The Development of a Model of Audio Comfort in an Aviation Context

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Master of Philosophy

By Max Alexander Ffolliott Aldridge

Date: 24/10/2018

Abstract

The Development of a Model of Audio Comfort in an Aviation Context – M.Aldridge

The aim of this research was to develop a model of audio comfort. This model was developed to facilitate the comparison of different stimuli and their impact upon participants' comfort.

The thesis provides details of the human factors research undertaken as part of an Innovate UK funded project **Im**proving the **P**ropulsion Aerodynamics and **Ac**oustics of **T**urboprop **A**ircraft (ImPAcTA) with Dowty Propellers. The thesis discusses the usage of the data obtained from this project in the development of a model of audio comfort.

The thesis describes the initial stages in which the impact on cognitive performance and comfort are assessed in context of the physical properties of noise (Chapter 3). In this initial stage the results showed that there was a significant impact of both spectral content and decibel level on cognitive performance and reported discomfort. These results confirmed the efficacy of stimuli specific cognitive performance tasks in the measuring of discomfort when paired with a comfort questionnaire.

With the usefulness of cognitive performance tasks being supported by the research carried out in Chapter the context in which noise was experienced was then examined (Chapter 4). To manipulate the context in which the noise was perceived the experiment varied the fidelity of a flight simulator. The results from this study showed that the reported immersion and presence participants experienced was related to both the comfort reported and task performance. The results demonstrated the need to consider the physical stimulus properties as well as environmental context and presentation methodologies in the assessment of comfort. The results also showed a significant impact of simulation fidelity on the assessment of comfort for an audio stimulus with consistent physical properties.

The proposed model in this thesis is based on Rolls (1990) model of emotion. This model posits that there are parallels between comfort and emotion as defined by Rolls. In this model these parallels are based on the concepts of; elicitation of autonomic and endocrine responses, requirements to learn new and flexible behavioural responses to avoid/attain discomfort/comfort, motivation to take action, and facilitating communication. In the proposed model, comfort is represented as the centre of an x,y axis, movement on these axis are caused by the introduction or removal of primary or secondary reinforcers. Primary reinforcers refer to the experience of the discomfort stimuli while secondary reinforcers refers to the expectations a participant has of the stimuli and surrounding context. The impact of these reinforcers can be measured by cognitive task performance for primary reinforcers and perceived annoyance for secondary reinforcers. This model provides a novel method of predicting comfort that can be objectively measured and tested. The models assumptions were tested with use of both data collected previously (Chapter 4) and data from a new large scale study. Through the use of this data the assumptions of the model were confirmed with the model providing useful predictions of comfort from the metrics (Speech Identification Task, comfort questionnaire and Tanker Tracking Task) used.

Acknowledgements

I would like to express my gratitude to:

Dowty Propellers for giving me the opportunity to be a part of such an interesting and exciting project. Nicholas Bown and Alison Starr for their kind guidance and expansive expertise in my introduction to industry.

My primary supervisor Dr Mark White whose patience, kindness, understanding, and dedication were endless and always appreciated.

My secondary supervisor Dr Georg Meyer whose relaxed and calm demeanour helped to offset my constant panic.

Thank you both for your support, bringing levity to every meeting, and punctuation to my sentences.

Rolf Zon, Olaf Brouwer, and Christopher Roos at the NLR. Dank je vel for working with me on the ImPAcTA project it was a wonderful experience and a privilege to work with you. Thank you for making me feel so welcome on my trips to see you.

Jack Carter-Hallam, you have done so much to help me through the past four years, through every crisis you've been there with a tissue, kind words, and a plan of action.

The Post-Graduates in the Engineering and Psychology departments, it was wonderful working with you.

My mother Emma, thank you for being understanding as I ramble about statistics and for all the support you've given me over the years.

My partner Alice, thank you for convincing me that breaks are something I can take, that not every set-back is a catastrophe, and for so much more. I couldn't have done it without you.

Declaration

I confirm that the thesis is my own work. That I have not presented anyone else's work as my own and that full and appropriate acknowledgement has been given where reference has been made to the work of others.

Max Aldridge

October 2018

Contents

Table of Figures

List of Tables

Notation

dB – Decibel

dB(A) – A-weighted decibels

dB(C) – Decibels relative to the carrier

DoF – Degrees of Freedom

e-/+ - exponent

Hz – Hertz

n – Number of participants

ms – Millisecond

p – Probability Value

SD –Standard Deviation

SDE – Standard error

r – Correlation Coefficient

RPK - Revenue Passenger Kilometres

t – t-statistic

Abbreviations

ANOVA – Analysis of Variance

BTA – Brief Test of Attention

CAA - Civil Aviation Authority

GUI – Graphical User Interface

HSE – British Health and Safety Executive

ICAO - International Civil Aviation Organization

IATA - International Air Transport association

ImPAcTA – Improving the Propulsion Aerodynamics and Acoustics of Turboprop Aircraft

MRI – Magnetic Resonance Imaging

PEP – Phillips Effect Paradigm

PI – Place Illusion

Psi – Plausibility Illusion

RPK - Revenue Passenger Kilometres

RPvdsEx – Real-time Processor Visual Design Studio

SIT – Speech Identification Task

SNR – Signal to Noise Ratio

VEPR – Visually Evoked Postural Responses

VR – Virtual Reality

WHO –World Health Organization

Chapter 1 Introduction

1.0 Background to Aviation Noise Problems

The airline industry is continuing to grow, with revenues increasing from $413 billion in 2005 to $718 billion in 2015 (International Air Transport association (IATA), 2016). As airlines have become more accessible, passenger numbers have also increased, with the IATA (2016) report showing increases in the Revenue Passenger Kilometres (RPK) increasing globally from 6.3% in 2011 to 7.4% in 2015. This increase in passenger numbers has an additional impact, noise pollution. Aircraft traffic has been linked to noise pollution by numerous different research bodies such as the World Health Organization (WHO, 1995 and 2011). Noise pollution has been found to have significant impact on physical and mental health (Civil Aircraft Authority, 2016) (CAA). These impacts, the CAA states, include sleep loss, hearing loss, cardiovascular disease, and annoyance as illustrated in Fig. 1.

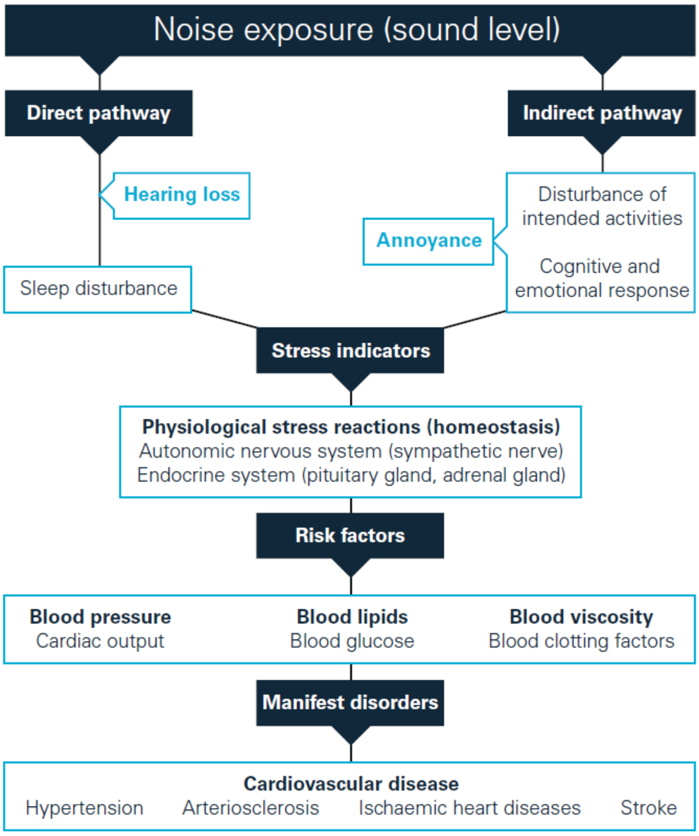


Figure 1: Noise exposure chart taken from CAA report 2016

Due to this impact the International Civil Aviation Organization (ICAO), as the world's aviation body, has set out the Balanced Approach for noise reduction (Fig 2). This approach consists of four aspects:

1) Reduction of noise at the source – informing the design of the aircraft ensuring that noise reduction methods are in place.

2) Land-use Planning and Management – detailing methods of reduction of the impact of aircraft noise on the population, in particular around airports.

3) Noise Abatement Operational Procedures – implementation of operational procedures which are low noise.

4) Operating Restrictions – These restrictions detail time of day for flights, as well as the phasing out of particularly high-noise aircraft.

With this set of guidelines set out by the ICAO it is clear that noise reduction is an important area of interest within aircraft design and the industry as a whole. Shown in Table 1 is the ICAO noise reduction goals for various categories of aircraft, which shows a significant short-term reduction requirement.

