10	Universal
Varied	1	2	3	4	5	6	7	8	9	10	Simple
Reverberant	1	2	3	4	5	6	7	8	9	10	Anechoic
Far	1	2	3	4	5	6	7	8	9	10	Near
Social	1	2	3	4	5	6	7	8	9	10	Unsocial
Meaningful	1	2	3	4	5	6	7	8	9	10	Meaningless
Calming	1	2	3	4	5	6	7	8	9	10	Agitating
Smooth	1	2	3	4	5	6	7	8	9	10	Rough

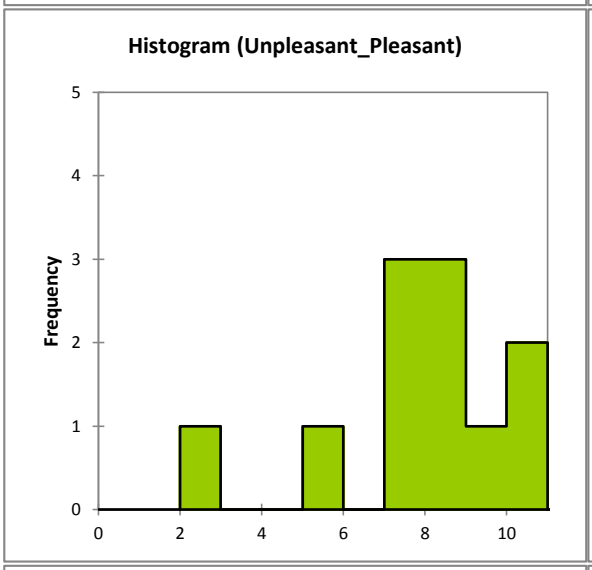
That's it! Thank you for taking part.

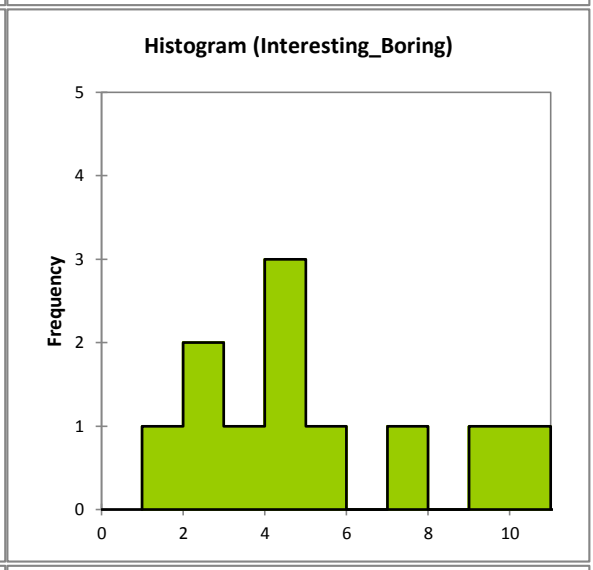
7.3 Histograms from Phase Two Semantic Scales Responses

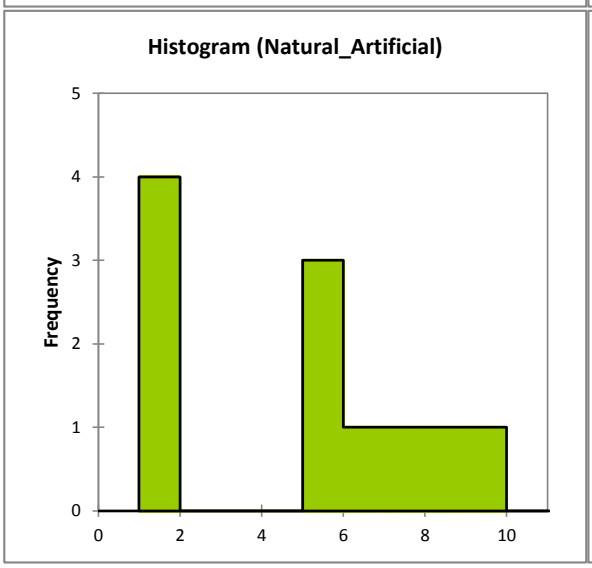
7.3.1 In Situ

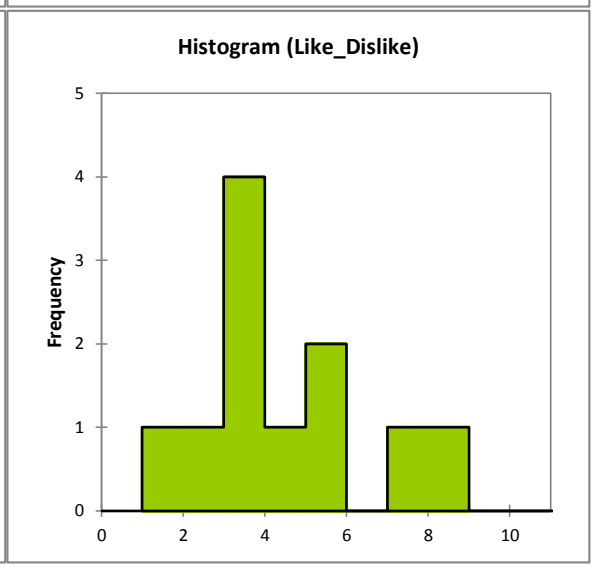

Histogram (Comfortable_Uncomfortable)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

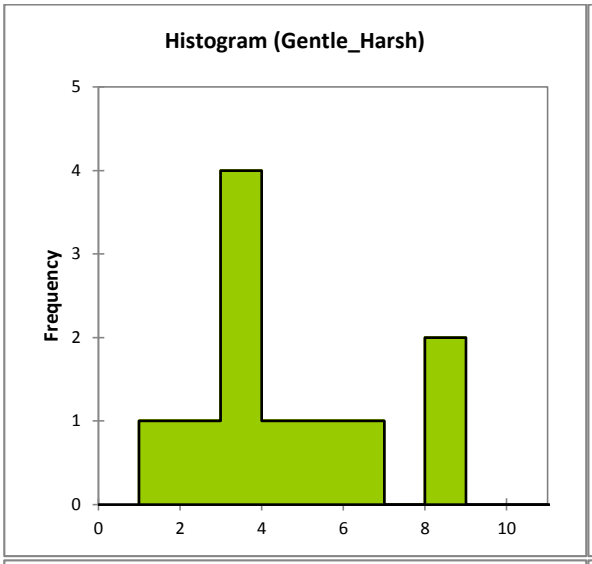

Histogram (Noisy_Quiet)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

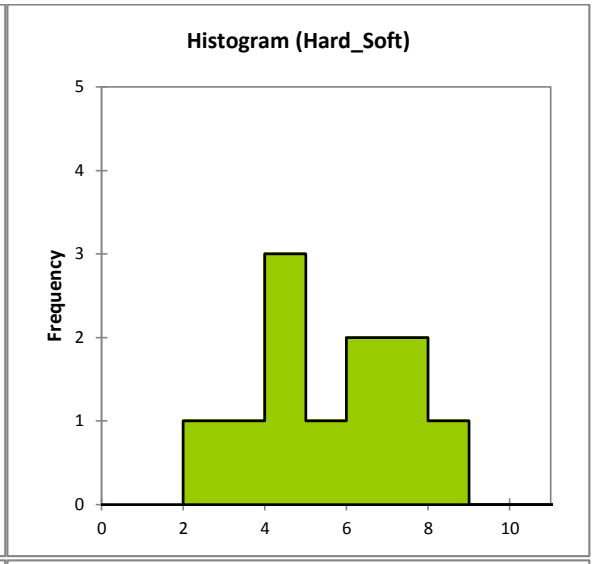

Histogram (Unpleasant_Pleasant)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

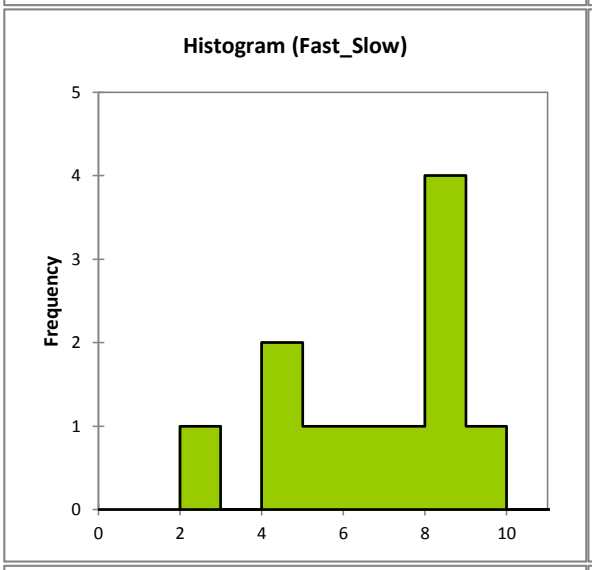

Histogram (Interesting_Boring)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

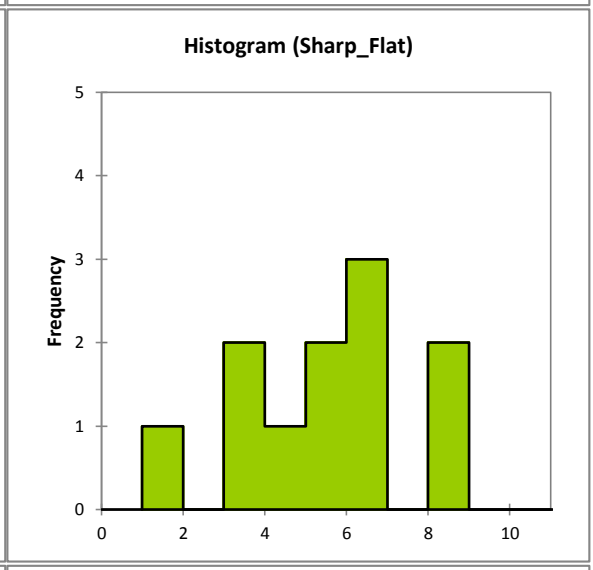

Histogram (Natural_Artificial)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

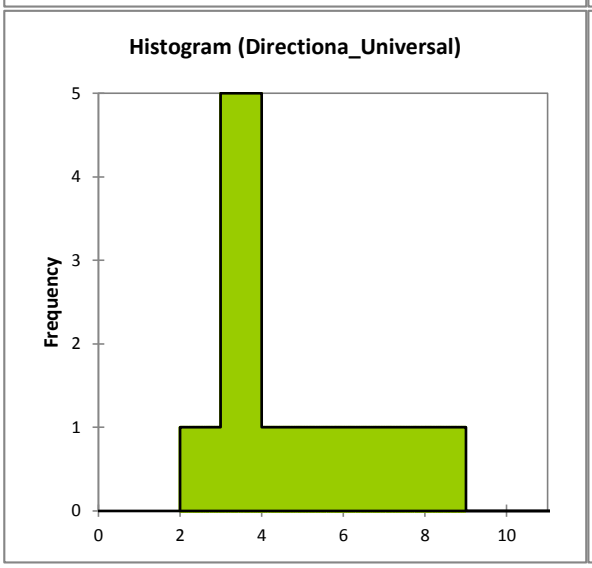

Histogram (Like_Dislike)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

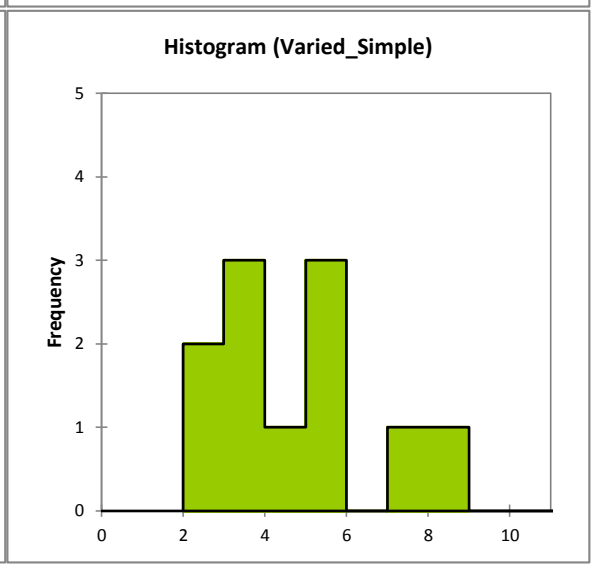

Histogram (Gentle_Harsh)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

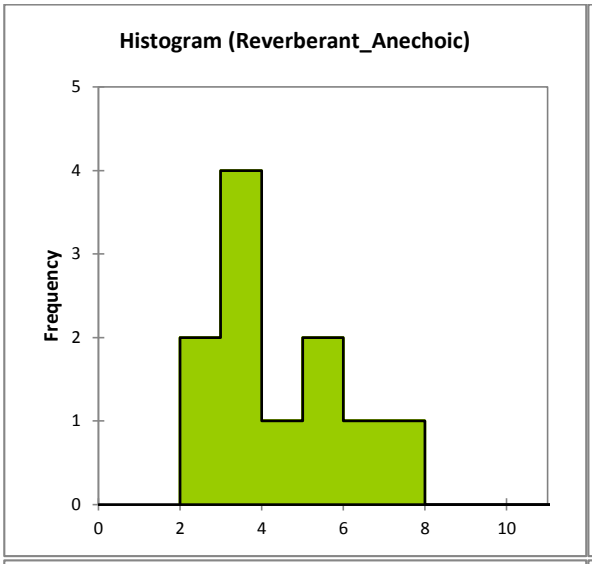

Histogram (Hard_Soft)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

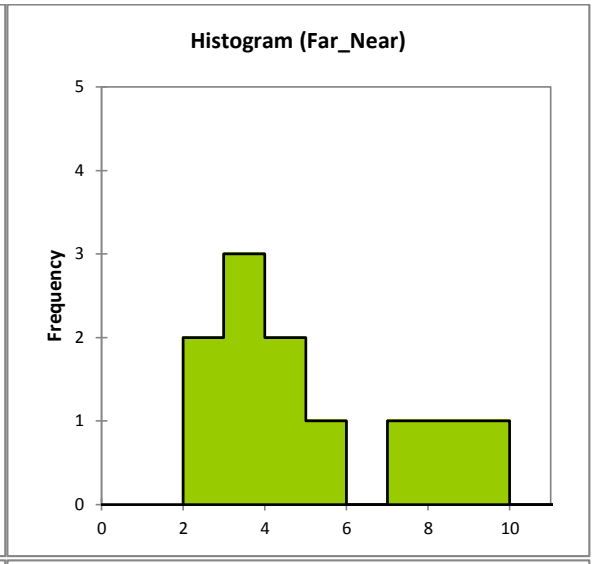

Histogram (Fast_Slow)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

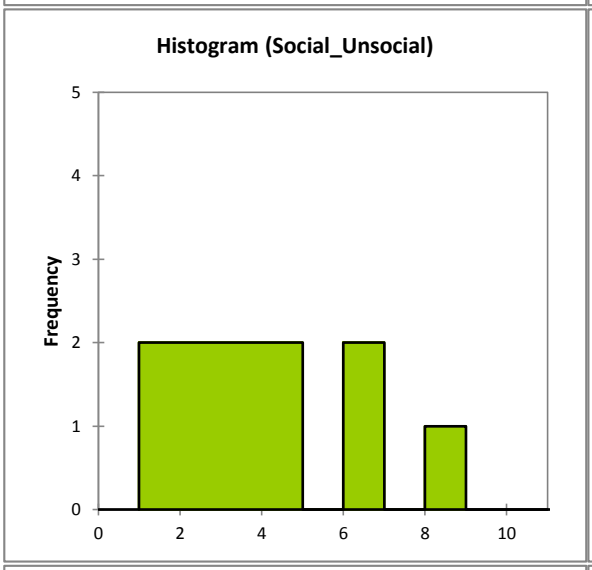

Histogram (Sharp_Flat)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

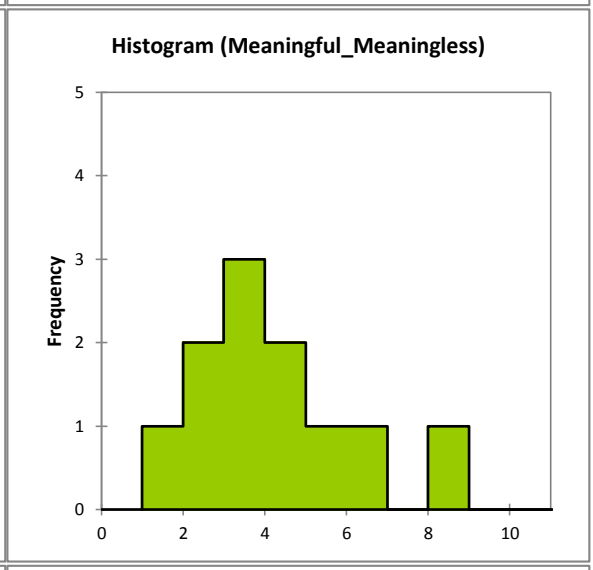

Histogram (Directiona_Universal)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

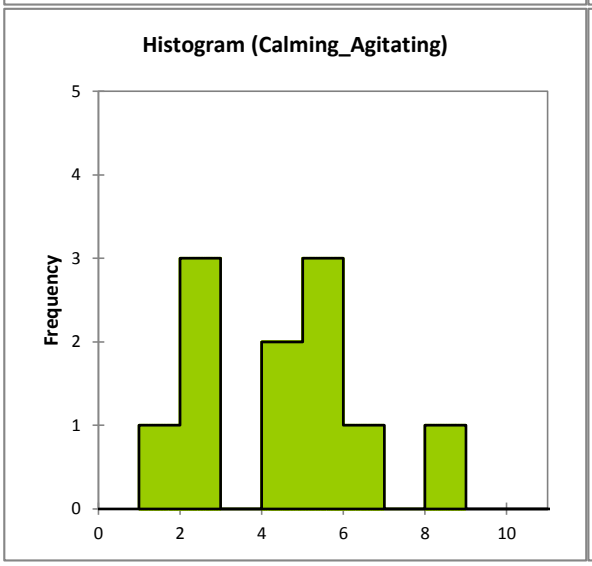

Histogram (Varied_Simple)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

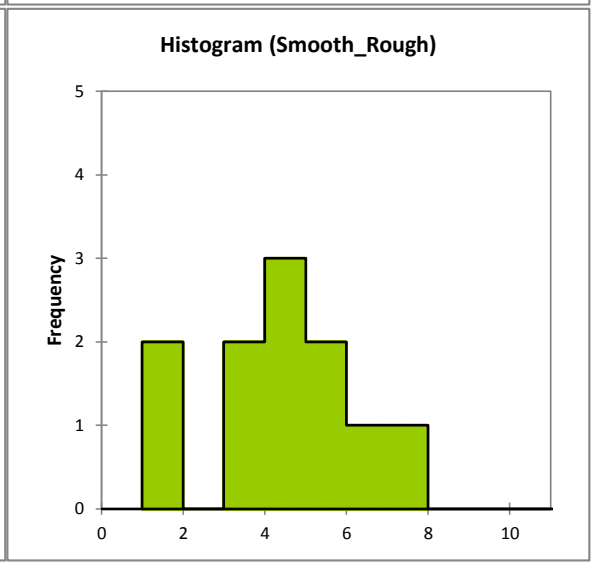

Histogram (Reverberant_Anechoic)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10


Histogram (Far_Near)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

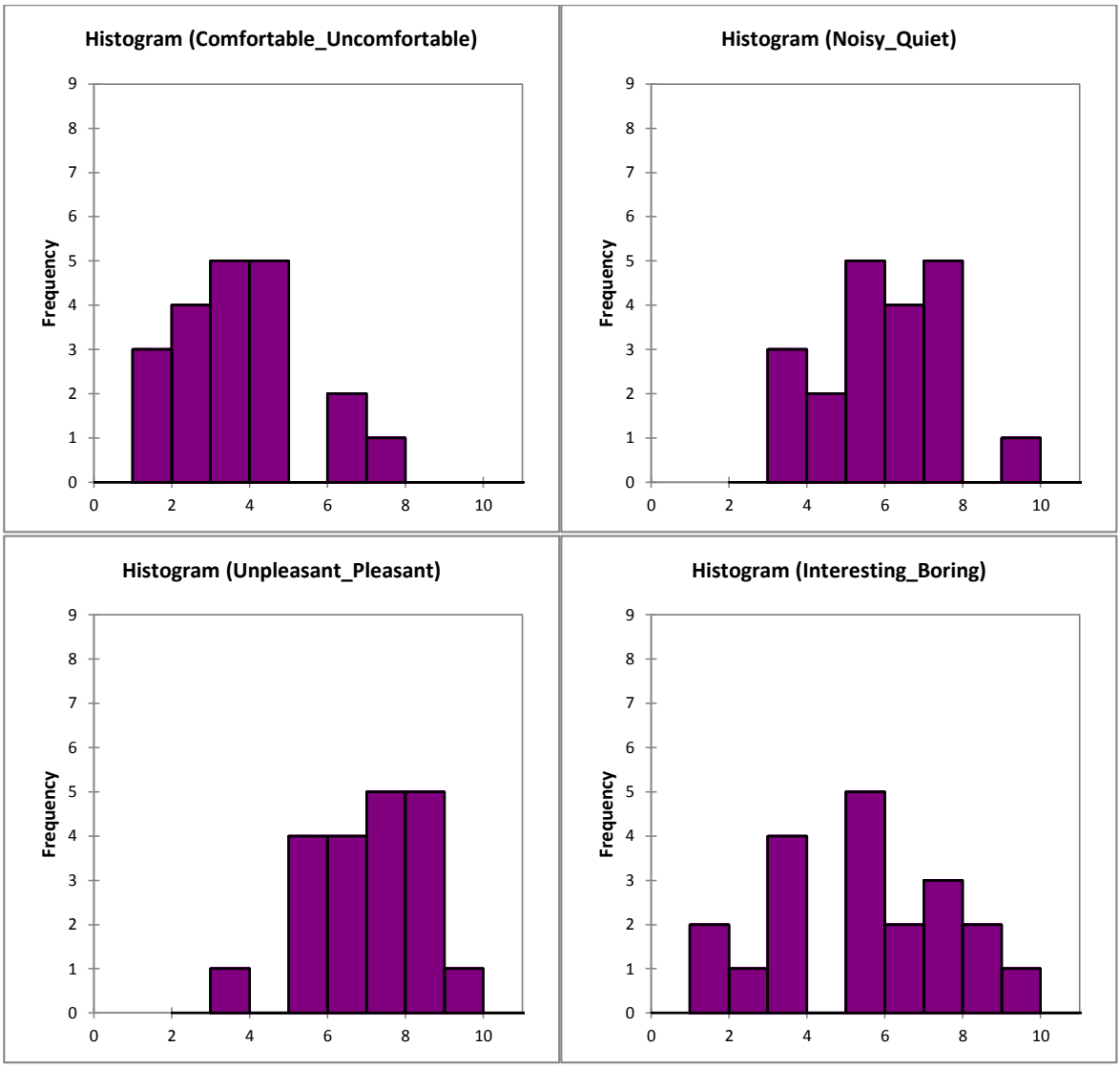

Histogram (Social_Unsocial)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

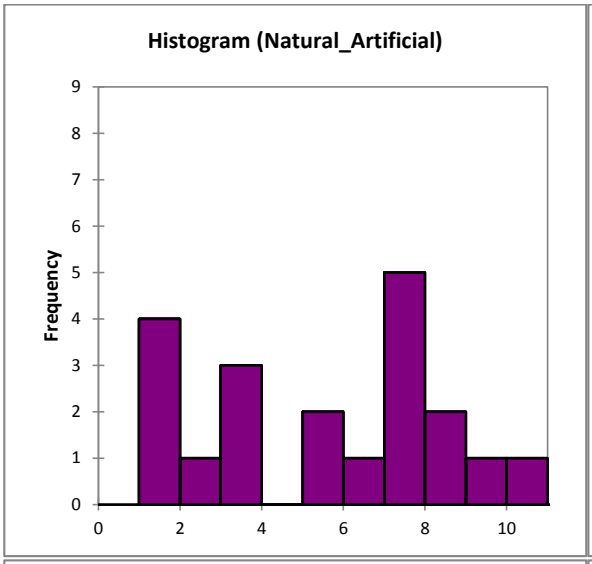

Histogram (Meaningful_Meaningless)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

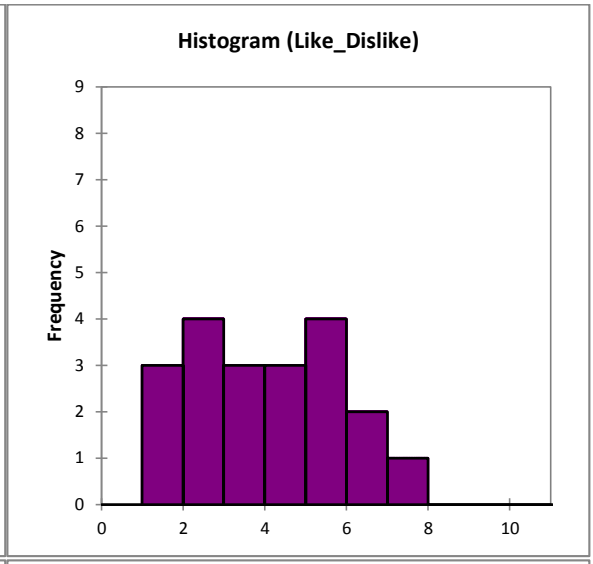

Histogram (Calming_Agitating)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

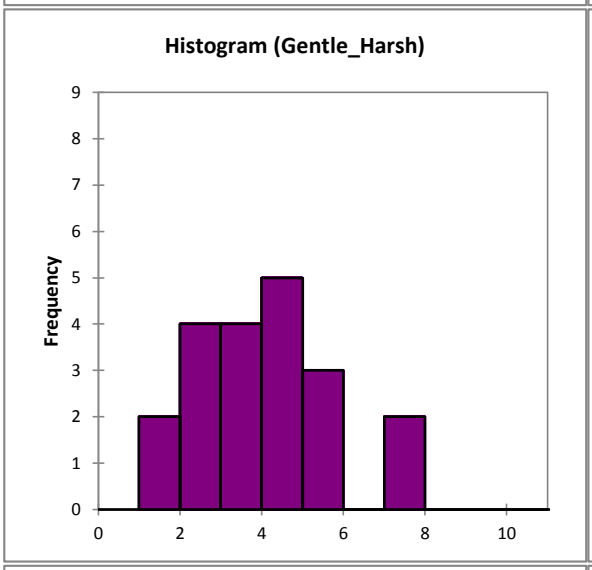

Histogram (Smooth_Rough)
Frequency
0
1
2
3
4
5
0
2
4
6
8
10

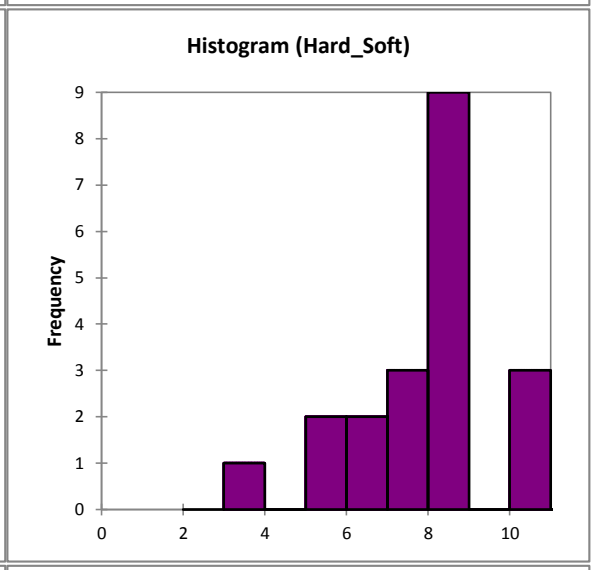
7.3.2 Ambisonic

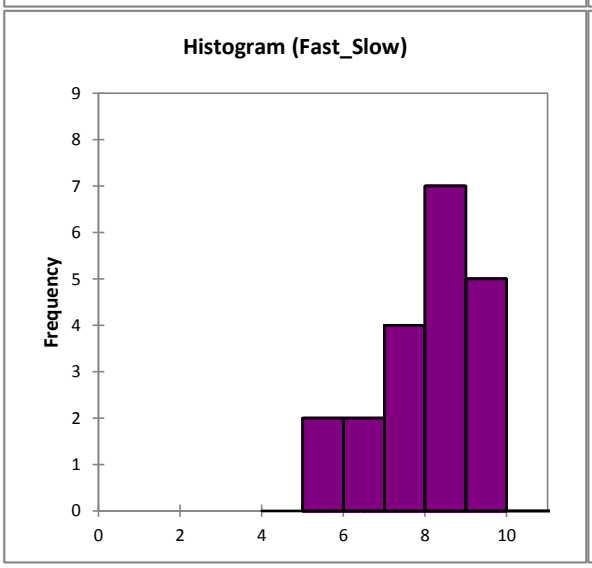


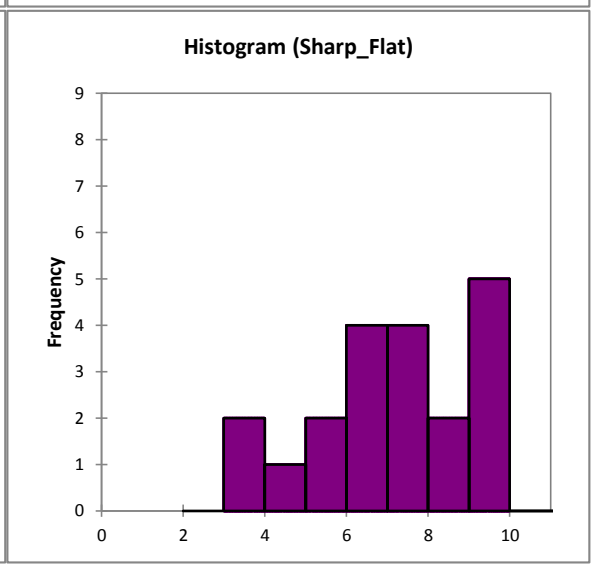

Histogram (Natural_Artificial)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10