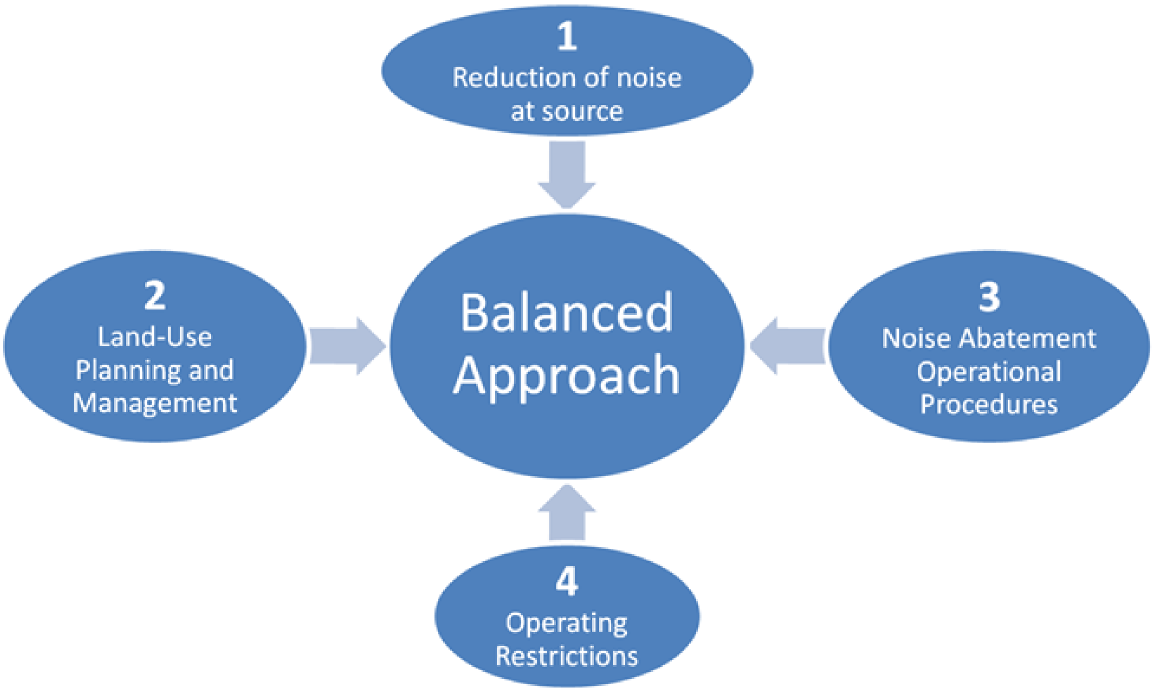


Figure 2: ICAO Balanced approach for aircraft noise reduction

Table 1: ICAO Noise Reduction Goals

Aircraft Category	Noise Reduction Goals (dB)	
	Mid-Term (2018)	Long-Term (2028)
Regional Jet	13.0±4.6	20.0±5.5
Small-Med. Range Twin	21.0±4.6	23.5±5.5
Long-Range Twin	20.5±4.6	23.0±5.5
Long-Range Quad	20.0±4	23.5±5.5

With the importance of aircraft noise reduction to the industry, it is no surprise that it is a common theme in aircraft research and design. The report produced by the CAA in 2016 discusses the human factors and medical impacts of aircraft noise. The CAA report details an increase in annoyance and related hypertension, cardiovascular diseases, negative impacts on child learning, and sleep disturbances, resulting from noise pollution (Fig 1).

The acoustic dimensions of propeller noise are typically made up of harmonically related pure tone components superimposed upon broadband noise. The controlling mechanisms for this aspect of propeller noise is cited by McCurdy (1988) as blade passage frequency, blade helical-tip Mach number, engine, and airframe noise.

Ivosevic, Mihalincic and Bucak (2010) note in their analysis that piston engines for light aircraft produce levels of noise with a minimum of 75 dB. They suggest that this is partially due to their air cooled rather than, as they recommend, liquid cooled engine systems. The majority of literature which focuses on aircraft noise reduction or negation tends to be on exterior aircraft noise. Examples of these include previously mentioned research on the impact of air traffic on residential areas from the WHO (1995, 2011) and CAA (2016). This focus can be seen in the ICAO balanced approach for aircraft noise reduction (Fig 2). In this approach three of the four points are focused on external noise and its impacts, while only "Reduction of Noise at Source" really highlights internal noise.

Studies such as Mellert, Baumann, Freese, Weber (2008) show reduced comfort, increased physical stress reactions (such as increased heart rate), and decreased task performance in participants exposed to noisy cabin environments (experimental range from 70 dB to 75 dB). With studies like Mellert et al. (2008) it is shown that reduction of cabin noise is beneficial to both passenger comfort, and aircraft companies; producing more "appealing" products and better staff health and performance. There have been attempts to reduce interior cabin noise through changes to the structure of the aircraft's fuselage design (Sengupta, 1977, Robinson, Fernholz, 1996). Sengupta (1977), for example, carried out tests on developments from Boeing which used intrinsic tuning and damping of the aircraft's fuselage to reduce interior low frequency cabin noise. Other factors have also been highlighted, such as non-aircraft-originating noises such as verbal communication and air conditioning which were discussed in Ozcan and Nemlioglu (2006). In their paper, passenger and flight-crew activities along with other environmental noise producers were noted to have a significant impact upon aircraft in cabin noise levels. Research into propeller noise reduction technology has been carried out previously but, as Metzger (1995) stated, the research was rarely implemented in aircraft design. This, Metzger states, was due to weight, cost, and performance penalties associated with the products of this research.

Mwanalushi (2012) predicted that turboprop aircraft numbers will increase. Quoted in the article, Rob Morris, a senior aviation analyst at Ascend, claims that turboprop aircraft are experiencing a renewed popularity. He supports his statement with the increase in production and annual order rates of turboprop aircraft. Mwanalushi suggests that a main driver for this popularity is their fuel efficiency and an increase in oil and fuel prices globally. With Mwanalushi stating that turboprop sales are expected to rise to 48% of total regional aircraft purchases, new research into optimising these aircraft has become an industry priority.

Dowty Propellers, a subsidiary of GE Aviation Systems Ltd, is one of the world's leading manufacturers of integrated propeller systems, being one of only two global manufacturers of large composite propeller blades and the associated systems for these aircraft. Through an Innovate UK funded project, **Im**proving the **P**ropulsion Aerodynamics and **Ac**oustics of **T**urboprop **A**ircraft (ImPAcTA) (BIS reference number 110110), Dowty had identified several of aerodynamic and acoustic innovations which could be included in future designs. These innovations have been developed to further improve the efficiency of turboprop aircraft, whilst also attempting to negate any audio discomfort caused by the design changes; thereby increasing market potential. Turboprop aircraft offer up to 30% savings in fuel burn compared to an equivalent turbofan powered aircraft (Whitlow and Sievers 1984). This fuel efficiency already makes the turboprop aircraft an attractive alternative to other competing types of aircraft.

In collaboration with the University of Liverpool (UoL), the National Aerospace Laboratory of the Netherlands (NLR), and the Aircraft Research Association (ARA), Dowty aimed to design, build, and test new turboprop propeller configurations and assess them for fuel efficiency, functionality, and impact upon passenger comfort via noise level produced by these propellers. The ImPAcTA project was focused on the engineering assessment of the Dowty propeller designs using wind tunnel testing, Computational Fluid Dynamics (CFD) and a

Human Factors approach to tackle the issue of cabin noise reduction. For the first time, the re-design of propellers for turbo-propeller aircraft has been undertaken with performance, fuel efficiency, and passenger comfort as the design parameters.

The designs for the propellers were made with efficiency in mind. These new propeller designs, as a by-product of changes in their configuration from a standard propeller configuration, were predicted to have different noise signatures compared to standard designs (see section 1.2). With the knowledge that these propeller noise signatures would be different from that of a standard design Dowty, UoL and NLR undertook research to determine which of these designs would have the least negative impact upon passengers in terms of audio comfort. Whilst passive and active acoustic systems are available for the fuselage to reduce noise transmission into the cabin, the aim of this work was to reduce, where possible, noise problems at the source (i.e. at the propellers). In this research, the UoL and the NLR have examined the simulated audio from the three Dowty propeller configurations (Base, Staggered, Unequal see Figure 3) and the impact of the differences between the propeller configurations on passenger comfort. This research was highlighted in the 2015 European Aerodays innovation event (Dowty 2015).

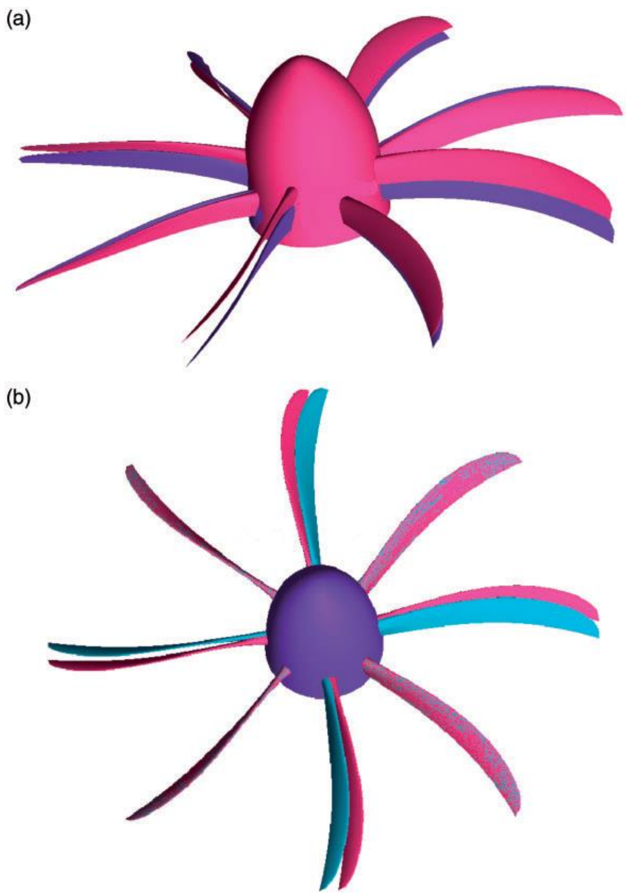


Figure 3: Comparison of (a) Base (blue) and Staggered (pink) and (b) Base (blue) and Unequal (pink) propeller designs

These three propeller configurations known as Base, Staggered and Unequal, were named after their design "styles" (see section 1.2). The Base configuration was based upon a standard propeller configuration, with blades evenly distributed around the central hub. The Staggered configuration had its blades with the same spatial distribution as the baseline configuration but the in-plane positions (i.e. to/from the hub) were moved fore/aft of the baseline positions. The Unequal configuration had an unequal spacing between the blades but are the same plane as the Baseline condition. These designs were developed by GE Dowty in advance of this study.

The University of Liverpool's CFD group produced acoustic simulations of the propeller configurations to generate acoustic signals (Barakos and Johnson, 2016), which were then used by the NLR to generate internal cabin noise files that were used in this thesis for the Human Factors research. Details of the generation of these files is provided in Section 1.2

These audio files were integrated into a test programme used in the Human Factors section of the ImPAcTA project. Participants of the experiments experienced the different sound files, and assessments were made of the impact of the audio signature on cognitive and perceptual experiences. In addition to the transfer function the NLR contributed a study into the comparison of monaural or binaural sound files of the propeller designs. This study was informed by the data gathered in Chapter 3 using the testing metrics tested and selected in Chapter 3. The metrics from this thesis that informed the NLR's study were the Speech Identification Task, and the comfort questionnaire. The NLR also used a forced choice preference task and an extended version of the comfort questionnaire. This experiment was used to determine if there was a significant difference between monaural or binaural propeller sounds. The noise presentation method chosen was monaural as the work from NLR had indicated that providing additional facilities to present binaural stimuli would not significantly benefit the research. The NLR found that there was no significant difference in comfort when audio was presented as monaural or binaural (see NLR technical report NLR-CR-2013-145-RevEd-b).

This thesis addresses the Human Factors element of the ImPAcTA project, developing a methodology to compare and predict passenger auditory comfort with new developments in reduction to cabin noise. The rationale for this research is that although there are several papers looking into passenger comfort (see Chapter 2), there is only a minimal amount of research into this area of interest in turbo-propeller aircraft, and a methodology has not previously been developed to assess passenger comfort at the design stage; this was the focus of this research.

In terms of passenger comfort assessment, human factors research is needed to produce a set of sensitive measures to analyse any changes in perception of the noise both subjectively and objectively. This thesis proposes a model of audio comfort which allows for predictions and comparisons between audio stimuli.

The aim of this thesis was to develop a novel model of audio comfort. This aim was achieved through;

- The identification of effective and sensitive task performance metrics (Chapter 3)
- Assessing the impact of level and spectral differences on comfort and task performance (Chapter 3)
- Assessing the impact of outside factors such as simulator fidelity on comfort and task performance (Chapter 4)
- Testing the model's functionality and assumptions (Chapter 5)

1.1 Structure of Thesis

Chapter 2 discusses the literature surrounding comfort as a concept. This literature discusses the impact of the physical properties of sound on participants' comfort including spectral differences and level. The Chapter then goes on to discuss the methods in which comfort has been measured and the issues and advances that have been made in this area. From this, the literature review then goes on to discuss the potential for comfort to be impacted by factors outside of the physical properties of the measured stimuli. With the literature surrounding comfort discussed the chapter then goes on to detail and discuss several of the current models of comfort.

Chapter 3 examines the primary issue of how to measure comfort through both objective and subjective metrics, using physical changes in audio stimuli as the independent variable such as spectral differences and dB level changes.

Once the comfort assessment metrics have been established in Chapter 3, the methods were then applied to variables which extended beyond the physical properties of sound such as simulator fidelity. The variable beyond the physical properties of sound used in Chapter 4 was simulator fidelity. With this secondary study, the thesis examines the impact of simulator fidelity upon participants' comfort levels.

The metrics developed in Chapters 3-4, were determined to be effective in assessing the impact of physical property change and the context and method of presentation of the audio stimuli upon participant comfort. With confirmation that the metrics used were effective, the results produced were then used to test a new model of audio comfort. The assumptions of this model were then tested in a large-scale cabin simulation study designed to assess the task performance, annoyance, and comfort in participants. This study is detailed in Chapter 5. The model discussed in Chapter 5 provides a method of predicting the impact of changes in an individual’s level comfort, determined by changes in the participants’ task performance and ratings of annoyance. While the model can be used to track changes at an individual level it can also be applied to predict general changes in a wider population through methods discussed within this thesis.

The audio comfort model presented in this thesis and the methods used to develop it not only serve in its primary function of providing a method of predicting changes in audio comfort but also bring up a discussion on the use of modelling and the methods of analysis applied to modelling. The model also fills out its original specification of being able to be applied to industry (in this case Dowty Propellers) to inform the impact of new designs upon humans.

1.2 Propeller Audio Files

Two new propeller designs were developed by Dowty for assessment in this research. Due to changes in their configuration, (Figs 4 and 5), it was anticipated that they would have different noise signatures compared to standard designs. The acoustic signals produced by the three propeller configurations, (Base, Staggered, Unequal), calculated using CFD for a turbo-propeller aircraft operating (Barakos and Johnson, 2016) in a cruise condition were used in the final assessment experiment on passenger comfort.

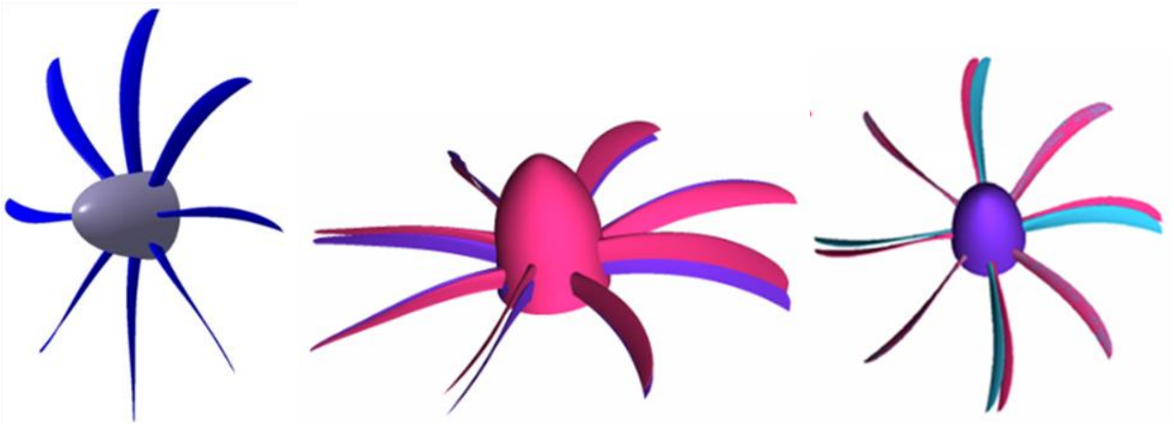


Figure 4: Propeller designs, (left to right) Base, Staggered, and Unequal.

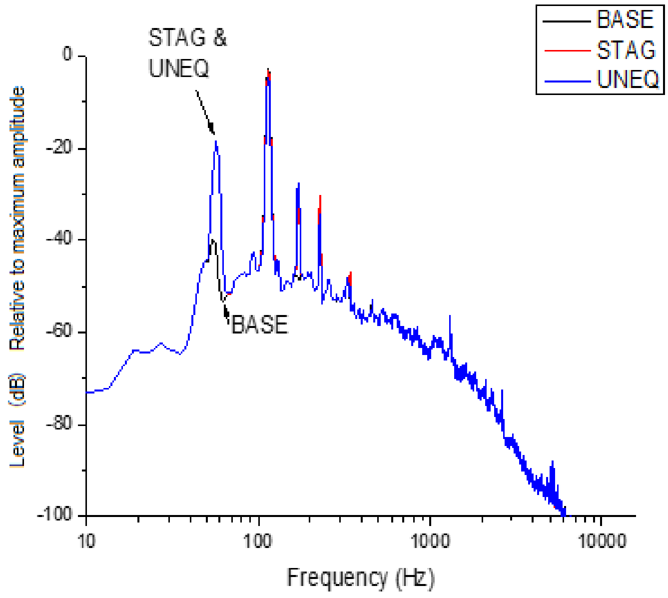


Figure 5: Log spectrum of all three configuration audio files. Y Axis shows attenuation.

Barakos and Johnson (2016) discuss their method in modelling the propeller designs, using fine multi-block grids with the sliding plane method to deal with motion between airframes and the propeller. These grids used 12 million cells per blade in the isolated cases and 50 million cells in the installed cases. To analyse the tones produced by the propellers and produce the CFD results Barakos and Johnson used the Reynolds-Averaged Navier-Stokes method.

The spectral content of the three configurations is shown in Figure 5. It is observed that there is a major difference in the spectrum between BASE (Base) and the two alternative configurations (STAG, UNEQ – Staggered and Unequal) at a frequency of 57Hz (Fig 5). The hearing threshold rapidly increases below 100Hz. The threshold at 60Hz is more than 50dB higher than at 2-4 kHz, so that very loud signals are required to reach threshold at low frequencies. At the intensity used in the subsequent experiments (70 dB and 76 dB) individual components below 100Hz are likely to be inaudible for most participants. Humans are most sensitive in the range between 200 and 10,000Hz, with steep increases in threshold at either end of the spectrum. The signal levels required for discomfort are comparable across the frequency range.

A second difference in the spectra is visible at 170 Hz, where the Staggered and Unequal configurations show a harmonic peak (15 and 20 dB above baseline) which is missing in the reference Baseline configuration. This peak is attenuated by 30 dB relative to the dominant frequency at 114Hz. Increases in energy at this peak are compensated by slightly higher levels at the surrounding peaks (Fig 6).