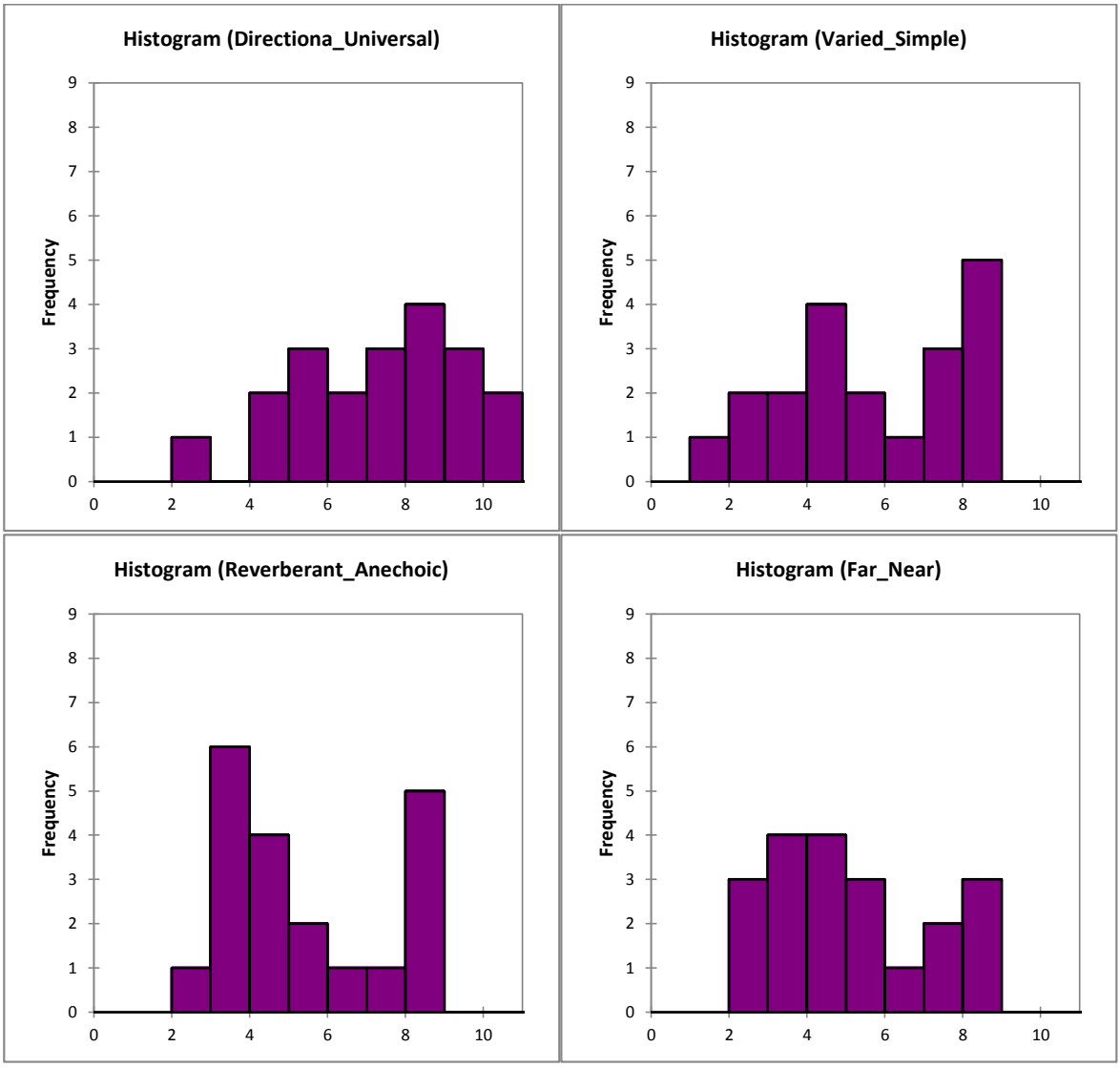

Histogram (Like_Dislike)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10

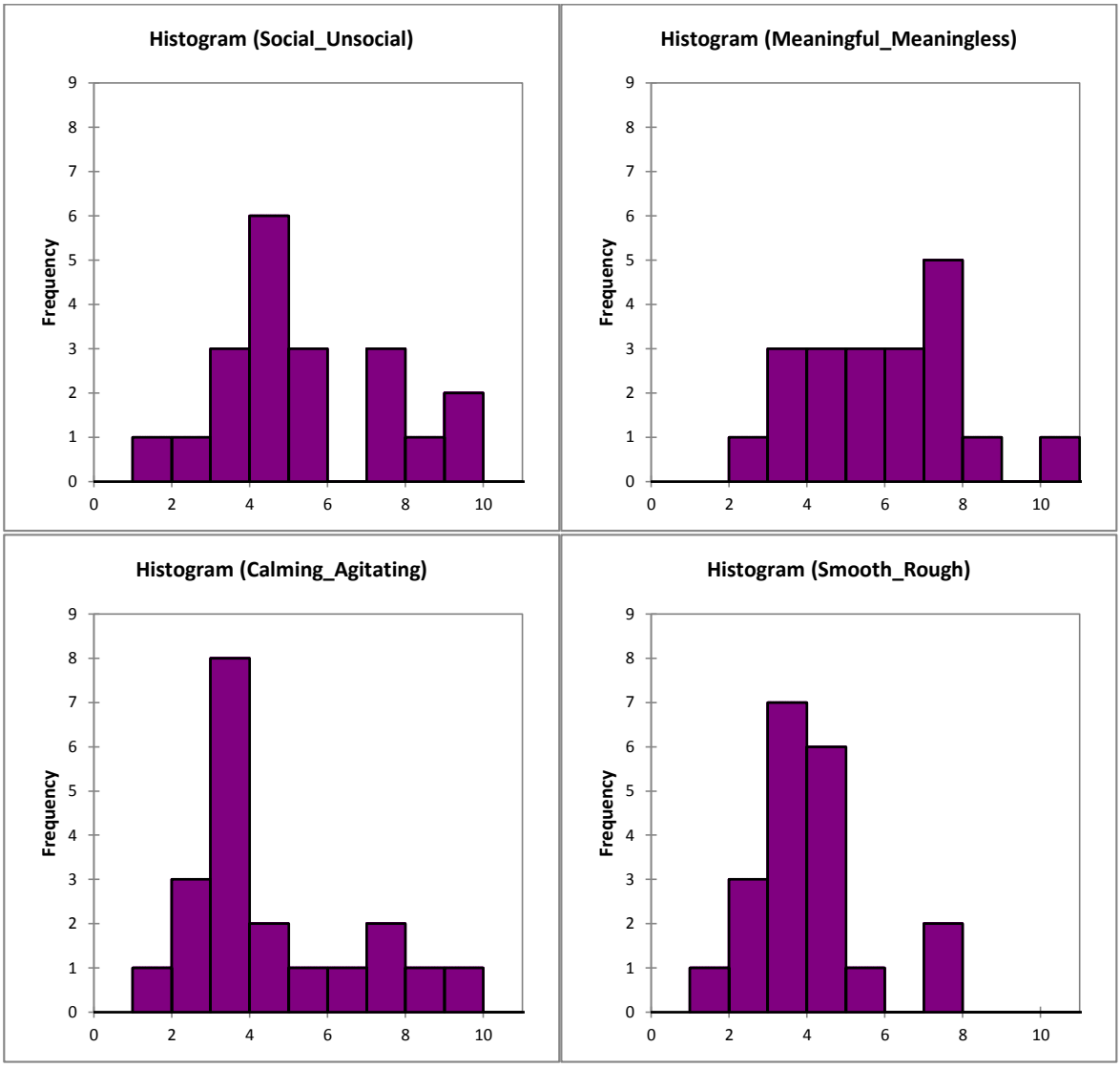

Histogram (Gentle_Harsh)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10


Histogram (Hard_Soft)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10

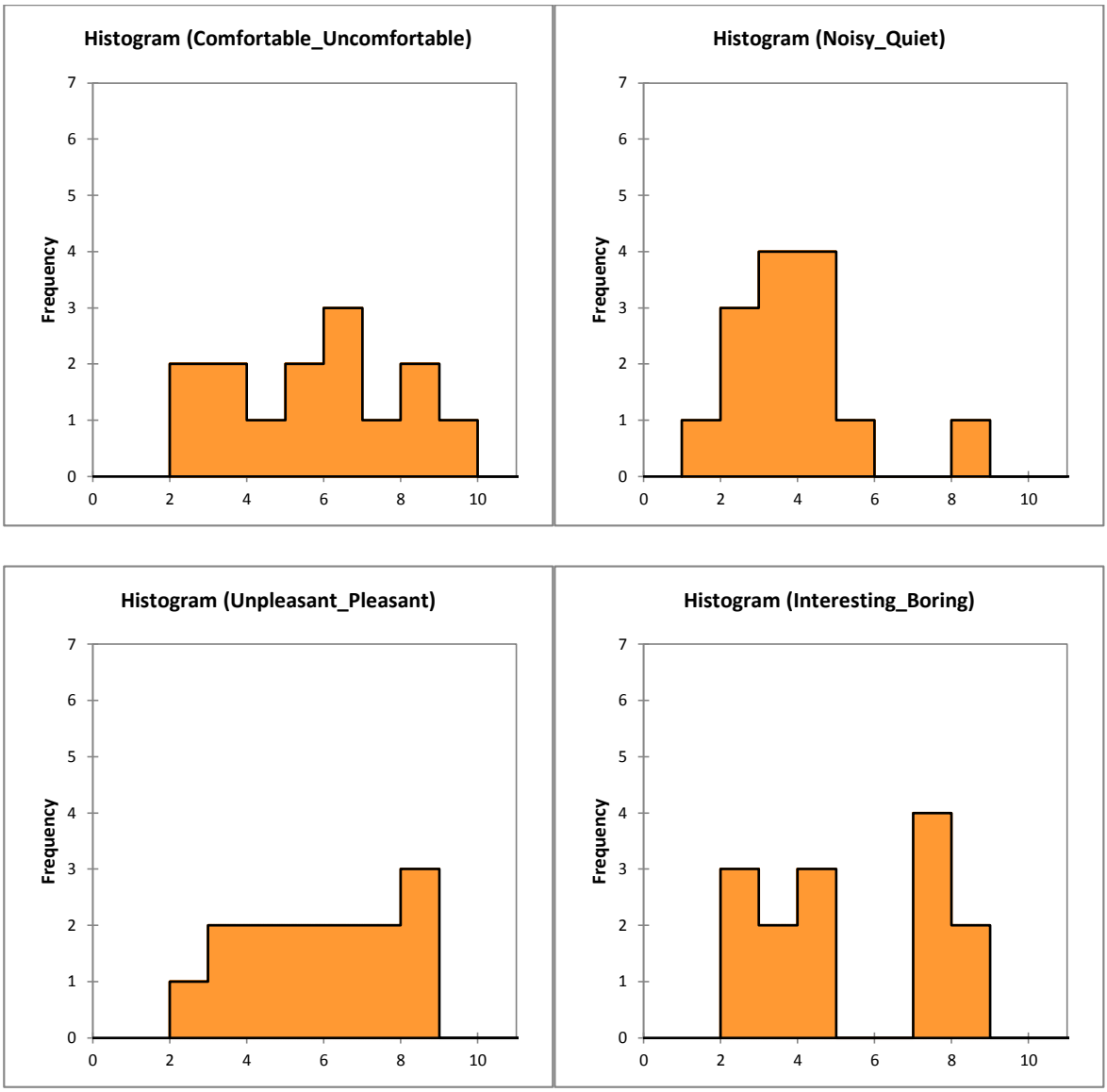

Histogram (Fast_Slow)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10

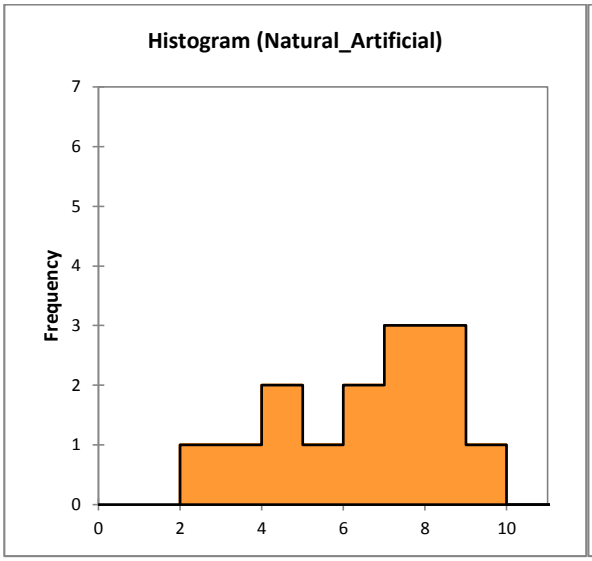

Histogram (Sharp_Flat)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10

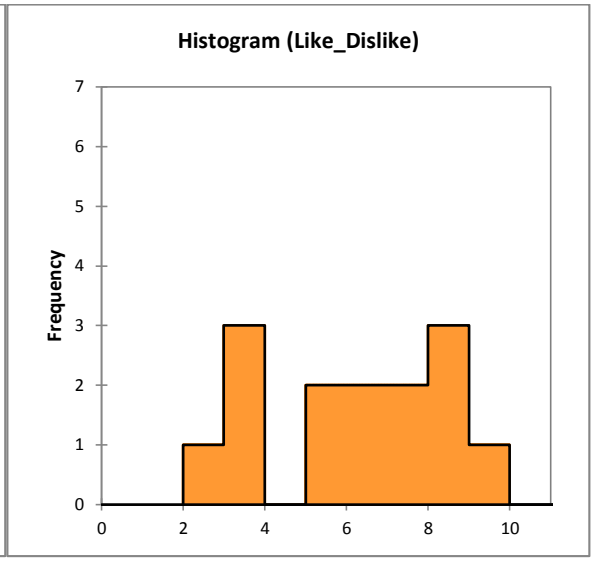

Histogram (Directiona_Universal)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10
Histogram (Varied_Simple)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10
Histogram (Reverberant_Anechoic)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10
Histogram (Far_Near)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10

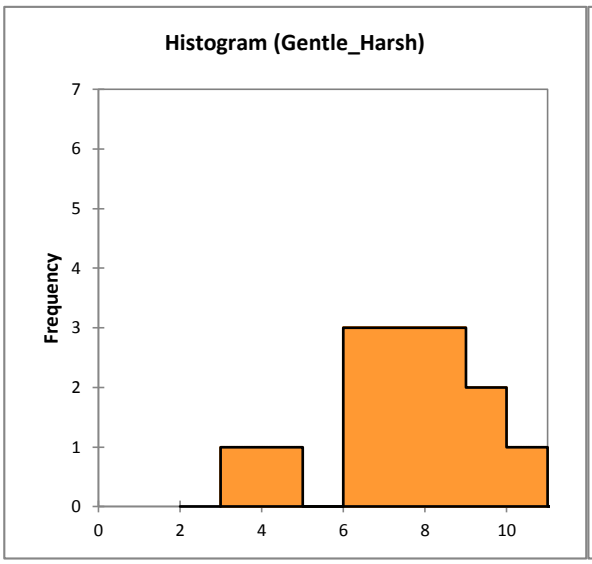

Histogram (Social_Unsocial)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10
Histogram (Meaningful_Meaningless)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10
Histogram (Calming_Agitating)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10
Histogram (Smooth_Rough)
Frequency
0 1 2 3 4 5 6 7 8 9
0 2 4 6 8 10

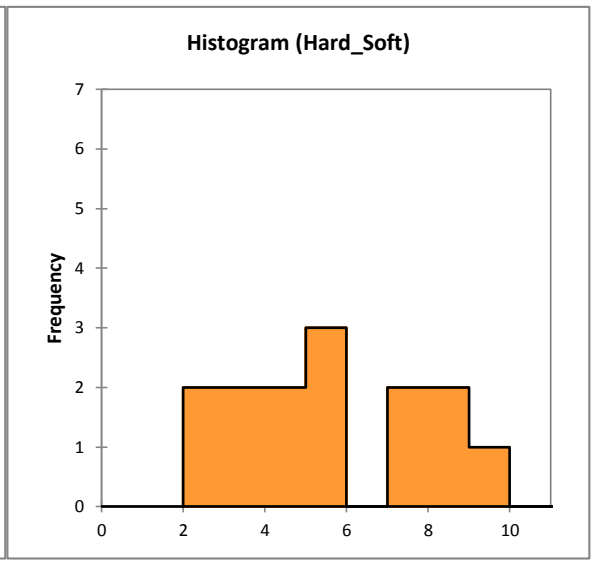
7.3.3 Binaural

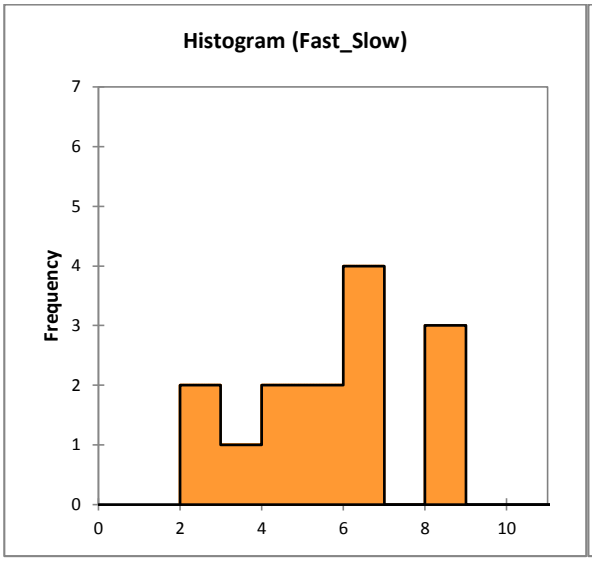


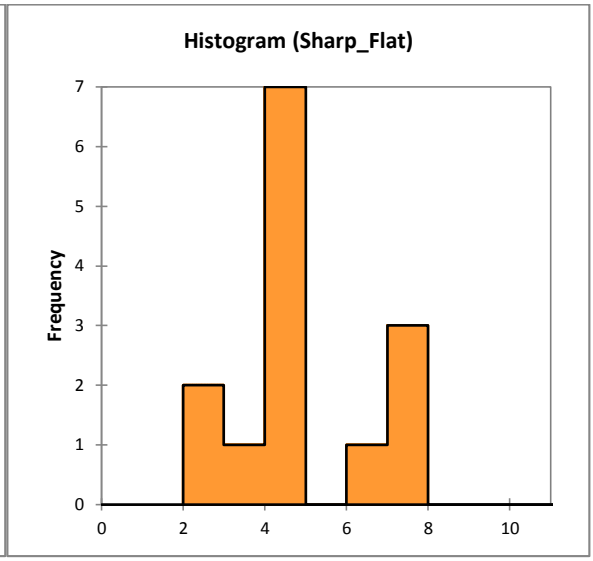

Histogram (Natural_Artificial)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