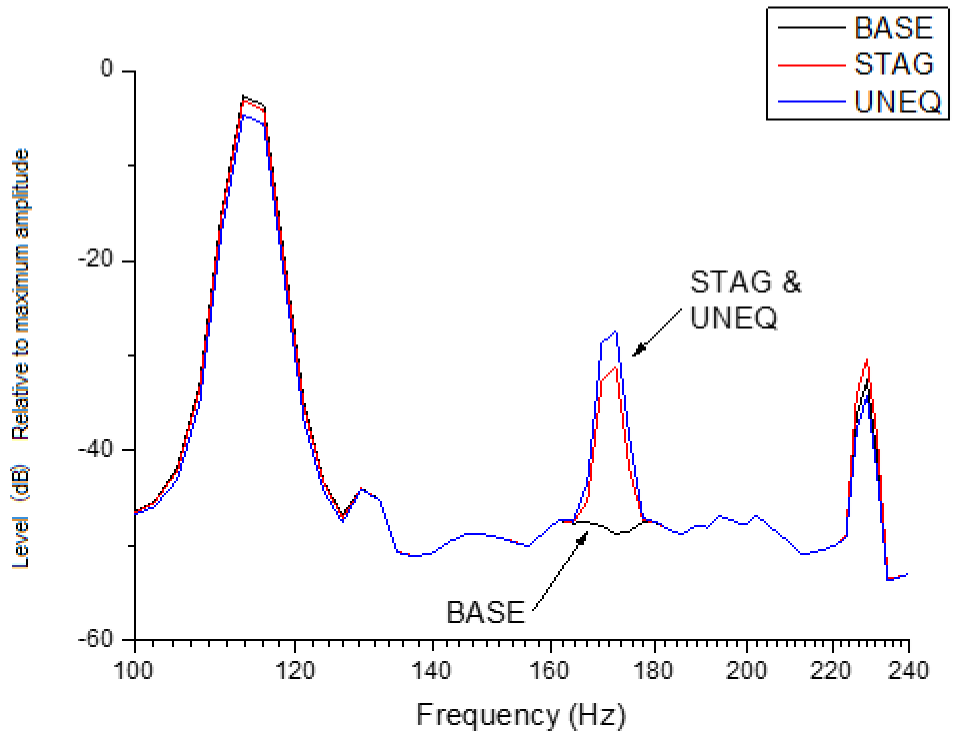


Figure 6: Detailed spectra for the three configurations between 100 and 240Hz. Y Axis shows attenuation.

These sound files were used by the NLR, who, via a transfer function, generated internal cabin noise sound files of a single engine (see NLR technical report NLR-CR-2013-145-RevEd-b). The transfer function used was developed through a testing of the sound signature (produced by the array in Fig 7) changes after passing through the fuselage of a Fokker 50 turboprop aircraft (Fig 7). This was recorded at multiple points in the interior of the aircraft with a set of dodecahedron microphones (Figs 7 and 8). These microphones were situated at key points throughout the aircraft (S1, S2 and S3 on Fig 7) to gain a view of the signature changes throughout the aircraft rather than at just one point. The transfer function was applied to the CFD acoustic signals to produce a sound file for each propeller configuration (Fig 9). The transfer function was carried out through a Fourier transformation of the input data from UoL's CFD, multiplication with the transfer function values, summation over all of the probe positions, and then an inverse Fourier transformation. The files produced by the NLR were not only passed through a transfer function but also had additional elements added to them to provide a realistic cabin environment; these included the input of ambient cabin noise such as wind and air conditioning. These files were used in various iterations throughout the thesis. In Chapter 3, during Experiment 2, an early version of these files from one point on the propeller was used. In Chapter 5 the final sound files including sound profiles from the whole propeller with a transfer function applied were used in the large-scale study. The reason for the difference in sound files used is that the simulation of these propeller sound files was carried out in parallel to this research project and therefore the early single point sound files used in Chapter 3 were those that were available at the time. The sound files used in Chapter 5 were only available later in the research due to production and simulation times.

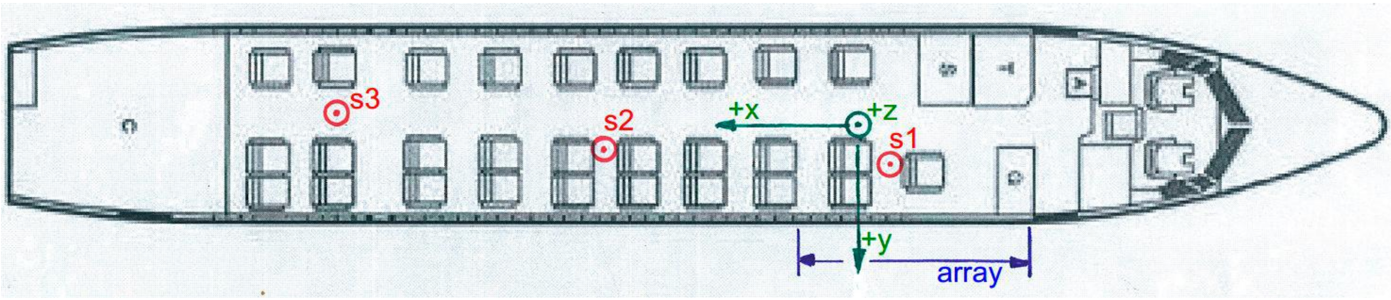


Figure 7: Source positions and array cover for Fokker 50 aircraft used by NLR

Figure 8: Dodecahedron microphone

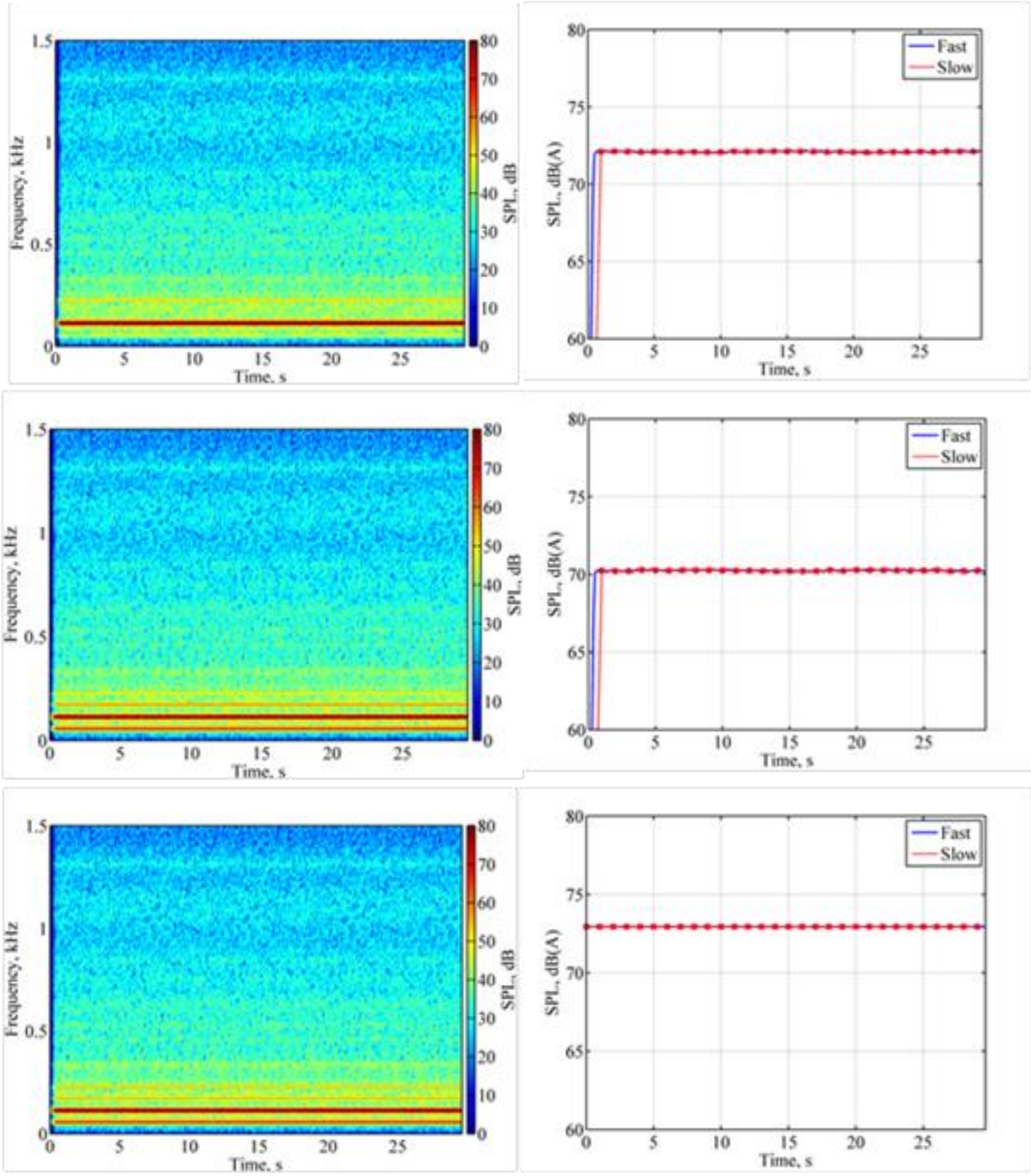


Figure 9: Resulting sound files from transfer function transformation (from top to bottom: Base, Unequal, and Staggered)

1.3 Research Contributions

This thesis provides a demonstration of the use of task performance as a measure of auditory comfort. In addition to this it also carried out novel research into the impact of various levels of fidelity in virtual environments on comfort perception. This research showed the significant impact of fidelity upon comfort, task performance, and annoyance. This study also displayed the importance of separating aspects of fidelity for manipulation within an experimental context. The results of this research were used to develop a new model of audio comfort. This model allows for a framework which can provide predictions of shifts in comfort either across populations using an averaged method or within individual participants. The thesis provides a visual representation of the theoretical aspect of the model, along with an analysis and interpretation method for the application of new data to the model's framework.