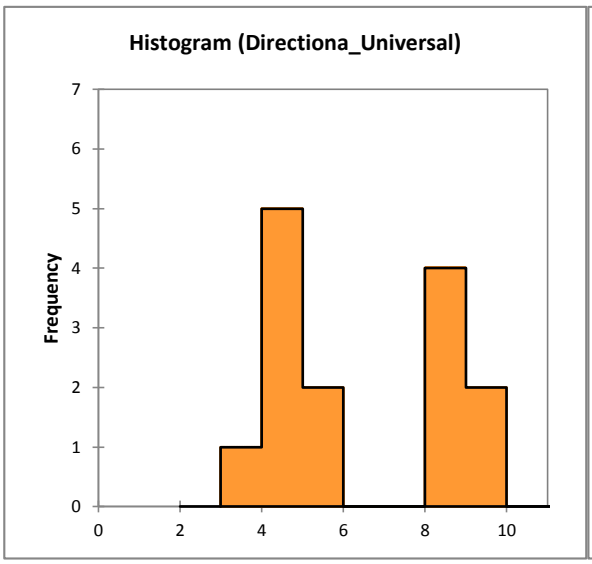

Histogram (Like_Dislike)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

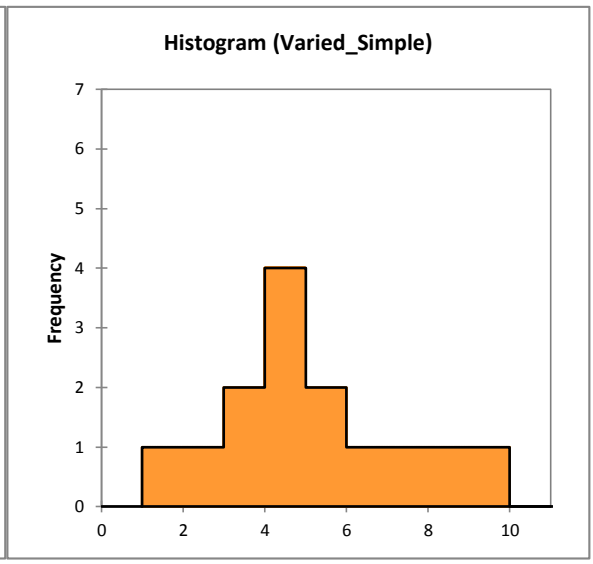

Histogram (Gentle_Harsh)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

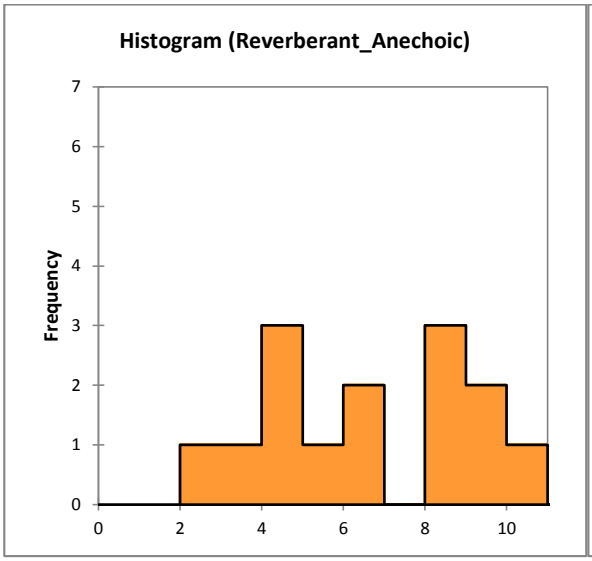

Histogram (Hard_Soft)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

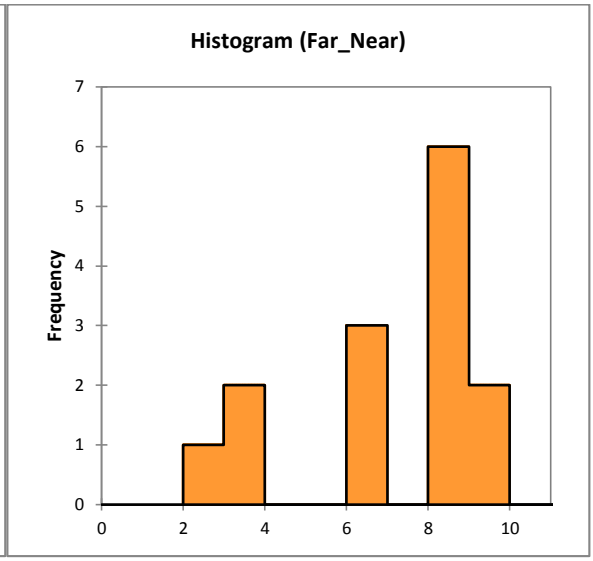

Histogram (Fast_Slow)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

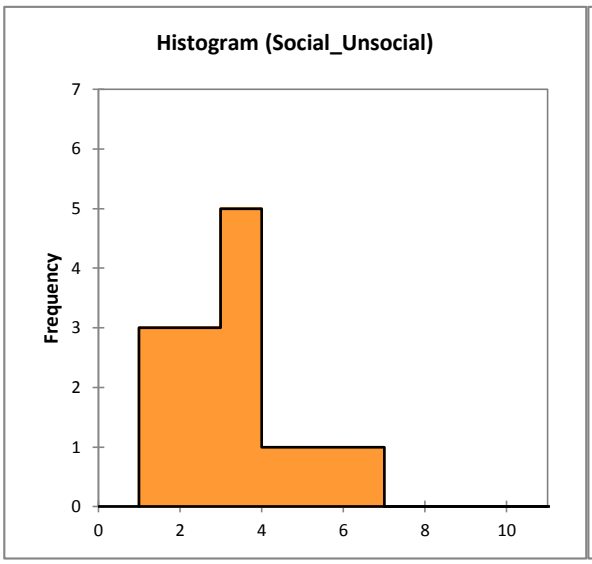

Histogram (Sharp_Flat)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Histogram (Directiona_Universal)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

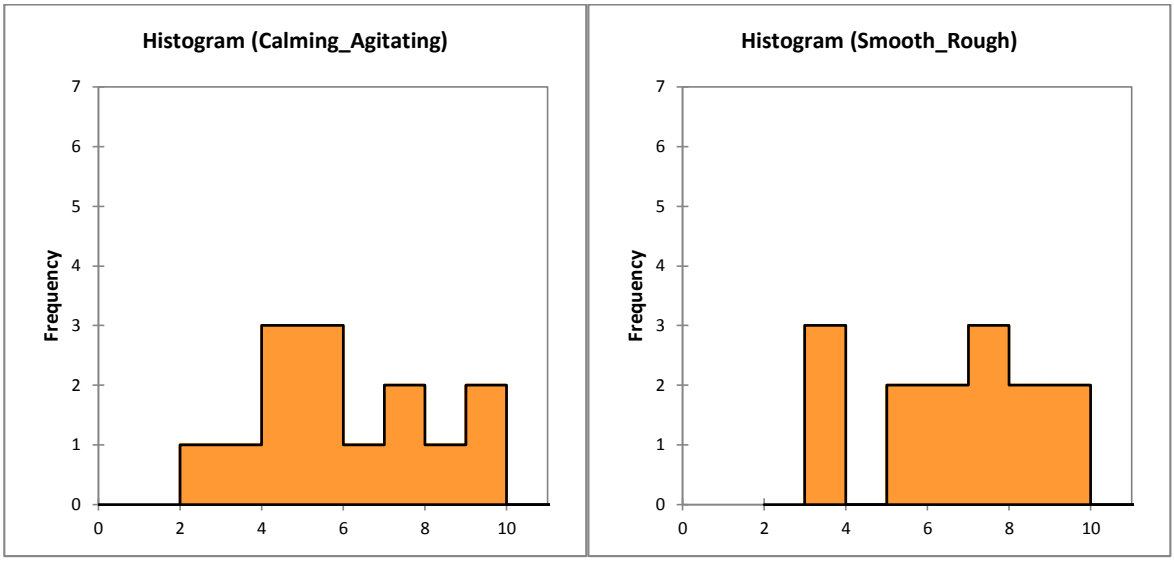

Histogram (Varied_Simple)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Histogram (Reverberant_Anechoic)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

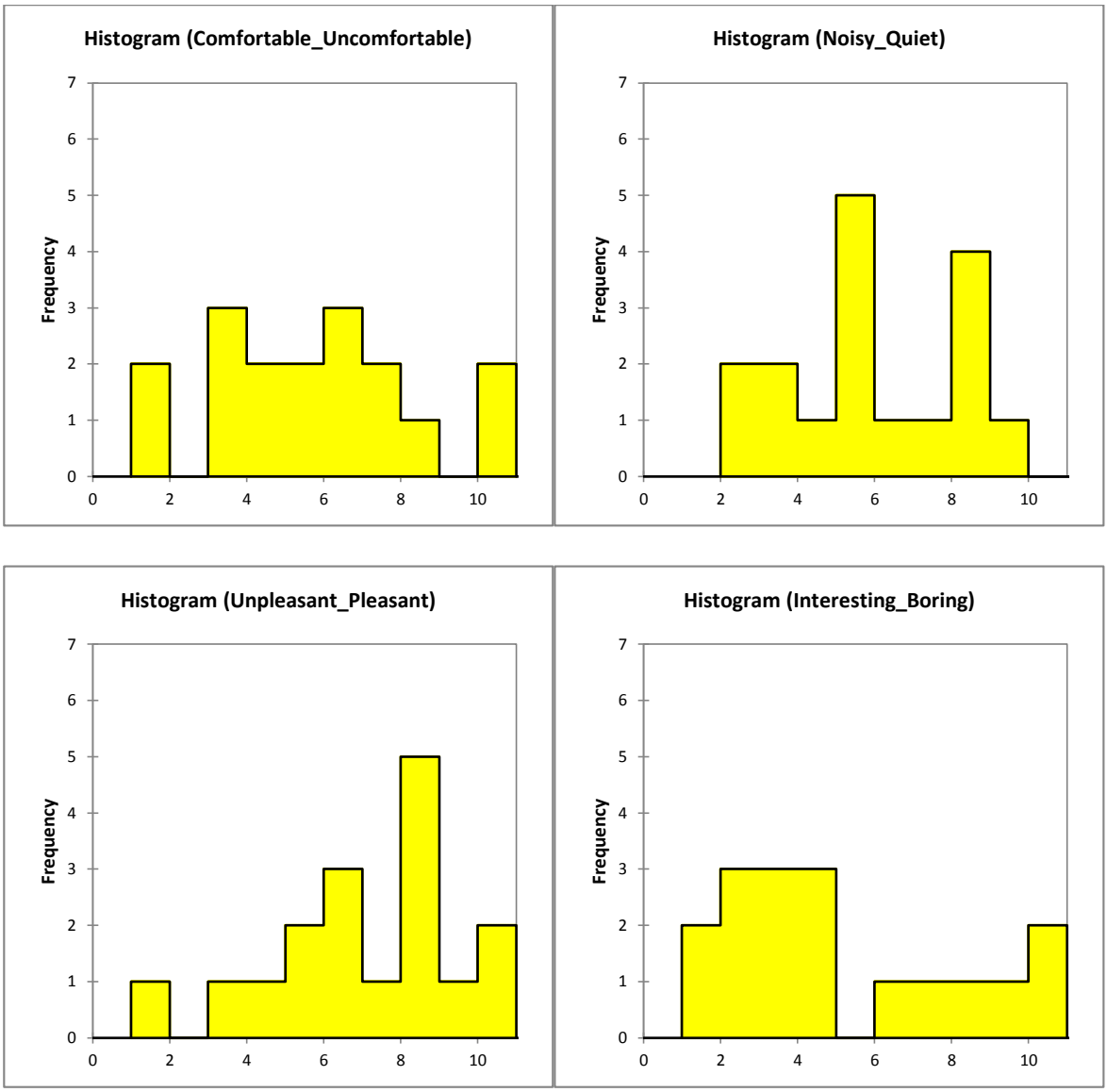

Histogram (Far_Near)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

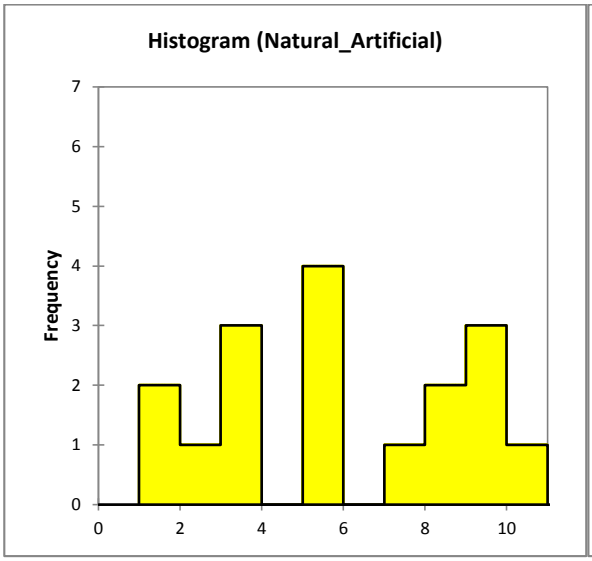

Histogram (Social_Unsocial)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

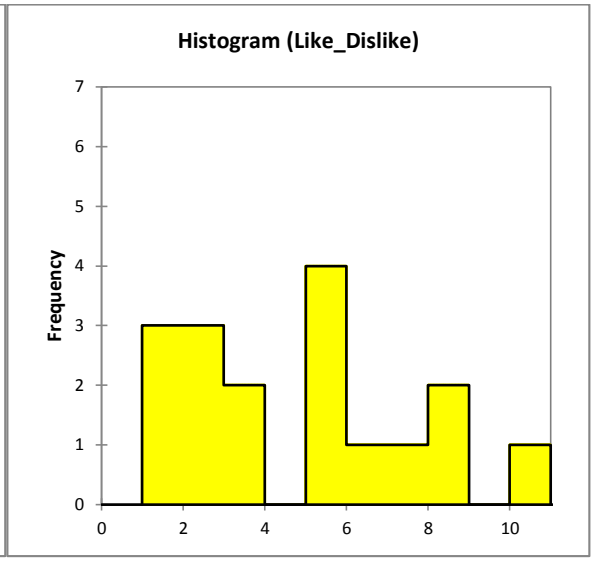

Histogram (Meaningful_Meaningless)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

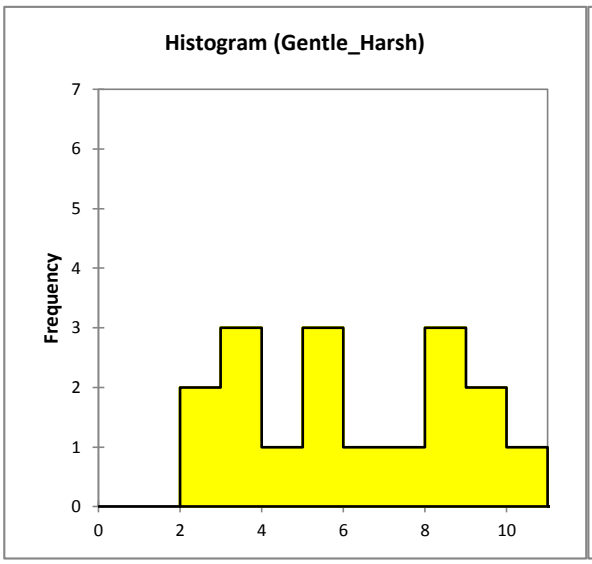

Histogram (Calming_Agitating)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10
Histogram (Smooth_Rough)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

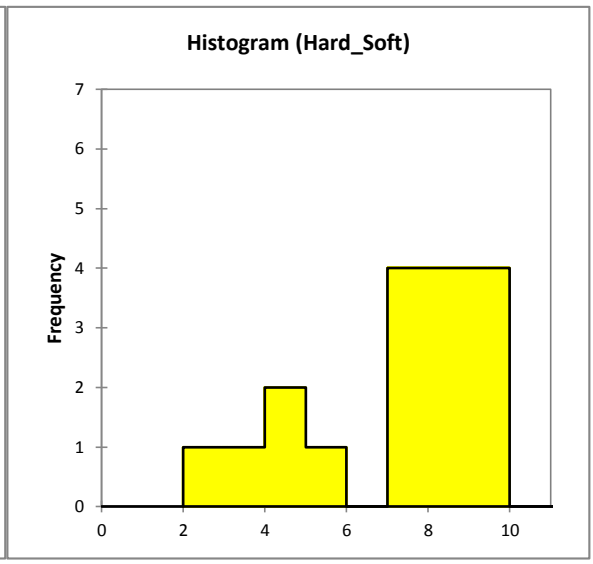
7.3.4 Real-time Binaural In Situ

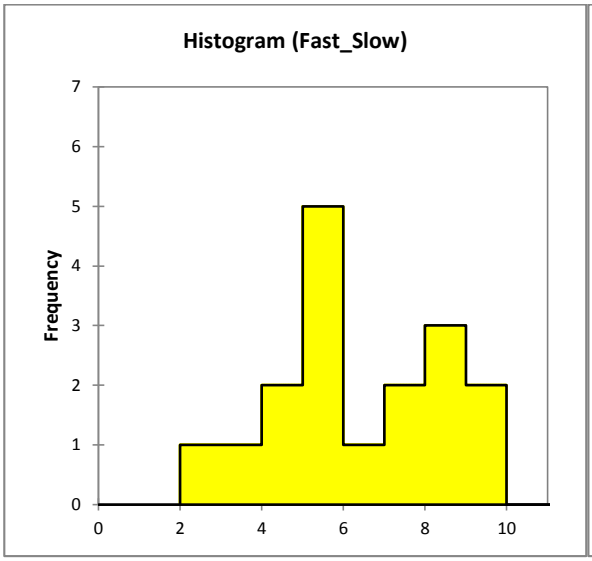


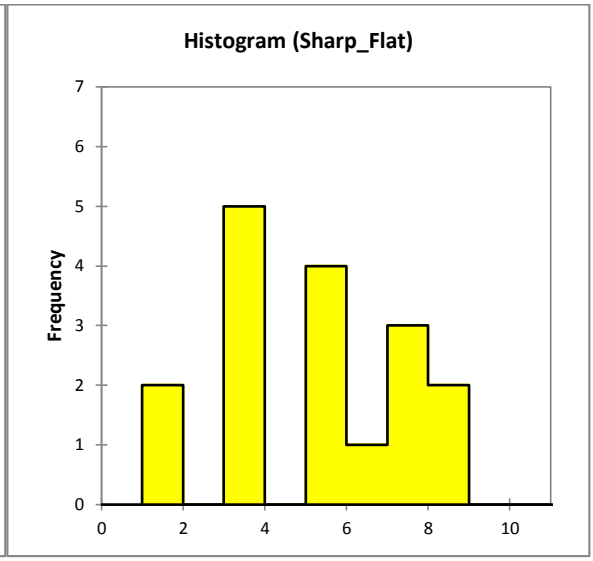

Histogram (Natural_Artificial)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

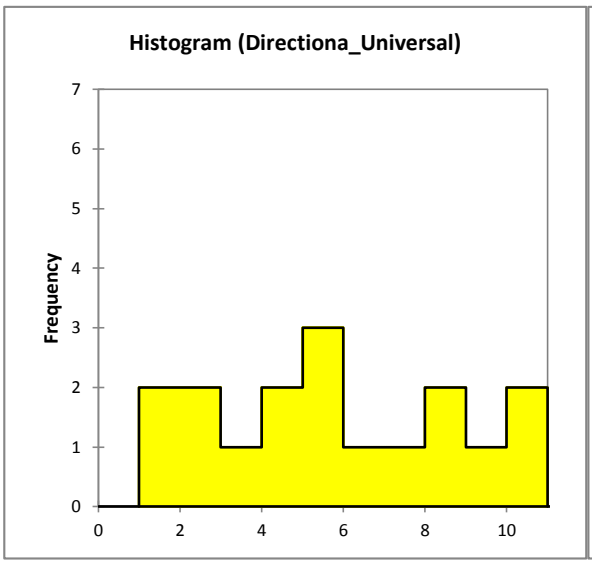

Histogram (Like_Dislike)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

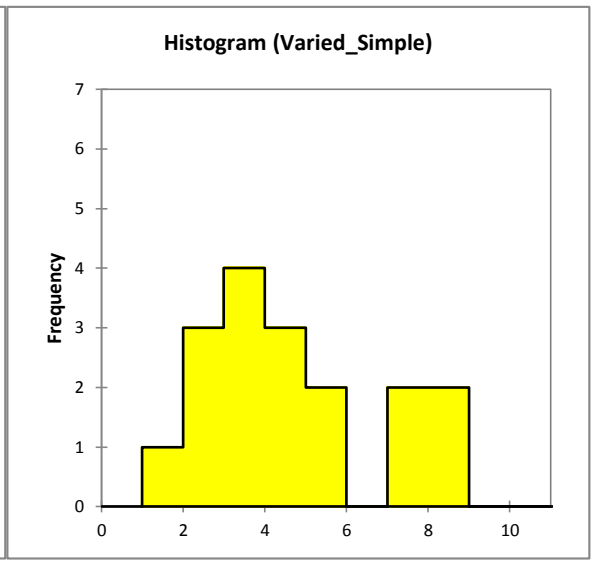

Histogram (Gentle_Harsh)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

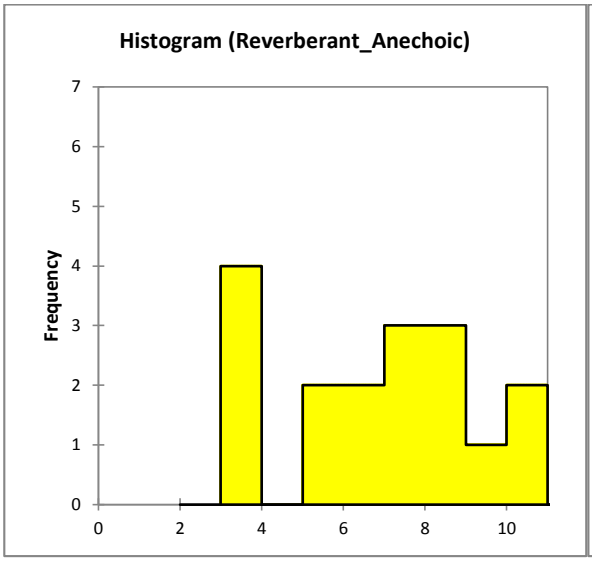

Histogram (Hard_Soft)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

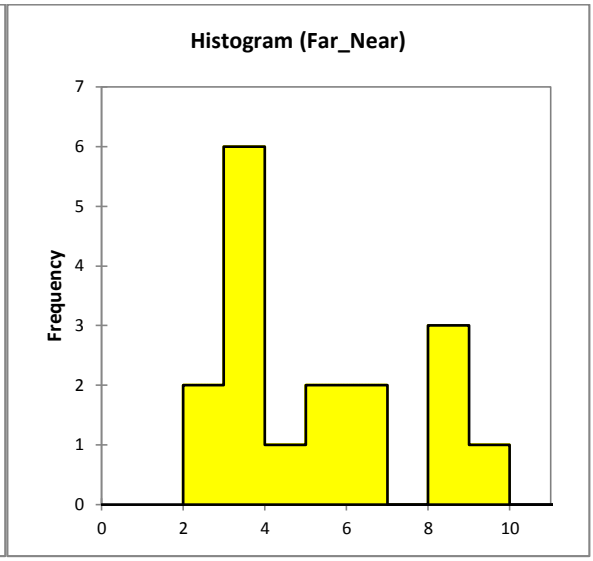

Histogram (Fast_Slow)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

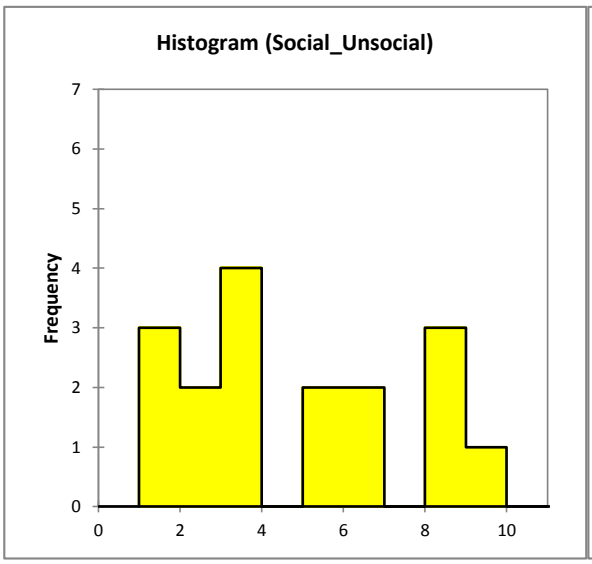

Histogram (Sharp_Flat)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

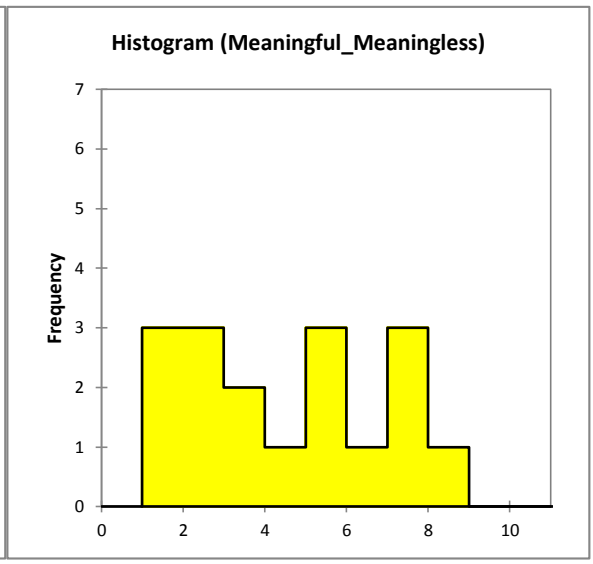

Histogram (Directiona_Universal)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

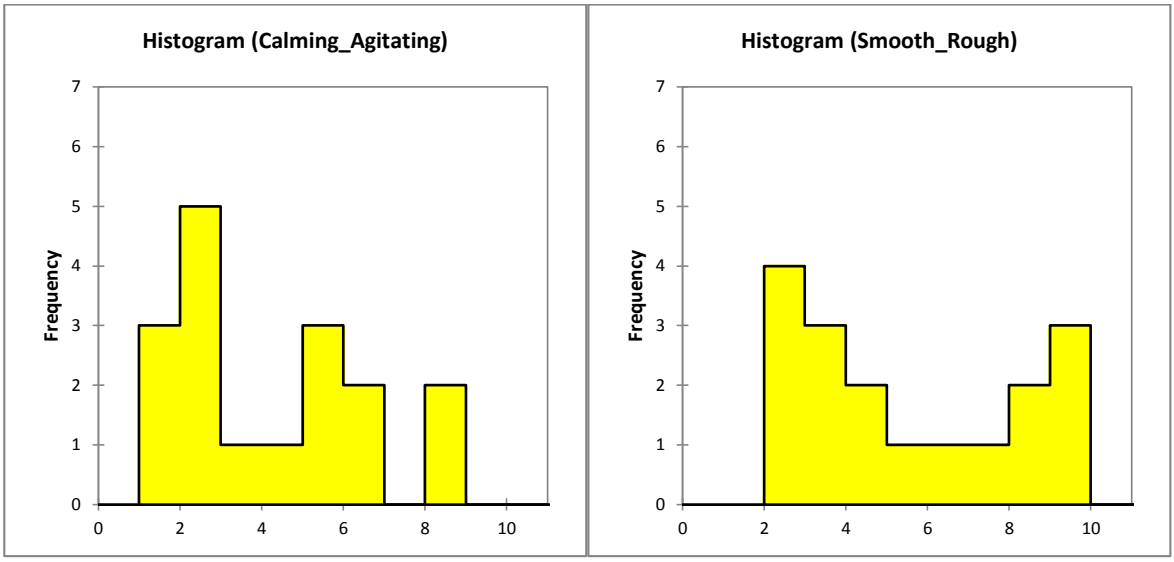

Histogram (Varied_Simple)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Histogram (Reverberant_Anechoic)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Histogram (Far_Near)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Histogram (Social_Unsocial)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Histogram (Meaningful_Meaningless)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10


Histogram (Calming_Agitating)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10
Histogram (Smooth_Rough)
Frequency
0
1
2
3
4
5
6
7
0
2
4
6
8
10

7.4 Appendix D

7.4.1 Phase One: Open Responses

7.4.1.1 In Situ

7.4.1.1.1 OUTSIDE THE NEWTON BUILDING, UNIVERSITY OF SALFORD, SALFORD

Amrit

1) People talking, background heavy machinery noise.
2) Lively. Energetic.
3) I like the sounds of people talking, it is cheery.
4) I would cut out the sounds of the machinery. It ruins the ambience and is abrupt.
5) The sounds of the machinery since it disturbs the overall ambience (cuts through) the chatter.
6) Car driving past is abrupt and disturbs the ambience.

Ben

1) Busy with people – general background conversation, noise from demolition nearby.
2) Conversation. Demolition – annoying.
3) Good feeling of people around.
4) Perhaps less people – a bit more peaceful.
5) Demolition noise is very irritating – doesn't help me think.

Chad

1) People talking very prominent
2) Sound of the digger. People. Sounds industrial.
3) The sound of people chatting is almost excitable – young people.
4) Less noise from the construction. It is annoying.

5) Sound of people chatting. They sound excitable with laughter etc.

Dylan

1) Conversation. Construction. Footfall.
2) Bustling. Human. Industrial.
3) Lots of people sounds make you feel part of something.
4) Lose the construction noise.
5) Conversation adds energy and place.

Ethan

1) Very busy. Students chattering. Construction noise.
2) Very lively, lots of people chatting, lots of different languages, multi-cultural.
3) Lots of people chatting makes it feel sociable.
4) Get rid of the construction noise.
5) I like hearing lots of languages – makes it quite exciting.

7.4.1.1.2 INSIDE SALFORD PRECINCT, SALFORD

Amrit

1) Children shouting, sound of footsteps, reverberant voices, the noise of a machine, sound of cart.
2) Lively, busy.
3) It is busy, lively, makes you feel active.
4)
5) The voices of children, cheers you up.
6) Reverberant sense of place.

Ben

1) Trolley wheels scraping on the floor, high pitched noise of babies and children.
2) Not very peaceful.
3) Nothing.
4) Maybe some relaxing music played.
5) Children's noise – irritating and annoying.

Chad

1) Children crying. People talking.
2) Reverberant / spacious. Busy. Unintelligable talking.
3) Nothing.
4) More absorbtion to reduce reverberation – children crying is emphasised by reverberation.
5) People accents are quite strong and can be heard when they talk loudly.

Dylan

1) Children, footfall, bean grinder, reverberation.
2) Artificial, human, hard.
3) Conversation makes it less dull.
4) Less reverb.
5) Conversation adds vibrancy.
6) Sounds too hard.

Ethan

1) Busy, lots of children and shopper. Reverberant sound. Footsteps on hard floor. Northern accents!
2) Busy. Lots of chatter. Children crying / shouting.
3) Very lively.
4) Maybe add some music / pleasant sounds.
5) Sound of children is quite annoying.

7.4.1.1.3 OUTSIDE AT MEDIACITY, SALFORD QUAYS, SALFORD

Amrit

1) Sound of distant and near traffic. Distant voices. Whirl of air conditioning units. Sond of rustling tree. Sond of security guard's radio.
2) Calming, peaceful. Country like.
3) The ability to tune into different aspects of the background noise or not.
4) Get rid of the traffic noise as it cuts through and disturbs the peace.
5) The peace and quiet. Allows escape.
6) How quiet it is for a place in the city.

Ben

1) Very quiet. Quite peaceful.
2) Large outdoor TV screen background talk but too quiet to hear properly – find myself straining to hear it.
3) Peaceful (fairly).
4) Turn the volume off on the screen, at this ditant straining to hear it (100m).
5) The peacefulness.

Chad

1) Sound of the large TV screen. Sounds very relaxed. Very limited sounds.
2) Sound of cars on a distant road. Sound of wind through the trees. Gain, very relaxing. Sound of people rehearsing singing in the distance – not frequent.
3) It sounds peaceful. Not very many people around.
4) The sound of the big TV screen removed – it's annoying probably because I am not watching it. Remove the 'ticking' sound – possibly A/C.
5) The lack of sounds and ambient sounds are relaxing.

Dylan

1) Big screen audio adds excitement. Dull traffic noise in background. Industrial tram.
2) Relaxing. Tranquil.
3) Relaxed atmosphere created by the soundscape.
4) Try to dampen the traffic noise. Increase the Quay sounds.
5) The tranquillity calms me down.
6) Very quiet!

Ethan

1) Slight road traffic noise. Quiet chatter, sound of bicycle spokes – peaceful. Sound from TV. Tram sounds. Choir singing – lovely
2) Peaceful, quiet.
3) Very peaceful.
4) Nothing, it's nice and peaceful.
5) Sound from the screen is interesting as you don't often hear that outside.
6) I enjoy quiet soundscapes, so enjoying this the most so far.

7.4.1.1.4 MANCHESTER'S TRIANGLE, CITY CENTRE, MANCHESTER

Amrit

1) Sound of glasses chinking, footsteps. Distant laughter, distant traffic. Screeching of car/bus brakes.
2) Busy, lively.
3) It sounds busy – lots of activity.
4) Nothing.
5) The distant bustle – it is interesting.
6) The sounds are very sparse, they come and go so you can listen to different elements.

Ben

1) Footstep and general chatter of people passing through.
2) See above. Also background noise of traffic and background noise of large TV screen.
3) Fairly quiet – not too busy.
4) Rid the noise fro the TV screen, it's the most annoing.
5) See above.

Chad

1) Feels busy. Feels social.
2) Urban – buses / cars.
3) Sounds lively and like there is a lot going on.
4) Nothing.
5) The sound of the buses or trams reminds you that you are in a city. (Screeching wheels).

Dylan

1) Conversation is calming and welcoming. Traffic noise is stressful.
2) Bustling. Urban. Human.
3) High energy feels exciting.
4) Reduce the traffic noise. More birdsong.
5) Human movement and chatter makes you feel part of something.

Ethan

1) Chatter – sociable, lively.
2) Quite lively.
3) Lots of talking and laughter gives a positive impression.
4) More people, it's relatively quiet for the city.
5) Lots of chatter makes the city seem lively.

7.4.1.1.5 PUB GARDEN, THE KINGS ARMS, SALFORD

Amrit

1) Distant sound of traffic. Screeching of brakes. Whistling noise. Hum of air conditioning unit.
2) Busy, lively.
3) Quiet mixed with sound of the traffic in the distance allows you to relax or tune into the distant sounds.
4) Nothing.
5) The noise of the air conditioning unit – it is annoying.

Ben

1) Silent apart from the conversation of our group.
2) Faint noise of the air conditioning? Very faint, whole feeling quite relaxing. Occasional noise of traffic from the nearby road – disturbs the peacefulness.
3) Nice and peaceful.
4) Block out the traffic noise.
5) Traffic noise – spoils the peacefulness.

Chad

1) Sounds of mechanical noise from inside the pub – sounds very dull. Cars – road nearby, not very exciting.
2) Industrial sounding noises.
3) Fairly peaceful.
4) Sound coming from inside the pub could be reduced or gone completely.
5) The sound of two nearby people talking.

Dylan

1) Dull traffic noise. Air con drone adds the artificial nature.
2) Monotonous urban drone.
3) Nowt.
4) Block traffic noise. Add more trees to encourage birds in.
5) Traffic noise 'cos it reinforces the urban atmosphere.

Ethan

1) Air conditioning sounds a little industrial. Pub chatter seems friendly.
2) Slight industrial noise from air con.
3) Quite peaceful for a pub.
4) A bit more chatter if the pub was busy would be nice.
5) Pub chatter is friendly.