The primary beneficiary of the work carried out was Dowty Propellers, as a portion of the work was used to assess three turbo-propeller designs for their impact upon passenger comfort. Further beneficiaries of the work include industries in which user or passenger comfort is an issue in relation to sound discomfort for example automobile manufacturers. The work presented in this thesis would allow for the assessment of the sound output of the product in comparison to potential alternatives. This assessment using the model presented in this thesis, would provide these industries with the capability to tailor their product to cause the least discomfort and therefore increase customer satisfaction.

1.4 Details of presentations/papers

Work in this thesis has been presented at conferences in Europe. The research carried out was presented as a poster at the International Multisensory Research Foundation (IMRF) Conference in 2014 in Amsterdam under the title of "Simulator Fidelity Affects Comfort" and at the IMRF 2015 Conference in Pisa titled "The Effect of Audio and Vibration Stimuli on Comfort and Performance". Research in this thesis was also presented at the European Conference of Visual Perception 2015 as a poster titled "Simulation Fidelity Affects Perceived Comfort".

Work from this thesis has also been presented to Dowty Propellers, NLR, and TSB at the quarterly ImPAcTA meetings and has been incorporated into a report to Dowty Propellers which was used to inform the selection of three different design choices of turboprop propellers.

Chapter 2 Literature Review

This chapter details a wide view of research carried out in relation to comfort and how it is impacted by different stimuli changes. The chapter goes on to discuss the literature relevant to the research and development used in the formulation of the proposed model of audio comfort discussed in Chapter 5. This literature review covers the impact of the physical properties of sound upon comfort and the methods of measuring comfort; further literature then discusses the impact of simulator fidelity upon comfort. In the penultimate section of this chapter, several prominent theories of comfort are covered.

2.1 Comfort

Comfort as a research area spans many different sensory inputs; thermal comfort (Bhiwapurkar, Saran, Goel, Mansfield, and Berg, 2009, Candas and Dufour, 2005), seating comfort (De Looze, Kuijt-Evers, and Van Dieen, 2003), vibration comfort (Melert, Baumann, Freese, and Weber, 2008, De Looze, Kuijt-Evers, and Van Dieen, 2003), and audio comfort (Kahn, and Sunderström, 2006, Zhang, Zhang, Lui, and Kang, 2016).

Research on comfort has been applied to a variety of areas from healthcare (Kolcoba 2001) to industrial design (Ziaran 2014). Comfort research has facilitated patient recovery and well-being practices by informing nursing practices (Kolcoba, 2001). Research into the health impacts of technology such as wind turbines and other producers of infra-sound have also stemmed from comfort research (Leventhall and Knopper, 2015). These are elaborated on in section 2.1.2. Industrial design is another area which highlights the importance of comfort research, this research facilitates more comfortable designs which in turn produce better work performance of employees (Waye, Rylander, Benton, and Leventhall, 1997). These are just a few of the many applications of comfort research being used today.

The research detailed in this thesis has been used to develop a novel model of comfort which can be applied in experimental and industrial assessment of stimuli, environment, or presentation method of stimuli. The proposed model allows for the direct comparison of discomfort stimuli through task performance and participants perception of the stimuli. This is a gap in the literature referring to sound and general comfort models; the model of audio comfort proposed in this thesis fills this gap. The proposed model is novel in several key aspects such as; presenting a new theoretical standpoint which draws on emotion and comfort research, reducing cross-participant variability issues, and using a framework which can display both numerical and abstract data.

2.2. Comfort and Sound

Sound, particularly at high levels, has a negative impact on comfort and therefore is often intentionally reduced or changed; this is particularly true in the case of industrial environments.

While comfort is considered to be a subjective experience, with tolerances and preferences being specific to the individual (Woszcyk, Bech, and Hansen, 1995), there are certain stimuli which are considered to be almost universally uncomfortable, in particular high intensity sound and vibrations (De Looze, Kuijt-Evers, and Van Dieen, 2003). In their model of seating comfort De Looze et al. (2003) argued that high intensity vibrations can lead to notable discomfort. Within these universal factors influencing comfort there is basis for the claim that comfort is affected by a variety of aspects which make up these universal factors or stimuli. These stimuli as Canadas et al. (2005), Yang and Kang (2004), and Jeon, You, Jeong, Kim, and Jho (2011) state can be a variety of different aspects of one variable. An example of this is audio stimuli which may influence comfort due to features such as the sound intensity or pitch. These universal stimuli, such as high intensity noise or vibrations, are considered to be uncomfortable due to their potential to cause damage to sensory organs (National Institute on Deafness and Other Communication Disorders (NIDCD)), and their impact on a person's capability to carry out tasks (Metzger 1994).

The pitch or frequency of a noise has been regularly cited as an important factor in audio comfort. Leventhall (2004), for example, discusses the negative effect of low-frequency noise (10-200Hz). Leventhall states that these low-frequency noises can have an impact upon both stress and annoyance levels. This was found to be particularly true for sensitive participants who showed significant increases in self-reported levels of stress and annoyance. While the WHO in their documentation state there is a greater need to reduce the intensity of an audio signal if it contains low-frequency components known as infra-sound (below 20Hz) due to both annoyance and health risk.

In contrast to the stress and annoyance from low frequency sounds, (Leventhall 2004) high frequency sounds are more closely tied to hearing loss (Salvendy 2012). Reinhold, Kalle, and Paju (2014) note that high frequencies are connected with hearing loss, hypertension, and fatigue. Reinhold et al. (2014) also state that high frequency sounds characterize the majority of industrial noise. In this industrial setting ultrasonic noise, as defined by Smagowska and Pawlaczyk-Luszczynska (2015) to be 10-40kHz, had an impact upon subjective comfort as well (Smagowska and Pawlaczyk-Luszczynska 2015). Pawlaczyk-Luszczynska, Dudarewicz, and Sliwinska-Kowalska (2007) showed with 25 operators of welding machines, which produced ultrasonic noise, the operators reported it as not only having unpleasant physical impacts (headaches - 12.1%, and fatigue 36.8%) but also that it caused discomfort with the sound being described as sharp and unpleasant by 44.4% and annoying by 36.8%.

The impact of high intensity sounds and spectral characteristics (discussed below and in Chapter 3) which are considered uncomfortable have been found to have effects upon consumer choice, satisfaction, and staff performance, with higher levels of sound intensity and vibration leading to a negative impact (Melert, Baumann, Freese, and Weber, 2008, Nor, Fouladi, Nahvi, and Affrim, 2008). This, therefore, should put these aspects of perception high upon the priorities of those whom research comfort, as well as those developing products with comfort in mind

2.2.1 Audio Intensity and Comfort

Audio comfort research has primarily, though not exclusively, focused on sound intensity which is measured by Decibel level (dB). This is due to the significant negative impact that high sound intensity can have on the human body and individuals' capacity to function normally. The British Health and Safety Executive (HSE), have placed regulations on exposure limits to high sound intensity due to their potential health risks. Those limits are 87 dB(A) for daily or weekly personal noise exposure, and a peak sound pressure exposure of 140 dB(C). dB(A) is defined as a weighting of dB levels based on frequencies of the audio signal. This weighting allows for a more accurate representation of what is perceived by a human, due to their lowered sensitivity to lower noise frequencies. dB(C) is similarly designed to follow human hearing but at much higher noise intensity. In addition to these upper limits the HSE has stated that at exposure levels of 80 dB(A) and in peak sound pressures of 135 dB(C) employers are required to carry out risk assessments and potentially reduce sound intensity and provide sound protection (Control of Nose at Work Regulations, 2005). Similarly, the WHO (World Health Organization, 1999 and 2011) define the adverse impacts of sound as a psychological, social or physical reduction in; functional capacity, capability to compensate for additional stress, or increased vulnerability to other environmental influences. The WHO provide statements and recommendations on the numerous health impacts of high intensity sounds in working and everyday life, these include; hearing impairment, interference with speech communication, sleep disturbances, potential cardiovascular disease, and potential anxiety and other psychiatric disorders. The WHO also states that high levels of environmental and occupational noise can adversely impact cognitive task performance in adults and children, these impairments are most strongly noted in reading, attention, problem solving, and memory (in particular incidental and material memory). Due to the reduction in cognitive task performance the WHO recommends an intensity of less than 35dB within classrooms to allow for good teaching and learning conditions.