7.4.1.2 Ambisonic Reproduction

7.4.1.2.1 OUTSIDE THE NEWTON BUILDING, UNIVERSITY OF SALFORD, SALFORD

Amrit

1. Background noise people talking. Footsteps. Birds. Door slam. Ticking sounds.
2. Busy. Spacious
3. Birds in background adds depth.
4. Background, it's annoying.
5. Bird sound – calming.

Ben

1. A monotonous wind-like sound. As if being in a windtunnel.
2. Footsteps and voices of people walking and talking around together with some kind of squeaking noise.
3. Quite peaceful.
4.

5. Trying to determine the origin of wind like noise – puzzling – too monotonous for wind?!

Chad

1. Not many people. Not busy.
2. Constant. Tranquil.
3. Not very noisy. Peaceful.
4. People having loud telephone conversations.
5. I find the background noise quite calming. It is not a very eventful environment.

Dylan

1. Traffic; footfall; door closures; birdsong; conversation
2. Hard and urban; grey?
3. Human buzz so not isolated
4. Drop the traffic noise as is annoying
5. Background traffic drone is annoying

Ethan

1. Occasional footsteps, chatter and birds tweeting add some liveliness, otherwise fairly monotonous
2. Quite monotonous
3. It' quite calm but occasional footstep and chatter make it more lively
4. Perhaps more variety in the soundscape
5. Not very exciting. Chatter and bird tweeting gives most variety.

7.4.1.2.2 INSIDE SALFORD PRECINCT, SALFORD

Amrit

1. People talking. Something being dropped (metal) (wood) objects. Wheels moving (prams) (pushchairs). Nail gun. Whistling baby crying. Reverb of Precinct.
2. Busy. Bustling.
3. It' quite cheery, because there are things going on.
4. Nail gun sound, because it cuts through, breaks soundscape abruptly.
5. Nail gun annoying.

Ben

1. Loud banging and general loud mixture of noise.
2. Annoying.
3. Nothing.
4. Perhaps quiet background music.
5. Banging – very annoying.

Chad

1. Reverberant. Busy. People are working or active.
2. A busy working environment. Lots of people moving.
3. It is lively. Has lots of different “characters”.
4. Removal of specific people speaking louder than others.
5. The general soundscape reminds me of a “shopping centre” experience. Child shouting is unpleasant.

Dylan

1. Reverberating humans; artificial signals reverberation; impulsive sounds shock.
2. Hard, reverberant and artificial; all human.
3. Energetic soundscape good but only if I wasn’t tired. Maybe too confusing.
4. Less reverberation as is a mush of sound; hard to pick out sources through the soup.
5. Humans all round, can’t localise.

Ethan

1. Lots of chatter and children makes it feel busy, a little chaotic.
2. Busy, lively, chaotic, dense.
3. It seems very lively, lots of people and business.
4. A little less reverberation, echoes make it very noisy
5. Lots of chatter is quite exciting

7.4.1.2.3 OUTSIDE AT MEDIACITY, SALFORD QUAYS, SALFORD

Amrit

1. Background traffic noise. Distant footsteps. Coughing. Squeak of tram.
2. Busy. Fast moving traffic.
3. Sparse fell because it allows you to tune in and out and listen to different sounds.
4. Nothing.
5. Tram sound, on the rails, because it is grating.

Ben

1. Loud squeaking/squealing noise – painful.
2. Quite peaceful footsteps
3. Fairly peaceful.
4. Some wildlife noise.
5. Loud squealing – painful.

Chad

1. Birds sound peaceful. People walk casually. No real sense of urgency.
2. Quiet. Tranquil.
3. Very quiet and calm.
4. Tram has loud brakes – should be quieter.
5. The sound of the shoes walking – no sense of urgency – calming.

Dylan

1. Can hear wind noise; car sounds; footfall; tram brakes
2. Relaxing background sounds; urban but well balanced.
3. Low overall level but not too low.
4. More nature, e.g. Ducks, brids, Shetland ponies.
5. The relaxing element; love urban and relaxing coexisting

Ethan

1. Tram noise adds slight industrial feel to otherwise quiet soundscape
2. Very quiet, calm, peaceful.
3. It is nice and peaceful for a relatively urban environment.

4. Remove tram squealing as it cuts through otherwise peaceful soundscape, perhaps add some pleasant garden sounds, birds, windchimes etc.
5. Tram noise is a little irritating.

7.4.1.2.4 MANCHESTER'S TRIANGLE, CITY CENTRE, MANCHESTER

Amrit

1. Busy. Footsteps. Distant voices. Slight reverb. Banging sound. Bus noise. Traffic noise.
2. Busy – people moving quickly through.
3. Indifferent.
4. Remove background noise. Disturbs soundscape.
5. The fact that it is busy makes me feel restless.

Ben

1. General conversation and footsteps – sound quite busy with people moving.
2. Too busy
3. Footsteps – nice sound as they pass
4. Less people – quieter.
5. Footsteps – quite relaxing.

Chad

1. People are shopping and seem to be happy – social.
2. Background noise is general city traffic and transport (public transport). Not too loud, not undesirable to listen to.
3. Not a very "busy" soundscape.
4. The "whine" of the buses acceleration.
5. The sound of the traffic brings back memories of London and general city-centre experiences.

Dylan

1. Footfall; conversation; traffic hum
2. Large urban centre; hard man-made
3. Energetic and full of life

4. Maybe reduce the traffic leaking in but not too much
5. People passing and chatting makes you feel part of something.

Ethan

1. Lots of footsteps and chatter adds liveliness
2. Lively, but not too loud (in contrast to Precinct)
3. Relatively lively but not too loud/chaotic for city centre
4. Try to remove some of the road noise, brake squeals of buses(?) etc.
5. Chatter makes it have a nice degree of liveliness

7.4.1.2.5 PUB GARDEN, THE KINGS ARMS, SALFORD

Amrit

1. Traffic noise. Car doors. Distant bus sound. Distant car horns. Wind chime. Truck reversing sound. Drain cover as car drives over it. Gate sound.
2. Busy.
3. Foreground – is quite calming.
4. Remove traffic noise. Disturbs the ambience.
5. Windchime – has a calming effect.

Ben

1. Horns and noise of traffic
2. Busy and annoying.
3. -
4. Take away the traffic – too noisy.
5. Horn of a tram (possibly)

Chad

1. All sounds appear to be fairly distant.
2. Traffic, fairly busy. Urgency of traffic is noticeable.
3. Nothing in particular.
4. The traffic sounds (beebing also).

5. The distant sounds and occasional unrecognisable elements make the scene fairly eery – cannot predict what will happen next. Could be a horror film background sounds.

Dylan

1. Car passes; gate closing; knocking; horn
2. Urban but not central urban
3. Quiet but still carries some of the cities energy
4. -
5. Traffic and passes keeps up cities energy; quiet enough to be relaxing

Ethan

1. Road traffic noise is quite dull and monotonous
2. Mostly quiet, monotonous road traffic noise
3. It's nice and quiet, but a little too dead for a pub
4. Add more chatter and liveliness from pub patrons, enough to mask road noise
5. The fact that you can hear road noise makes the pub seem very dead/empty.

7.4.2 Phase Two: Open Responses Categorised

7.4.2.1 In Situ

P1

1. Industrial. Wind. Water. City. Trams.
2. One volume, never stops, humming.
3. People, you can always hear people. I like that.
4. Right now… stop that man with his jetwash.
5. Trams – I like travel and they have a unique sound on the track.

P2

1. Nice humming of the traffic. People speaking, wind blowing, machinery.
2. Nice, industrial, business-like.
3. It changes dramatically from one moment to the next.
4. I would increase the sound of people enjoying themselves.
5. Background noise of people walking.

P3

1. The sound of the wind is very calming. I can hear the noise of the power washer in the background.
2. Soothing. Happy, I'm hearing a lot of laughter.
3. There's nothing I like about this soundscape.
4. Music on speakers. Radio.
5. Laughter.

P4

1. Industrial, empty, business, sterile.
2. White noise, engines, workmen.
3. Stuff is happening, important stuff.
4. More people, warmth, chatter.
5. The business of the few people here, catching glimpses of conversation.

P5

1. Openness, peace, tranquil.
2. Background noise is a gentle buzz that supports how free the space feels.
3. Slow, stressless and effortless the place seems.
4. More of it.
5. The buzz of the background noise.

P6

1. Howling wind. Water pump. Paper bag.
2. Still.
3. I don't like it. It's too quite that the only noises I can hear are industrial noises of men at work.
4. Nothing.
5. Wind and industrial sound.

P7

1. The wind gives a feel of openness. Odd voices and chuckles are warming although distant machine is overly dominant.
2. Quite deep background ambience. Little to identify.
3. Combination of openness and people.
4. Lower machine noise (albeit not permanent).
5. People and footsteps and the echo is rather characterful – cold but gives a sense of place (hard edges).

The following four collected on 14/5/14. Sunny day, 1-3pm, many people chatting and lunching, some minor construction works (close to P8).

P8

1. The loud humming and drone of the commute from the day to day life.
2. The character is strong and profound.
3. It makes me feel welcome and lets my mind wonder, to go to another place.
4. I would change nothing, let it be whatever it wants to be.
5. Wind -> it is quite fulfilling.

P9

1. Birds, wind rustling the trees, bell ringing. Engine noise. Feet walking / run for tram.
2. Workmen / clanging. Faint laughter chatter. Bike wheels moving faintly.
3. Feels vibrant, people around a busy social and work environment.
4. Less mechanical noise so it's calmer.
5. Feeling of space – not hemmed in.

P10

1. People chatting – light, casual, social. Wind – gentle waves of white noise in my ear and faint tree leaves rustling. Heavy vehicles moving and clanking – invisible but implies construction, expansion. Wheeling office chairs across paving – bit comical. Segway going past. Helicopter.
2. Quite unintrusive, not much 'natural' noises of birds or water or insects. No key noises like a clock tower / big ben / public transport that locates it.
3. Background chatter is quite relaxing.
4. Would like more water noise like a fountain like Geneva. More birds. Maybe more music.
5. (Segway sounds futuristic). Chatting, laughing is relaxing.

P11

1. Human conversation – right sound for purpose. Helicopter – not typical I'd say. The sound of being outside, not a specific sense / awareness. Distant.
2. Distant, unintrusive, if you're consciously listening towards it all you're aware of people doing work.
3. Peace. Not brash or in your face. Evidence of a community through the sound of peoples' conversations. Peaceful having tress blowing. Unintrusive.
4. Nothing. As a work environment, it serves a purpose. You want a mental break from work. Quite peaceful.
5. Conversation of others – only thing actively noticed.

7.4.2.2 Ambisonic Reproduction

P1

1. Footsteps are very prominent and their passing adds to the shape of the environment (feeling of how it looks?) These I like. The occasional distant coughs by contrast were hard to place and the occasion snaps of chatter were sudden and did not create the environment like the moving feet and tram did. (I only know it's a tram from being there.)
2. I think the background noise is the distant cars as it's always there and does not feel as three-dimensional as the rest. In the beginning when all I heard was background noise I did not like this. It felt oppressive. But as it progresses I heard layers of sound.
3. The build up of layers of things, when the tram passes and the feet.
4. The people talking in the distance – either that they'd do it continuously and stay in one place so i could locate them or if you could hear them closer in but walking past because the movement of the sounds making the space shape was enjoyable.
5. Tram going by – I can't think of a word for the emotion but it is positive as it takes me to that place fully (transported?).

P2

1. Events slowly enter and exit the space. Light voices appear and disappear. Footsteps regular and circulating.
2. The traffic noise blends events together, provides smooth transition, lightens construction.
3. I like its envelopment and how it seems to move from one source to another. Sounds blend into each other with a nice echoy feel.
4. An engine-like rattly noise. Doesn't play well with the atmosphere.
5. The noise of a breaking train (or tram) echoes beautifully. There's something tragic about it, sounds like a lament.

P3

1. Footsteps, tram sounds sound close, a lot of detail. Personal feels to the sounds, brief windows of buy-ness.
2. White noise-like, wide frequency range, constant loudness. Presence of an oscillating sound source.
3. The background noise is soothing due to its uniformity. No real loudness peaks.

4. Remove the oscillation in the background sound – it is distracting – the least pleasant part of the soundscape.
5. The intimacy of the fore-ground sounds VS the generic nature of the background sounds – good contrast.

P4

1. Subtle voices and footsteps give it a sense of realism. Tram noise and echo gives a sense of the size of the space (also quite exciting).
2. Constant distant road noise gives a sense of openness and distance. Constant with little variation. Somewhat boring/unexciting.
3. Calm. Variation, but not too much. No loud 'crash' and 'bangs'.
4. Slightly less road noise (or whatever the constant noise is throughout). Bring events such as footsteps, voices and other human noise a bit closer.
5. I really like the sound of peoples footsteps, interesting to hear them move around.

P5

1. Footsteps; fan; construction noise; coughing; tram noise. The entire scene seems relatively noisy. It is not calm or inviting.
2. The background noise is annoying and distracting.
3. The train and few footsteps are interesting. They seem to give you a sense of place.
4. The industrial fans and construction noise is really annoying. It seems to distract from the experience and drown out the interesting stuff / movements.
5. The tram sound because it reminds me of the train station I lived near back home.

P6

1. Tramline gnissel, constant level of infrastructure and traffic noise. Still somewhat pleasant.
2. Urban, high noise floor i.e. urban soundscape, fans, ducts, infrastructure. Yet non obtrusive, similar in aspects to pink noise. A calming effect.
3. I like the open feel / transparency of the reproduction, also the detailed vertical localisation.
4. Noise barrier to subdue traffic and building noise. Perhaps a row of hedges.
5. Very exciting, nice envelopment, nice full frequency content.

P7

1. I can hear the sound of tram, footsteps, brake.
2. I feel the background noise is like the sound in a plane.
3. I like the envelopment and I can feel the dynamic of the soundscape esp the moving sound.

4. I think the noise is too dominant and because I don't really like the sound of plane, I hope I can move the sound.
5. Footstep – because I can feel it moving.

P8

1. Conversation (very quiet). Steps. Tram approaching, sound of wheels rolling along rails, squeaking (high pitch noise). Coughs occasionally. Construction work. I feel neutral about everything apart from tram squeak which I did not like.
2. I feel neutral about background noise. It seems to be sound of the centre in the middle of the day. Perhaps slightly relaxing sound because I've got used to it. It could also be slightly relaxing.
3. It reflects the sound of the city centre. I usually enjoy being there.
4. Nothing.
5. The sound of the tram. It emits annoying sound.

P9

1. It feels like a suburban environment. It envelops you. A constant blanket of what almost sounds like white noise yet isn't.
2. Distant. Barely audible but you can feel it. Your eardrums become accustomed to the noise.
3. I can relate to it. It reminds me of my daily commute.
4. Due to curiosity I would like to hear what the voices of the pedestrians are saying. Making the voices more audible.
5. The screeching of the train. It reminds me of excitement as I rarely use the train except when I'm going to a different city.

P10

1. Footsteps. Ambient. Talking. Indistinct. Possibly cars.
2. It's very difficult to pinpoint individual sounds – could be cars, or almost flowing water.
3. The sound of people (grunts, footsteps, talking). The soundscape gives a very realistic representation of ambient noises.
4. Perhaps it would be possible to amplify certain aspects more.
5. A screeching sound like brakes as well as footsteps.

P11

1. Walking. Windy. Something spinning.
2. Calm. Busy. Subtle. City lifestyle. Constant.

3. Calmness – no loud sudden sounds. Consistency – shows the quietness of the environment. Walking – makes the environment feel less isolated.
4. The walking sound disturbs the quietness and subtleness slightly, along with the tram noise.
5. The wind and traffic noise.

P12

1. Sounds realistic, the atmosphere stands out, tram.
2. Sounds like sitting on a bench beside the waters and people walking by.
3. Gives me the opportunity to imagine what's going around in the environment.
4. Sounds like it's recorded in just in 1 position… perhaps if you did a 360degree turn on one position… might be able to capture more sounds?
5. It builds up curiosity to know if it a car passing by nearby or people gonna cross by.

P13

1. Traffic, footsteps, talking, metro.
2. Soft, calm.
3. The sound of (what I think) is the metro tracks resonating.
4. Background noise (traffic) is a bit high.
5. The sound mentioned in Q3 – seems to have a musical quality.

P14

1. The most prominent elements are the footsteps and tram(?) squeal. Although it sounds like there are people in the area, the lack of voices makes the *** feel subdued. The voices you can hear sound hushed.
2. The background noise makes the place feel subdued and lifeless. The broadband traffic noise(?) dominates other activity making the place sound empty.
3. It's quite peaceful.
4. It perhaps needs more lively elements to make it feel like something's going on. As it sounds like a big urban space, it feels like there should be more activity like cafes, markets, etc.
5. The squealing noise is quite jarring compared to the background elements.

P15

1. You can distinguish the difference between the whereabouts of cars and pedestrians. Such as human voices being played ahead of me and cars going past on the left.
2. n/a

3. It gives an actual feel of what is happening around you, if you stood in one place in a busy city.
4. Nothing to change.
5. The voices, cars and footsteps.

P16

1. The fuzzy noise sounds (maybe wind) seem to be the most dominant. A tapping sound also seems to occur away with the other background noise.
2. The background noise sounds to be people walking and driving by. Also trams stopping and possibly other daily things that would happen on a main road.
3. The soundscape seems to represent a typical afternoon in mediacity. People walking by and it not being too busy.
4. Maybe record at a busier time to have more going on and more to listen out for.
5. The sounds of footsteps, as it makes it seem busy.

P17

1. Hear rain; wind; like white noise on a tv. Hear screeching like bicycle brakes, chain rattling – bike lock? People talking; high heels walking.
2. Feels relaxed.
3. I like the bike noises because they don't blend in, also the sound of walking.
4. Add a different layer of noise traffic, makes it more interesting. Or add weather like thunder etc.
5. Bicycle screech – thought something was going to happen… but it didn't – anti-climax.

P18

1. Traffic noise – spacious – background noise/cough. Voices – distant unclear. Footsteps – distant, slow, crunchy. Tram noise – harsh, resonant. Sound of chain – fast. Engine.
2. Monotonous, spacious, enveloping. Industry. Pulsating.
3. Movement of the footsteps gives impression of vibrancy and purpose.
4. The background noise; general city noise dominates. There is some wind noise also that seems to some from behind. The level of the background noise makes it seem unreal.
5. Tram it's very noticeable, in a real situation I would notice and be sure to get out the way.

P19

1. I think it sounds like train when approaching to a certain point.

2. The background noise sound. Like river, trains passing by.
3. I think it gives good environemntql sound record. Recording environmental sounds.
4. Changing the level of recording and making it in scaling.
5. The drop sound like dropping a stone into the river.

P20

1. Broadband murnour of background noise is quite constant/static with quite a lot of low frequency content. It feels a bit like pink(ish) noise.
2. Broadband murnour, static. There is not much fluctuation in the level of the background noise, as it lacks character.
3. I like that the relatively bland background noise is punctuated by sounds such as the loud screech of the tram, people coughing and footsteps etc coming from various locations.
4. I would reduce the level of the background noise especially in the low frequency region. I might also add in more elements of interest such as discrete sources, people walking (signs of life).
5. The background noise from traffic etc in the distance makes it feel a bit bland and arid but the punctuation of discrete sound events make it more exciting/vibrant. But I think it is the background that very much dominates and affects my feeling of the soundscape.

7.4.2.3 Binaural Reproduction

P1

1. It feels noisy, there is sound of children and people talking in distance. I can also hear the footstep. There are many sound that I can't describe.
2. There is constant noise like white noise in the background.
3. I can feel the movement of sound
4. The sound of the child on my left ear at certain point was really loud and feels annoying. It feels noisy and plenty of sound that I can't describe.
5. The sound of children, it feels really loud and makes me feel uncomfortable.

P2

1. The child talking and the people laughing give the soundscape a friendly character. The prominent footsteps give the soundscape a dynamic feel.
2. The background noise is fairly homogenous, sound like low level traffic noise. Although its difficult to pick out the source, it give the soundscape the character of being somewhere fairly busy.
3. I like that there is a lot going on (different people talking and interacting) but it is still relaxing and friendly.
4. I can't think of anything I would change.
5. The people talking, particularly the child's voice. It gives the soundscape a friendly character.

P3

1. Busy, happy, transient, noisy
2. Busy, noisy, integrated hard to pick out single elements
3. The sound of the child, because it is a happy sound
4. Nothing
5. The distant background sounds, I can make them out but cannot easily identify them.

P4

1. Noisy like a marketplace and windy too or perhaps a artificial waterfall in a small pond.
2. Background noise is more like white noise as said like a artificial waterfall or might actually be heavy rain.
3. Nothing because its quite irritating.
4. The sharp sound of the little girl. It sounds harsh at times.
5. The talking between people. It attracts my attention immediately but its not exciting rather unpleasant.

P5

1. Childs voice and talking – shrill! Footsteps – distracts attention from general ambience
2. Sounds like a water feature – calming in otherwise busy space. General hubbub – crowd, laughing, talking – café culture feel – like sat in a busy town park.
3. Generally relaxing even though busy – atmosphere of people socialising in a busy world – like town centre park – excapism.
4. Get rid of child! – partially serious, just a bit shrill. Prob prefer slightly quieter general background noise. Gives impression perhaps bit too busy.
5. Voices that you suddenly pick out above crowd – grabs you suddenly. Generally lively.

P6

1. Chatter of people – sounds bustling, busy and vibrant. Occasional laughter sound happy. Shouting child(?) sounds happy too.
2. Very busy – lots of people footfalls, chattering.
3. It sounds very busy and potentially quite exciting. Lots of things happening at once.
4. More natural sounds would be nice – birdsong, water etc. overall noise levels would be more pleasant if it was a bit quieter.
5. The sound of what might be a child cuts through the background noise quite dramatically. It sound like it might not be English making it more diverse and exciting.

P7

1. Children voices. People chatting.
2. Water, coffee shops.
3. It feels like recording was made near a shopping centre. People are related and enjoying themselves.
4. It's quieter.
5. Sound is very relaxing I wouldn't want to change any aspect of it.

P8

1. There interested thing such as child events and other collection of various activities.
2. It is noisy, but I can recognise some events.
3. –
4. Make it less noisy.
5. Restaurant of playing events.

P9

1. Sharp, happy/positive, busy, human, fragmented (child's voice comes in and out)
2. Bustling, lots of layers, seems filled with people but non-threatening. Bell of the shop = sociable.
3. I feel comfortable in the middle of it. There's lots going on, lots of interest, but it feels secure, not exposed. The footsteps sound natural, absorbed into the ground, not harsh.
4. The child's voice is very shrill and piercing!
5. The hum of people – it's interesting and human – like aural people watching! (I'm nosy.)

P10

1. Child's voice – penetrating / piercing, annoying. Bell/chime, - synthetic, (only heard once?) footsteps – close, purposeful stride but not hurried.
2. Background - roar, loud, fairly consistent. – busy, lots of people perhaps.
3. I like it that it is busy but I'm not in a hurry. I sometimes like to sit and watch the world go by.
4. Reduce the child's volume – too loud and piercing. Also reduce the background level.
5. The child's voice was distracting and annoying – even uncomfortable.

P11

1. The children's voices are most prominent and stand out, other occasional sounds like footsteps, a chair sliding on the floor and laughter also stand out from the background.
2. The background noise is a mixture of voices and a general hiss. Quite harsh sounding as if the floor and walls are hard, maybe tiled. Lots of high frequency content.
3. It sounds like everyone is having fun. Laughing and children playing are nice to hear. The sounds are recreational.
4. The sound is a little harsh. If the low frequencies were more prominent than the highs it would sound more relaxing as an environment.
5. The sound is busy, typical of a filled cafeteria at lunch time (quite a large space). Sound like a generally happy place.

P12

1. A lot of high frequency energy. Busy but quite static. People moving furniture gives a feel of energy in the place.
2. Busy and quite noisy as constant.
3. It feels like there's a lot going on as the people seem happy in the soundscape.
4. I find the background noise too loud. Felt a bit unrealistically loud.
5. The element that affects me most was the children's voices as this sound is very noticeable and jumps out of the background noise.

P13

1. Harsh, young, noisy. Piercing, annoying.
2. Constant, busy, noisy, bustling, rustling, clinking.
3. Movement of some of the sounds, left to right.
4. Too noisy, reduce the noise. Remove the child's voice. It is very piercing.
5. Child talking, I found it harsh and always from one side.

P14

1. Some background chattering – particularly a child. Clattering of plates/cutlery. Footsteps.
2. Busy, confusing, outdoors, wide open space, full of people.
3. Good community atmosphere – sounds like a nice place to be, not boring or dead.
4. Get rid of the screaming child, conversation is nice to hear, it is good to be around other people, but noisy children are vey annoying.
5. The clip isn't too exciting, it sounds like a nice place to be. Maybe a bit noisy but a good place to be out, like being outside on a nice sunny day rather than coupe up inside.

7.4.2.4 In Situ Real-time Binaural Reproduction

P1

1. High pitched sounds along with speech.
2. Very deep and drowny
3. Listening to the small details such as paper rustling, because they are not sounds we're used to hearing or paying attention to
4. Take out any monotones because they don't add anything and drown out anything else, eg car engines
5. The sound of the motor its really horrible and boring, however the bird call was very nice and uplifting.

P2

1. Most prominent is the sound of background noise, sounds like rain
2. Sounds like rain storm
3. I don't like the soundscape
4. –
5. Its just background noise

P3

1. Nature – sounds of insects? Water
2. Feels like night time / dusk

3. Calming, sleepy – evening / camping sounds
4. –
5. Water – soothing, calm, fluid

P4

1. I feel noisy, I can hear the sound that normally I cannot hear. I also feel a little strange because I can hear "unnatural sound" on the background. It is like listening to a recording from cheap speakers. Even I can hear footstep and the sound when I flip this paper.
2. For me, the dominant character is the background noise. I describe them on the first question.
3. I like the fact that I can do localisation easily. And because I can hear where the sound is coming from I like to check where it is from.
4. I like to reduce the background noise because it is strange to be more sensitive in hearing.
5. I feel I can hear more sensitive than before and for me it feels unnatural.

P5

1. Laughter, cheerful, relax, chilled, taking in the ambience of a summer day.
2. Building works, people chatting, glasses clinking with ice in it, people walking on the grass, paper rustling.
3. Apart from the building works, I enjoy the cheerful, laughter and relaxed conversations between people. Also enjoy hearing the glasses clinking as it is making me want to hit the bar!
4. Building works – distracting and depressing in the sense that people have to work on such a nice summer day.
5. The sound of glass – I need a drink / laze about by the bars!

P6

1. The voice around and noise very far like the watery sound.
2. I feel a common wind sound.
3. Sound like a film.
4. Nothing
5. Yeah, exciting

P7

1. The loudest, scraping sounds – digger, wind sound most prominent. There is a lot of texture to the sounds, the sound sharper, scraping etc.
2. Muted, more muffled than normal – seems to be all at similar volume.
3. The crispness of sound, and also the sound sounds like you are listening to so0mething like an orchestra, hearing sounds coming from different directions.
4. The close-up sounds – talking, he instructions were more difficult to hear.
5. Clearness of sound and 3d nature – how it pricks up from all around you that possibly I might not normally hear.

P8

1. The more prominent ones are quite jarring, but it has the feel of everyday life and activity going on.
2. The background noise is more regular and quite calming.
3. It sounds like all the elements of life.
4. Can't think of anything – sorry!
5. The background sound has the most effect – it's quite soothing and regular.

P9

1. Constant drone. Revving of engine annoying.
2. Before the nearby work started the soundscape was relaxing, soothing and an enjoyable break from work.
3. The stillness and tranquillity of the garden and nearby canal (apart from when the construction work started).
4. No nearby workmen!
5. Tranquillity and relaxation.

P10

1. Children's crys. Distant children's speech. Quite a bit of hiss (electrically induced?)
2. Similar to real life. Spaced away from me.
3. Feels real. That's eiry??
4. Add some birdsong – and perhaps a Bach cantata!
5. The wind blast is annoying.

P11

1. Very muffled and high passed
2. The background seems very surreal.

3. Being able to hear something from a distance but seems like its next to you.
4. -
5. Interesting as it seems surreal and like an outer body experience as well as creating more of a 3D experience.

P12

1. Everything sounds high passed with the talking sounding more prominent.
2. -
3. The stereo image
4. Add lower frequencies
5. The music in the background is very audible.

P13

1. Footsteps and snippets of passing conversations. Plates and glasses clatter from cafes cut through the background hum of chatter. Has a slow paced nature (it's lunchtime)
2. Feels busy and ambient. Very non-specific in terms of country/language, no desernable words not even from tram tannoy.
3. The non-specific nature described above (as a drama editor I like that a lot!)
4. Nothing. Its relaxing but clearly urban.
5. Non-specitivity of the location.

P14

1. The more promt soundfield I can hear is the leaves blowing / or should I say crackling noises eg plastic bag.
2. People talking in the background, music playing in the bars.
3. I like the atmosphere, it sounds nice and relaxing. I can hear the wind blowing.
4. Maybe add more low end.
5. It's the overall sound of the area we're in, relaxing.

P15

1. Voices more prominent… less low frequencies
2. General mish-mash of higher frequencies

3. Easier to pick out speech, can hear a lot more going on… feel more aware of more sounds
4. Hard to accurately locate where a sound is coming from (source)
5. Trees rustling in the wind… almost 'tickles'

P16

1. Slightly harsh/noisy with more prominent high frequency noise. Lots of higher frequencies? Leaves / voices more prominent.
2. White noise-esque… not much clarity
3. The 3D environment itself
4. Unsure
5. Strange to hear sounds coming from a different location to where you would expect.

7.5 Appendix C

7.5.1 A comparative view of the semantic differentials' soundscape dimensions from all three phases

	In situ				Ambisonic responses				Binaural responses				In situ real-time binaural			
	Component				Component				Component				Component			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
% of Variance	22.2	21.3	13.9	12.7	23.3	15.7	15.9	12.2	25.8	16.0	12.6	18.3	24.8	24.4	21.1	12.5
Comfortable_Uncomfortable	**0.728**	0.318	-0.057	-0.218	**0.726**	-0.141	0.133	0.178	**-0.370**	0.199	-0.260	0.834	-0.214	**0.886**	0.004	-0.205
Noisy_Quiet	-0.071	0.474	**0.676**	-0.183	0.008	**0.495**	0.402	0.068	**0.639**	0.116	0.211	-0.254	**0.805**	-0.288	0.125	0.365
Unpleasant_Pleasant	**-0.847**	-0.305	0.031	0.394	-0.238	-0.040	**-0.699**	0.293	0.328	-0.217	0.416	**-0.742**	**0.754**	-0.509	0.091	-0.007
Interesting_Boring	-0.200	**0.815**	0.071	0.189	0.394	0.306	**0.437**	0.352	0.122	**0.888**	0.093	-0.018	-0.003	**0.641**	0.062	0.511
Natural_Artificial	0.332	0.296	**-0.612**	-0.113	0.341	0.249	-0.104	**0.678**	-0.041	0.131	-0.475	**0.518**	**-0.941**	-0.021	0.038	-0.124
Like_Dislike	0.542	**0.572**	-0.198	-0.087	0.490	0.029	**0.588**	0.367	0.112	**0.831**	-0.220	0.442	-0.316	**0.907**	-0.100	-0.106
Gentle_Harsh	0.265	0.094	**-0.571**	0.024	**0.903**	0.090	0.063	-0.111	-0.323	**0.781**	-0.107	0.265	**-0.859**	0.208	-0.327	-0.008
Hard_Soft	0.141	0.186	**0.857**	0.050	**-0.830**	-0.119	0.146	-0.029	**0.777**	0.161	-0.024	0.040	0.333	-0.023	**0.847**	-0.368
Fast_Slow	**0.781**	-0.153	-0.220	0.238	-0.172	0.123	**0.496**	0.329	**0.727**	0.335	0.305	-0.302	0.001	-0.368	**0.769**	0.208
Sharp_Flat	0.541	-0.235	**0.625**	0.116	-0.072	**0.851**	0.054	0.293	**0.962**	-0.155	0.081	0.032	-0.089	0.278	**0.758**	0.400
Directiona_Universal	**0.588**	-0.285	0.036	-0.522	0.259	**0.838**	-0.009	0.044	0.007	-0.007	**0.600**	0.002	-0.011	0.067	0.564	**0.617**
Varied_Simple	-0.019	0.112	0.201	**0.887**	-0.076	**0.538**	0.496	0.373	0.536	**0.600**	0.107	-0.212	0.574	**0.739**	0.023	0.210
Reverberant_Anechoic	**-0.796**	0.245	-0.021	-0.151	0.504	0.137	**0.539**	-0.141	0.093	0.111	**0.395**	-0.120	-0.173	0.069	-0.019	**-0.949**
Far_Near	-0.052	-0.076	-0.017	**0.513**	0.116	-0.131	0.132	**-0.800**	-0.105	0.180	**-0.804**	-0.009	0.234	**-0.483**	-0.149	-0.014
Social_Unsocial	-0.172	**0.607**	-0.060	0.324	0.387	-0.424	0.165	**0.511**	0.622	-0.048	0.194	**0.704**	**0.856**	0.223	0.193	-0.179
Meaningful_Meaningless	0.509	0.241	-0.277	**0.672**	-0.027	-0.070	**0.876**	-0.184	-0.232	0.013	0.438	**0.684**	0.274	**0.826**	-0.203	0.121
Calming_Agitating	0.181	**0.898**	0.060	-0.039	**0.848**	-0.086	0.165	0.056	**-0.681**	0.314	-0.207	0.453	-0.283	0.393	**-0.800**	0.174
Smooth_Rough	0.074	**0.907**	-0.030	-0.117	0.590	**-0.651**	0.016	0.210	**-0.763**	0.090	0.381	0.346	-0.217	-0.315	**-0.847**	-0.193

REFERENCES

8 REFERENCES

[1] R. M. Schafer, *The Soundscape: Our Sonic Environment and the Tuning of the World*. Inner Traditions / Bear & Co, 1993.