The impact of high sound intensity on comfort has been thoroughly researched, with the general consensus being that higher intensity sound, such as those discussed by the HSE, are universally considered more uncomfortable (De Looze, Kuijt-Evers, and Van Dieen, 2003) and have an impact upon task performance. This was shown in questionnaire assessments of ability to carry out standard aircraft crew tasks by Mellert, Baumann, Freese, and Weber, (2008). In a study of soundscapes on Han Chinese Buddhist temples Zhang, Zhang, Lui, and Kang (2016) showed correlations between sound intensity and evaluation of subjective comfort, this was reported to be more noticeable when the sound intensity of a temple was 60dB(A) or over. It is, however, important consider the idea that context is important in the perception of uncomfortable noise. While Zhang et al. (2016) state that discomfort was perceived as more noticeable in a temple at 60dB(A), an article by Harrison (2013) notes that reductions to cabin noise in the making of the Nextant 400XT (a fixed-wing business jet) were set at an unprecedentedly low level of 66dB. This level reported by Harrison is significantly higher than those reported from Zhang. To frame the reductions noted by Harrison this class of aircraft tends to produce sound intensities around 83dB - 91dB. Ozcan and Nemlioglu (2006), in their study they show sound intensities in passenger aircraft during flight and landing ranging between 75dB(A)-80dB(A). Ivosevic, Mihalincic, and Bucak (2010) similarly place an acceptable idling cabin noise at 75dB. This notable difference between Zhang et al.'s temple at 60dB(A) and the varied noise intensities of aircraft, with an acceptable intensity at 75dB(A) indicates that context and therefore the participants expectations of the environment are important to understanding comfort.

Knobel and Sanchez (2006) determined standard hearing discomfort levels for pure tones presented for 2 seconds was between 86 decibel hearing level (dBHL) and 98 dBHL. Knobel and Sanchez's study used normal hearing participants and measured their discomfort using a test used when fitting hearing aids known as the Loudness Discomfort Level. This test is a stepwise threshold test which uses a self-report response method with seven items ranging from 'Very soft' at its lowest response to 'Uncomfortably loud' at its highest (Table 2). However, they also found that there were noticeable cross-participant differences in their responses and stated that the interpretation of the results should consider the patients' history. Other factors related to dB level have also been found to impact discomfort responses, Sanchez, Moraes, Casseb, Cota, Freire and Roberts (2016) carried out a study into sound level tolerance in adolescents. The results of Sanchez et al.'s study indicated that the participants who experienced tinnitus also had a lower sound level tolerance, which was reported again through Loudness Discomfort Levels.

Table 2: Loudness Discomfort Level, Loudness Chart (Cox, Alexander, Taylor, and Gray, 1997)

Loudness Categories

7. Uncomfortably loud
6. Loud, but O.K.
5. Comfortable, but slightly loud
4. Comfortable
3. Comfortable, but slightly soft
2. Soft
1. Very Soft

An interesting factor in interpreting the results from studies which use subjective comfort assessments is that from older studies, such as Hood (1968) which examined the relationship between discomfort level and auditory fatigue with sound pressure level and sensation level, to more recent studies such as Sanchez et al. (2016) it can be seen that while the results are reliable the authors often report cross-participant variations as a consistent issue within their metrics. This would imply that while there are audio stimuli which are universally uncomfortable, each participant has their own tolerance level and vary in their response. This cross-participant variability would call for a predictive model of comfort to either be re-calibrated for each individual or provide a method of examining a population as a whole while removing this variability. The issue of cross-participant variability is addressed within the proposed model of audio comfort in this thesis (see Chapter 5).

2.2.2 Audio Frequency and Comfort

Intensity of noise, however, is not the only factor when considering discomfort responses in participants. Spectral differences also can lead to variations in preference and ultimately comfort levels. The WHO, in their documentation, state that there is a greater need to reduce the intensity of an audio signal if it contains low-frequency components due to the potential of both annoyance and health risk. Leventhall (2004) states that low-frequency noise (from around 10Hz to 200Hz) can cause discomfort to those sensitive to it. This sensitivity, Leventhall states, tends to develop in middle age. These low-frequency sounds when constant (referred to as the Hum) are classified as a background stressor (Benton, Waye, Rylander, and Leventhall, 1994, Benton, 1997). These low-frequency sounds, as previously discussed, cause annoyance. The effects can be exacerbated to severe levels with long-term exposure and for those with exaggerated susceptibility (Benton, Waye, Rylander, and Leventhall, 1994, Benton, 1997).

Further evidence for the impact of low-frequency noise can also be seen in Ziaran (2014). When assessing the impact of low-frequency noise produced by cars on participants Ziaran found that in addition to lower attention and a negative impact on driving safety, annoyance and discomfort were also increased by low-frequency noise. To determine the nature of the low-frequency noise Ziaran (2014) used a sound quantification measure from Zwicker and Fastl (2007) (Fig 10). Dempsey, and Leatherwood (1976) also show the importance of considering spectral differences with the assessment of comfort levels. Dempsey et al. (1976 and 1979) (Fig 11) showed that in an aircraft environment, spectral differences impact upon comfort levels significantly more than the intensity of the sound would predict. Dempsey et al. (1979) display in Figure 11 that while the dB(A) level of the audio signal remains at the determined levels discomfort ratings increased significantly, this was therefore attributed to the octave band the audio signals were produced in.

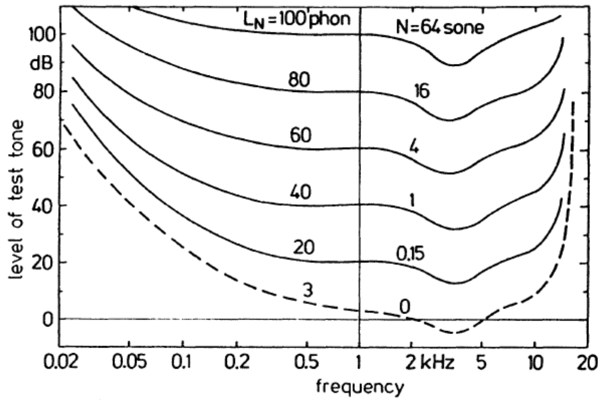


Figure 10: Graph of equal loudness contours (Zwicker and Fastl, 2007). Each line shows a specific combination of tone level and frequencies that are perceived as equally loud

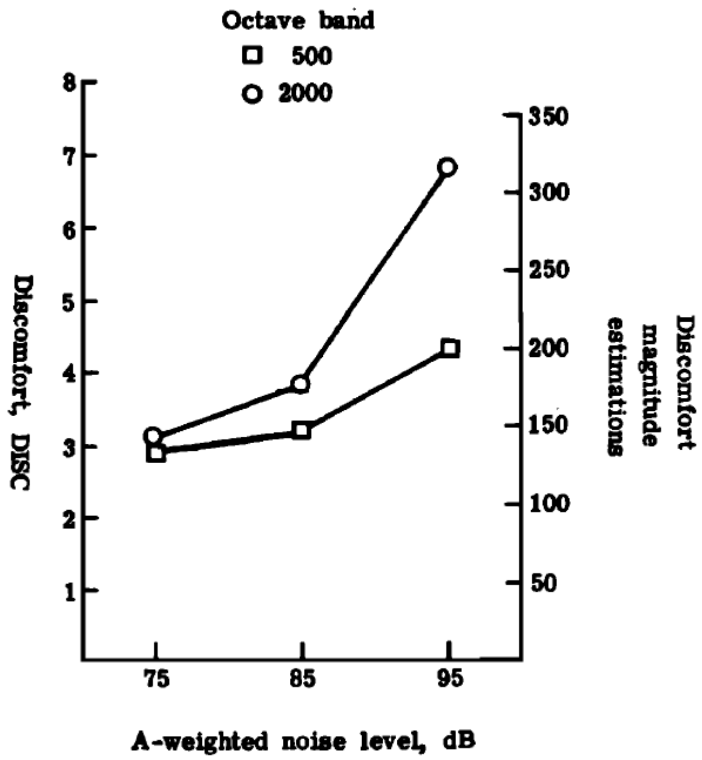


Figure 11: Discomfort results at 500 and 2000Hz bands across dB levels from Dempsey et al.(1979)

Work from McCurdy (1988) explored the impact on annoyance caused by turboprop aircraft flyover noise. McCurdy used recordings of flyovers and a scoring metric to assess the annoyance increase caused by different turboprop aircraft. In this study McCurdy found that interactions between frequency and tone-to-broadband noise ratio were driving factors in the impact on annoyance. McCurdy showed that annoyance could be predicted using duration-corrected, A-weighted sound pressure levels with a modified tone correction. When comparing aircraft types McCurdy found inconsistent results either indicating no differences or small differences in preference of lower frequency advanced turboprops (67.5 and 125Hz).

The NLR have developed a high-fidelity method of simulating flyover noise of aircraft; the Virtual Community Noise Simulator (VCNS). The VCNS can be used to experience aircraft noise in a virtual representation of standard flight paths from the ground with 360-degree video and 3D audio presentation. The VCNS is a useful tool in simulating the potential impacts of aircraft noise. The development of the VCNS, as with most aircraft noise assessment tools, is focused on assessing the effect of external aircraft noise on participants with a view to future noise mitigation activities. The VCNS has also been used in other similar projects, with the Ministry of Infrastructure and Environment commissioning the development of the VCNS to assess visual and auditory impact of wind turbines (http://annualreport.nlr.nl/2011/Projects/Spin-offs/Wind%20turbine%20simulation/) in 2011.

Zhang, Zhang, Lui, and Kang (2016) also show a significant correlation between the psychoacoustic parameter Sharpness (a method of reporting high frequencies in audio signals) along with sound intensity and sound preference evaluations. In studies where it was tested for, not only was discomfort noted in these situations, but also a reduction in cognitive ability and task performance (Dempsey et al., 1976, Ziaran, 2014). The results showing impaired task performance and cognitive function would fit in line with part of the WHO definition of audio discomfort in which they name a reduction of functional capacity observed in participants.