[2] M. M. Smith, *Hearing History: A Reader*. University of Georgia Press, 2004.

[3] H. Westerkamp, "Linking soundscape composition and acoustic ecology," *Organised Sound*, vol. 7, no. 01, pp. 51–56, Sep. 2002.

[4] M. J. Epstein, "Growing an Interdisciplinary Hybrid: The Case of Acoustic Ecology," 2003. [Online]. Available: http://www.ucalgary.ca/hic/files/hic/epsteinpdf.pdf. [Accessed: 12-Nov-2014].

[5] B. Truax, "Sound in Context : Acoustic Communication and Soundscape Research," *Communication*, no. 1987, pp. 1–11, 2007.

[6] B. Truax, *Acoustic Communication, Volume 1*. Greenwood Publishing Group, 2001.

[7] M. Zhang and J. Kang, “Towards the evaluation, description, and creation of soundscapes in urban open spaces,” *Environ. Plan. B Plan. Des.*, vol. 34, no. 1, pp. 68–86, 2007.

[8] B. Schulte-Fortkamp and K. Voigt, “Why soundscape? The new approach to ‘measure’ quality of life,” *J. Acoust. Soc. Am.*, vol. 131, no. 4, pp. 3437–3437, Apr. 2012.

[9] F. Mossberg, “Getting it together—Interdisciplinary sound environment research,” *J. Acoust. Soc. Am.*, vol. 130, no. 4, p. 2530, Oct. 2011.

[10] R. Kull, “Natural and urban soundscapes: the need for a multi-disciplinary approach,” *Acta Acust. united with Acust.*, vol. 92, no. August, pp. 898–902, 2006.

[11] B. Schulte-Fortkamp, “The Meaning of Annoyance in Relation to the Quality of Acoustic Environments.,” *Noise Health*, vol. 4, no. 15, pp. 13–18, Jan. 2002.

[12] C. Lavandier, C. Drake, and S. Viollon, “Influence of visual setting on sound ratings in an urban environment,” *Appl. Acoust.*, vol. 63, no. 5, pp. 493–511, May 2002.

[13] R. G. Oldfield, “The analysis and improvement of focused source reproduction with wave field synthesis,” Ph.D. Thesis, University of Salford, 2013.

[14] D. Sonnenschein, *Sound Design*. Michael Wiese Productions, 2001.

[15] L. Russolo, “The Art of Noise,” 1913.

[16] B. Shulz, *Resonanzen : Aspekte der Klangkunst*. Kehrer: Heidelberg, 2002.

[17] A. Engström and Å. Stjerna, “Sound Art or Klangkunst? A reading of the German and English literature on sound art,” *Organised Sound*, vol. 14, no. 01, p. 11, Mar. 2009.

[18] B. Truax, “Soundscape, Acoustic Communication and Environmental Sound Composition,” *Comtemporary Music Rev.*, vol. 15, no. 1–2, pp. 49–65, 1996.

[19] C. Sharp, J. Woodcock, E. Peris, G. Sica, A. Moorhouse, and D. Waddington, “Analysis of railway vibration signals using supervised machine learning for the development of exposure-response relationships,” in *Proceedings of Meetings on Acoustics*, 2013, vol. 19, no. 1, pp. 040107–040107.

[20] D. de Ruiter, “Reclaiming land from urban traffic noise impact zones - the great canyon,” Technical University Delft, 2004.

[21] S. McAdams and E. Bigand, *Thinking in sound: The cognitive psychology of human hearing*. Clarendon, 1993.

[22] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, 1994.

[23] S. Handel, *Listening*. Cambrid: MIT Press, 1989.

[24] T. Ackroyd and A. Catalani, “Inheriting slavery: making sense of a difficult heritage,” *J. Herit. Tour.*, vol. 8, no. 4, pp. 337–346, Nov. 2013.

[25] J. A. Ballas and J. H. Howard Jr, “Interpreting the language of environmental sounds,” *Environ. Behav.*, vol. 19, no. 1, pp. 91–114, 1987.

[26] B. H. Repp and G. Knoblich, “Action can affect auditory perception,” *Psychol. Sci.*, vol. 18, no. 1, pp. 6–7, 2007.

[27] A. Corbin, *Village bells: Sound and meaning in the 19th century French countryside*. New York: Columbia University Press, 1998.

[28] S. Htouris, “A comparative interpretation of soundscape and noise,” in *International Congress on Acoustics*, 2001.

[29] S. R. Payne and C. Guastavino, “Measuring the perceived restorativeness of soundscapes: is it about the sounds, the person, or the environment?,” in *Inter-Noise 2013*, 2013, pp. 1–7.

[30] W. Gaver, “What in the world do we hear? an ecological approach to auditory event perception,” *Ecol. Psychol.*, vol. 5, pp. 1–29, 1993.

[31] S. Payne, “Research into the Practical and Policy Applications of Soundscape Concepts and Techniques in Urban Areas (NANR 200),” DEFRA, 2009.

[32] M. Haverkamp, “Essentials for description of cross-sensual interaction during perception of a complex environment,” in *Inter-noise*, 2007.

[33] T. Ingold, “The temporality of the landscape,” *Society*, vol. 25, no. 2, pp. 152–174, 2012.

[34] M. I. Posner, N. J. Nissen, and R. M. Klein, “Visual dominance: an information-processing account of its origins and significance,” *Psychol. Rev.*, vol. 82, no. 2, pp. 157–171, 1976.

[35] I. Frissen, J. Vroomen, B. de Gelder, and P. Bertelson, “The aftereffects of ventriloquism: generalization across sound-frequencies.,” *Acta Psychol. (Amst).*, vol. 118, no. 1–2, pp. 93–100, 2005.

[36] J. E. Lugo, R. Doti, W. Wittich, and J. Faubert, “Multisensory integration: central processing modifies peripheral systems,” *Psychol. Sci.*, vol. 19, no. 10, pp. 989–997, 2008.

[37] H. McGurk and J. Macdonald, “Hearing lips and seeing voices,” *Nature*, vol. 264, no. 5588, pp. 746–748, Dec. 1976.

[38] A. Kohlrausch and S. van de Par, *Communication Acoustics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[39] J. Heron, D. Whitaker, and P. V McGraw, “Sensory uncertainty governs the extent of audio-visual interaction.,” *Vision Res.*, vol. 44, no. 25, pp. 2875–84, Nov. 2004.

[40] V. Maffiolo, M. Castellengo, and D. Dubois, “Qualitative judgements of urban soundscapes,” in *Inter-Noise*, 1999.

[41] J. Carles, I. Lopez Barrio, and J. de Lucio, “Sound influence on land values,” *Landsc. Urban Plan.*, vol. 43, no. 4, pp. 191–200, 1999.

[42] J. Carles, F. Bernaldez, and J. de Lucio, “Audio-visual interaction and soundscape preferences,” *Landsc. Res.*, vol. 17, no. 2, pp. 52–56, 1992.

[43] S. Viollon, C. Lavendier, and C. Drake, “Influence of visual setting on sound ratings in an urban environment,” *Appl. Acoust.*, vol. 63, no. 5, pp. 493–511, 2002.

[44] S. Marsella, A. Way, and M. Rey, “Modeling Coping Behavior in Virtual Humans: Don t Worry, Be Happy,” in *International Joint Conference on Autonomous Agents and Multiagent Systems*, 2003.

[45] C. A. Smith and L. D. Kirby, “Consequences Require Antecedents,” in *Feeling and Thinking: The Role of Affect in Social Cognition*, Cambridge University Press, 2001, p. 83.

[46] S. C. Marsella and J. Gratch, “EMA: A process model of appraisal dynamics,” *Cogn. Syst. Res.*, vol. 10, no. 2000, pp. 70–90, 2009.

[47] I. J. Roseman, “Cognitive determinants of emotion: A structural theory.,” *Rev. Personal. Soc. Psychol.*, vol. 5, pp. 11–36, 1984.

[48] I. J. Roseman, “Appraisal Determinants of Emotions: Constructing a More Accurate and Comprehensive Theory,” *Cogn. Emot.*, vol. 10, no. 3, pp. 241–278, May 1996.

[49] J. Blauert and U. Jekosch, “Sound-Quality Evaluation: A Multi-Layered Problem,” *Acta Acust. United with Acust.*, vol. 83, no. 5, pp. 747–753, 1997.

[50] U. Jekosch, “Basic Concepts and Terms of ‘Quality’, Reconsidered in the Context of Product-Sound Quality,” *Acta Acust. united with Acust.*, vol. 90, pp. 999–1006, 2004.

[51] A. Zeitler and J. Hellbrueck, “Sound quality assessment of everyday-noises by means of psychophysical scaling,” in *Inter-Noise*, 1999.

[52] A. Fiebig, B. Schulte-Fortkamp, and K. Genuit, “New options for the determination of environmental noise quality,” in *Inter-Noise*, 2006.

[53] K. Foale and W. J. Davies, “A listener-centred approach to soundscape evaluation,” in *Institute of Acoustics*, 2012.

[54] C. Mydlarz, I. Drumm, and T. Cox, “Application of novel techniques for the investigation of human relationships with soundscapes,” *Society*, pp. 1–7.

[55] A. Drechsler, H. Raffaseder, and B. Rubisch, “Klang.Reise,” in *Audio Mostly Conference on Interaction with Sound*, 2012, pp. 44–46.

[56] H. Coolican, *Research methods and statistics in psychology.*, 3rd ed. London: Hodder & Stoughton, 1999.

[57] R. Penwarden, “Comparing Closed-Ended and Open-Ended Questions - FluidSurveys,” *FluidSurveys*, 2013. [Online]. Available: http://fluidsurveys.com/university/comparing-closed-ended-and-open-ended-questions/. [Accessed: 17-May-2015].

[58] G. Brambilla, L. Maffei, M. Di Gabriele, and V. Gallo, “Merging physical parameters and laboratory subjective ratings for the soundscape assessment of urban squares.,” *J. Acoust. Soc. Am.*, vol. 134, no. 1, pp. 782–90, Jul. 2013.

[59] J. Kang and M. Zhang, “Semantic differential analysis of the soundscape in urban open public spaces,” *Build. Environ.*, vol. 45, no. 1, pp. 150–157, Jan. 2010.

[60] M. Zhang and J. Kang, “Evaluation of urban soundscape by future architects,” *J. Acoust. Soc. Am.*, vol. 115, no. 5, p. 2497, May 2004.

[61] J. Ge and K. Hokao, “Applying the methods of image evaluation and spatial analysis to study the sound environment of urban street areas,” *J. Environ. Psychol.*, vol. 25, no. 4, pp. 455–466, Dec. 2005.

[62] D. a. Hall, A. Irwin, M. Edmondson-Jones, S. Phillips, and J. E. W. Poxon, “An exploratory evaluation of perceptual, psychoacoustic and acoustical properties of urban soundscapes,” *Appl. Acoust.*, vol. 74, no. 2, pp. 248–254, Feb. 2013.

[63] Ö. Axelsson, M. E. Nilsson, and B. Berglund, “A principal components model of soundscape perception,” *J. Acoust. Soc. Am.*, vol. 128, no. 5, pp. 2836–46, Nov. 2010.

[64] A. Mehrabian and J. A. Russell, “A Conceptual Framework,” in *An Approach to Environmentnal Pschology*, Cambridge, MA: MIT Press, 1974, pp. 1–9.

[65] A. Chan and A. Noble, *Sounds in Translation: Intersections of Music, Technology and Society*. ANU E Press, 2009.

[66] D. Dubois, “Categories as acts of meaning: The case of categories in olfaction and audition,” *Cogn. Sci. Quaterly*, pp. 35–68, 2000.

[67] M. Raimbault, “Qualitative judgements of urban soundscapes: Questionning questionnaires and semantic scales,” *Acta Acust. united with Acust.*, vol. 92, pp. 929–937, 2006.

[68] M. Raimbault, “Qualitative Judgements of Urban Soundscapes : Questionning Questionnaires and Semantic Scales,” *Acta Acust. united with Acust.*, vol. 92, pp. 929–937, 2006.

[69] W. J. Davies and J. E. Murphy, “Reproducibility of soundscape dimensions,” in *Inter-Noise*, 2012.

[70] C. Guastavino, B. F. G. Katz, J. Polack, and D. J. Levitin, “Ecological Validity of Soundscape Reproduction,” *Acta Acust. united with Acust.*, vol. 91, no. September 2004, pp. 333–341, 2005.

[71] D. Dubois and C. Guastavino, “A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories,” *Acta Acust. United with*, vol. 92, pp. 865–874, 2006.

[72] V. Maffiolo, “Caractérisation sémantique et acoustique de la qualité sonore de l’environnement urbain,” Ph.D. Thesis, Université du Maine, 1999.

[73] D. Dubois, “A Cognitive Approach to Urban Soundscapes : Using Verbal Data to Access Everyday Life Auditory Categories,” *Acta Acust.*, vol. 92, pp. 865–874, 2006.

[74] P. Power, C. Dunn, W. Davies, and J. Hirst, “Localisation of elevated sources in higher-order ambisonics,” *BBC R&D Publications*, 2013. [Online]. Available: http://www.bbc.co.uk/rd/publications/whitepaper261. [Accessed: 07-Nov-2014].

[75] C. Hartshorne, P. Weiss, and Q. W. Burks, Eds., *The Collected Papers of Charles Sanders Peirce*. Cambridge, MA: Harvard University Press.

[76] W. James, “Philosophical Conceptions and Practical Results,” *Univ. Chronical, Univ. Calif.*, vol. 1, no. 4, p. 290, 1898.

[77] W. James, “Pragmatism’s conception of truth,” in *Pragmatism*, New York: Longmans, 1907, pp. 197–236.

[78] J. W. Creswell, *Research design*. Thousand Oaks: Sage, 2003.

[79] D. M. Mertens, *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*, 2nd ed. Thousand Oaks: Sage, 2004.

[80] N. Mackenzie and S. Knipe, "Research dilemmas: Paradigms, methods and methodology," *Issues Educ. Res.*, vol. 16, 2006.

[81] A. Tashakkori and T. Charles, *Handbook of Mixed Methods in Social & Behavioral Research*. London: Cassel, 2002.

[82] B. Somekh and C. Lewin, *Research Methods in Social Sciences*. London: Sage, 2005.

[83] S. Hawking and L. Mlodinow, *The Grand Design*. London: Bantam, 2010.

[84] R. Burke Johnson, "Mixed Research: Mixed Method and Mixed Model Research," *University of South Alabama*. [Online]. Available: http://www.southalabama.edu/coe/bset/johnson/lectures/lec14.htm. [Accessed: 10-Oct-2014].

[85] J. M. Condie, "Living Alongside Railways : A Discursive Psychological Analysis of Adapting to Disruption and Identities of Place," Ph.D. Thesis, University of Salford, 2013.

[86] N. Bruce, "The effects of expectation on the perception of soundscapes," Ph.D. Thesis, University of Salford, 2011.

[87] R. Nicol and M. Emerit, “3D-Sound Reproduction Over an Extensive Listening Area: A Hybrid Method Derived from Holophony and Ambisonic,” in *Audio Engineering Society Conference: 16th International Conference*, 1999.

[88] F. Hollerweger, “Periphonic Sound Spatialization in Multi-User Virtual Environments,” Ph.D. Thesis, USCB, 2006.

[89] C. Guastavino and B. F. G. Katz, “Perceptual evaluation of multi-dimensional spatial audio reproduction,” *J. Acoust. Soc. Am.*, vol. 116, no. 2, p. 1105, 2004.

[90] A. McKeag and D. S. McGrath, “Sound Field Format to Binaural Decoder with Head Tracking,” in *Audio Engineering Society Convention 6r*, 1996.

[91] E. Y. Choueiri, “Optimal Crosstalk Cancellation for Binaural Audio with Two Loudspeakers,” *Posted on the 3D3A Lab’s website*, 2008. [Online]. Available: https://www.princeton.edu/3D3A/Publications/BACCHPaperV4d.pdf. [Accessed: 23-Aug-2014].

[92] N. Bruce, W. Davies, and M. Adams, “Expectation as a factor in the perception of soundscapes,” in *Euronoise*, 2009.

[93] M. Mitchell and J. Jolley, *Research Design Explained*, 8th ed. New York: Cengage Learning, 2012.

[94] A. S. Sudarsono, “Currently Unpublished Work,” University of Salford, 2014.

[95] M. Chion, *Audio- Vision: Sound on Screen*. New York: Columbia University Press, 1994.

[96] P. Bertelson, *Cognitive Contributions to the Perception of Spatial and Temporal Events*, vol. 129. Elsevier, 1999.

[97] B. Smith and D. W. Smith, *The Cambridge Companion to Husserl*. Cambridge: Cambridge University Press, 1995.

[98] T. H. Pedersen, *The Semantic Space of Sounds*, vol. 1, no. May. 2008.