2.3 Measuring Comfort

Comfort research was initially dominated by the use of subjective questionnaire studies, which examined participants' personal conscious experience of stimuli. For example, McNulty and McNulty (2009) indicated that sound based discomfort could indeed be examined through subjective Likert scale questionnaires focused on comfort ratings. McNulty et al.'s (2009) Likert scale questionnaire consisted of two questions one focusing on perceived noise intensity ranging from "Quiet" to "Very Loud" and a second question focused on the acceptability of noise based on a 1-5 rating scale where 1 was "Completely Acceptable" and 5 was "Completely Unacceptable". Other examples of the bias towards primarily questionnaire-based research in early comfort studies include Dempsey and Leatherwood (1975) who examined the impact of vibration on passenger comfort. Their assessment used estimations of vibration discomfort in comparison to a baseline. While other research such as Richards and Jacobson (1974) used two longer questionnaires to assess passenger comfort on aircraft taking into account multiple factors such as pressure changes, noise, and temperature.

Measures of discomfort were initially focused around these subjective questionnaires which examine the aspects of participants' experience and their perception of such (Oborne, 1975, 1977, 1978), for example, those seen previously in McNulty et al. (2009) and Cox et al. (1997). The use of questionnaires in comfort research has been noted to have one primary issue which is cross-participant variation (Sanchez et al. 2016). Both Hood (1968) and Sanchez et al. (2016) have both noted that cross-participant variations are apparent in subjective comfort assessments such as the Loudness Discomfort Level chart Cox et al. (1997) (Table 2). While this can lead to complications with the interpretation the data this is only the case when not examined with the cross-participant variability in mind. The cross-participant variability noted in previous studies does have a useful implication. This implication is that it is possible that these variations in ratings on a comfort scale indicate that there are variances in participant's comfort tolerance. Essentially, a stimulus that causes discomfort will likely have the same directional impact on a participant (i.e. causing a participant to rate the stimulus as uncomfortable). This impact however, due to cross-participant variability, will not necessarily be of the same magnitude or have the same starting point on a rating scale. This is due to higher or lower participant tolerances to discomfort. In addition there is no common externally defined reference point for discomfort across participants (Meyer, White, Cant, Pinto, Milella, and Cooper 2018). A difference in comfort tolerance would explain why comfort ratings gathered by Hood (1968) and Sanchez et al. (2016) are consistent within participants and follow their predicted trends while showing cross-participant variability. More recent research has employed other measures, which provide a more objective view into participant's experience and comfort levels. These studies tend to use measures involving cognitive performance through task performance metrics (such as ease of reading comprehension) (Iachini et al., 2012) and physiological responses (such as heart rate, or neurological changes) (Mellert et al., 2008) in tandem with subjective questionnaires. This combination of objective and subjective measures has provided further insight into the effects of discomfort and the identification of the stimuli which cause discomfort. The use of multiple measures also provides a potential standard to define the impact of these effects in different contexts from personal experience with questionnaires to changes in capability such as with task performance; this is explored in this thesis. This methodology is incorporated into the building and testing of a new model of audio comfort, presented in Chapter 5.

It was argued by Straker et al. (1997) that reactions to stimuli which would elicit discomfort would include impairment of cognitive performance. Numerous studies have shown that as discomfort is experienced cognitive performance declines making it a useful method of objectively measuring discomfort (Clark and Sörqvist, 2012, Khan et al., 2004, Iachini et al., 2012, and Mellert et al., 2008).

2.4 Fidelity and Comfort

Audio comfort is not only impacted by the physical properties of the stimuli but also the context and way it is presented. Methods of varying participants' experience of the context in which stimuli is presented vary, however one of the most practical methods of varying context is the use of simulators with varying fidelity levels. These variations can take multiple forms including; vestibular cues (Zeyada and Hess, 2000), Field of View (FoV) (Duh, Parker, and Abi-Rached, 2002), and sound and visual cues (Perfect, White, Padfield, Gubbels, 2013, Larsson, Vastfjall, and Mendel, 2008). These multiple variations that impact fidelity not only show that simulator fidelity is a complex issue, but also led to usage of two different experimental manipulations using fidelity in the research in Chapter 4.

High fidelity simulations are used in a variety of industries as training methods ranging from healthcare workers (Lewis, Strachan, and Smith, 2012), to Heavy Goods Vehicle drivers (Allen and Tarr, 2005).

Examples in comfort research such as Iachini et al., (2012) used simulated subway environments to assess multi-sensory comfort through presentation and removal of visual cues. These presentations or removals were employed to vary the fidelity levels experienced by the participants. Iachini et al. determined that there was a negative impact on comfort reported by participants in the case of a lower fidelity environment.

2.4.1 Fidelity in Simulations

Fidelity, in the context of simulation, refers to the perceived realism of an environment produced within the simulator. This can vary in level through changes to the simulation such as audio and visual presentation (Perfect, White, Padfield, Gubbels, 2013). Simulators are used in a variety of industries and development. Applications can be seen in examples such as training for healthcare workers (Lewis, Strachan, and Smith, 2012). In a study by Sinha, Johnson, Hunt, Woolnough, Vidal, John, Villard, Holbray, Bellow, and Gould (2009) it was found that 83% of participants who used a training simulation for a percutaneous nephrostomy procedure considered it to be a useful model for training. Subsequent studies on other simulated medical procedures have found similar effectiveness in training. With Luboz, Zhang, Johnson, Song, Kilkenny, Hunt, Woolnough, Guediri, Zhai, Odetoyinbo, Littler, Fisher, Hughes, Chalmers, KLessel, Clough, Ward, Phillips, How, Bulpitt, John, Bello, and Gould (2013) showing that participants found that the simulation was effective for both learning basic skills (78% of participants) and learning equipment use (86% of participants). The use of high-fidelity simulators can be seen used in other contexts such as training and assessment of Heavy Goods Vehicle drivers (Allen and Tarr, 2005). The variations in fidelity are linked to the successfulness of training and fitness for purpose, low fidelity being linked to poorer quality training and learned behaviours.

The creation of a realistic simulation consists of multi-sensory application of stimuli to increase immersion. These stimuli include: visual representation of the simulation such as FoV (Duh, Parker, and Abi-Rached, 2002 vestibular cues (Hess and Zeyada, 2000), and audio cues (Larsson, Vastfjall, and Mendel, 2002). In Larsson et al.'s (2002) study they showed that there was a significant increase in presence not only when there was consistency between sound information and spatial cues but also other aspects of audio cues including room acoustic cues and source content. As it is possible to determine the level of fidelity with the use of these cues within a simulated environment, it therefore follows that the level of fidelity and the feelings of presence (i.e. the feeling of being present within the virtual world) caused by the simulation and experienced by the participants can be increased or decreased. This capability to vary the fidelity and the feeling of presence allows the perceived fidelity of a simulation to be measured through presence assessments.

2.4.2 Effect of Presence on Participants' Experience

Presence, in the case of simulations, is the feeling of being engaged and immersed within a simulated environment. This differentiates presence from fidelity by determining that an increase in fidelity can cause a sense of presence. These two can both be differentiated from something being perceived as realistic which itself is a result of both a feeling of presence and high fidelity causing a person to perceive the simulation as realistic. Increased simulation fidelity influences upon participants' perception and reaction to the simulated experience; Place Illusion (PI) which is defined by Slater (2009) as "The strong illusion of being in a place in spite of the sure knowledge that you are not there" and Plausibility Illusion (Psi) which is defined by Slater (2009) as "The illusion that what is apparently happening is really happening (even though you know for sure that it is not)". These refer to cues in the environment (including task response) which lead to the feeling of PI and that virtual reality (VR) events are occurring which causes Psi. While Slater (2009) states that these lead to a deeper feeling of presence, Slater also states that in both PI and Psi participants are aware of the VR and that these events are not occurring in reality. Instead, the feeling of presence is based on the reactions of the simulation (Psi), and the sensation of presence within a simulated environment (PI). It is further postulated by Slater that while PI and Psi are in effect the participant's reactions to the virtual reality will be realistic. Essentially, that with an increased level of simulation fidelity participants are more likely to treat the simulation as real and to react accordingly. This postulation is supported by research from Meyer, Shao, White, Hopkins, and Robotham (2013), in a study in which they manipulated visual motion signals in both a VR and non-VR settings. This manipulation was carried out to determine if Visually Evoked Postural Responses (VEPR) would occur in a VR environment as they do in non-VR environments with visual motion signals. Meyer et al. (2013) found that in high fidelity VR environments these VEPRs were present similarly to those in non-VR environments where one would expect to see VEPRs. This provides an effective method of objectively testing presence and fidelity in VR environments. The study also displays evidence to the previous statement that with an increased level of fidelity participants are more prone to treating the simulation as real and reacting accordingly.

2.4.3 Audio Beyond its Physical Properties

While spectral characteristics and intensity of sounds have an effect upon comfort, these are not the sole contributors to perceived comfort. Studies have indicated that height of the source of sound in physical space can lead to different interpretations of pitch of the sound (Rusconi, Kwan, Giordano, Umilta, Butterworth, 2005), such as finding higher pitches being placed higher in physical space than lower pitched noises. Data and research like this indicate that perception of audio stimuli transcends its simple physical qualities such as intensity. This therefore indicates that if research is to assess the impact of noise upon comfort levels the context in which the noise is presented cannot be ignored. This would include, but is not limited to, such contextual aspects as environment and audio-lag.

Simulator fidelity is noted to influence task performance, in which a higher level of fidelity is often postulated to have a positive effect upon performance on a given task. Meyer, Wong, Timson, Perfect, White (2012) showed that when provided with training a higher level of fidelity led to a significant improvement in task performance. The importance of fidelity in a simulated environment in relation to comfort ratings however is less fully investigated aspect. Iachini et al. (2012), through manipulation of the visual virtual environment from low fidelity to high fidelity (i.e. visuals on or off), assessed the impact of general comfort. The manipulations in the visual virtual environment led to a change in comfort ratings; in which comfort was positively correlated with fidelity level. Research such as this provides a strong indication that with sound and visual fidelity changes comfort ratings will also follow suit, with fidelity increases being mirrored by an increase in comfort tolerance to audio signals. The addition of audio cues and the segmentation of the definition of fidelity is a novel point explored in this research.

In this thesis, it is argued that Slater's PI and Psi provides an explanation for the expected increased tolerance of participants to audio disturbances in high fidelity environments. This tolerance would primarily stem from increased fidelity which in turn enacts stronger PI and Psi. This, therefore, would lead participants to be more tolerant of audio signals heard in high fidelity environments than those heard within a lower fidelity simulation.

2.4.4 Theoretical Impact of Simulator Fidelity on Participants

The increased simulation fidelity method of impact has been theorised on by Slater (2009). Slater posits that a participant's perception and reaction to simulated environments and events are represented by the phenomena of PI and Psi.

PI is enacted in the case of high-fidelity cues in the environment including responses to tasks, while Psi is caused through feeling that high-fidelity VR events are occurring. Though Slater states that these illusions both lead to a greater feeling of presence, he also confirms that the participants are aware of the VR and that these events and cues are simulated rather than occurring in reality. Instead, the feeling of presence in the simulation is based on the reactions of the simulation (causing Psi) and the cues within the simulation (PI). Slater states that while the illusions are in effect, the participant's reactions to VR will be realistic. Participants will react and interact with the simulation as they would outside of VR and act accordingly. Slater goes on to say that this phenomenon becomes stronger as the level of fidelity increases (i.e. the higher the level of fidelity) in the simulation the stronger PI and Psi are and so the more participants will interact with the simulation as they would a real-world scenario.

2.4.5 The Link from PI and Psi to Comfort

Within this thesis, PI and Psi are posited to increase the participant's tolerance to audio disturbances. Using this as a basis for the explanation as to fidelities impact upon comfort (see Chapter 4), the explanation puts forward that the decrease in comfort in lower fidelity environments, when presented with the same audio signal, is due to a decreased tolerance to the signal as PI and Psi are reduced. This thesis puts forward that this tolerance is theorised to stem from the previously mentioned effect of PI and Psi leading to participants treating simulations as reality. This change in the perspective of the participant not only changes how participants interact with the simulated environment but also their expectations of the cues received. In essence, if the participant treats the simulation as reality and the simulated environment these audio cues are expected, then the audio cues are easier to legitimise and so the participants experience a higher tolerance to these cues. This increase in tolerance would lead to higher fidelity environments being considered more comfortable while lower fidelity environments would be experienced as less comfortable due to the weaker PI and Psi experienced.

2.5 Theories of Comfort

Theories of comfort were originally produced primarily in one area of research, nursing. These theories focus on patient health and wellbeing, facilitating rehabilitation for the patient. Kolcaba (2001) discusses nursing theories of comfort in a review in which she notes the three bases that theories of nursing focus on; human needs, adaptation, and health/illness. With the basis of human needs Kolcaba states that this is viewed as what the patient requires to grow and be sustained. This has parallels in the basis of comfort as described by the theory defined by this thesis, in which comfort is seen as a state where there are no needs that are in need of being resolved. While of course there is a notable difference between comfort in comparison to patient health and wellbeing and comfort in relation to uncomfortable stimuli parallels can be drawn between the two. Nursing theory should be acknowledged for the practical beginnings of comfort theory.

Nursing has provided conceptual considerations for psychophysical comfort models and theories. These included; comfort needs being driven by expectations, which is mirrored in the removal or addition of positive or negative reinforcers in this comfort model, changing the environment from the participants' expectations, and the concept that participants have implicit and explicit requirements for comfort which when fulfilled lead to better motivation and performance in therapy and rehabilitation which is similar to the use of task performance as a metric of comfort assessment.

Psychophysical theories of comfort hold a similar basis to nursing theories of comfort, but instead generally are used to explain comfort and discomfort through experimental means. Metzger's (1994) theory of comfort states that comfort is primarily associated with: ease, relaxation, convenience, and well-being. Furthermore, Metzger noted that a distinction should be made between luxury factors and comfort. With these concepts, Metzger put forward four tenets of comfort:

- freedom from physical complaints – physical stimuli causing no distress,

- ease – an ease of carrying out tasks or activities,
- efficiency – objective performance not impacted upon,
- individuality – providing participant opportunity to express self through freedom of choice and personalisation.

Metzger's theory of comfort provides an interesting perspective on comfort. This follows on from previous nursing accounts of comfort. These include both the physical complaint that stems from discomfort and additionally an implication of discomfort impacting on an individual's capability to carry out a task. This not only gives insight into the causes of discomfort but also provides a point from which to make assertions on the methods of measuring comfort. Those assertions can be applied to perceived difficulty of tasks (ease), objective performance in a task (efficiency), and perceived discomfort (freedom from physical complaints). What limits Metzger's theory of comfort is that while it provides an insight into the make-up of discomfort, it does not provide a framework to measure or assess discomfort. The theory also crucially does not provide an explanation for how these tenets interact with one another or if they are indeed wholly separate entities acting on an individual.

De Looze, Kuijt-Evers, and Van Deen (2003) provide a model of seating comfort (Fig 12), which puts forward a two-sided, four stage model. The two sides refer to the physical aspect of the experience and the mental aspect of the model. The first stage of the model is the context that the stimuli is presented which De Looze et al. (2003) states includes psychosocial factors such as job satisfaction and social support on the mental side, and the environment the seat is presented on the physical side. Both sides of this model in the context stage also include a task aspect. This task aspect, De Looze et al. state, can expose a participant to movements or positions which can change the level of comfort experienced. The second stage refers to the seat itself, this category discusses both the aesthetics of the seat and the physical features of the seat; these fall into the mental and physical categories respectively. The third stage which De Looze et al. claim determines seating comfort is the human aspect. This human aspect is once again split the mental aspect refers to the expectations and emotions experienced by the participant, while the physical aspect looks upon the impact of the seat upon the participant’s physiology such as propriocepsis (muscle spindle, tendon, and joint sensation), pain, interocepsis (sensation from internal organs), and exterocepsis (sensation on the skin). De Looze et al. claim that in the final stage of the model discomfort is informed from the physical side of the model. This, along with the mental side, subsequently feed into the final comfort perception.

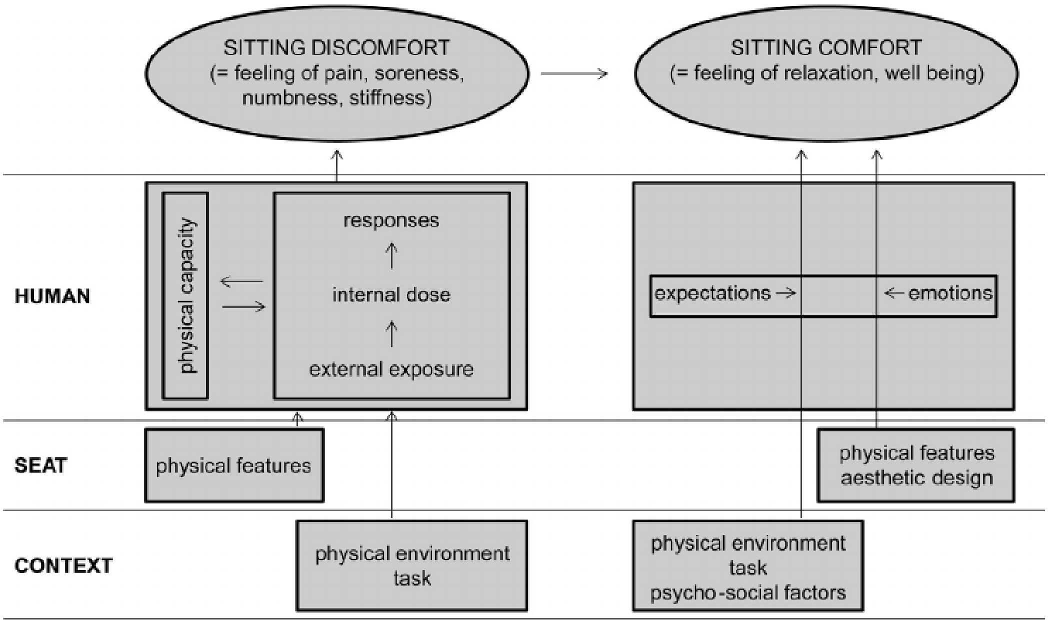


Figure 12: De Looze et al. (2003) model of seating comfort

The most salient points from De Looze et al. (2003) in the context of this thesis are; firstly, that at every point throughout the process of determining level of comfort or discomfort both the physical and psycho-social aspects of seating comfort inform the model. The second is that De Looze et al. predict objective measures will have a stronger relationship with discomfort than comfort. This being due to both objective measures and discomfort being primarily part of the left pathway of the model.

A more recent, but similar, model of seating comfort is presented by Naddeo Cappetti, Vallone, and Califano (2015) (Figure 13). This model builds on the previous work by Vink and Hellbeck (2012). This model adds explicit instrument-based evaluation of components for interaction as well as the role of expectations. Naddeo et al. (2015) state that with a higher cost of a product comes a greater expectation of comfort, they claim that this changes the perceived comfort experienced by individuals. While this model does provide the comfort model literature with some interesting additions it (much like its predecessors) does not provide specifications of how to quantify or predict perceived comfort.

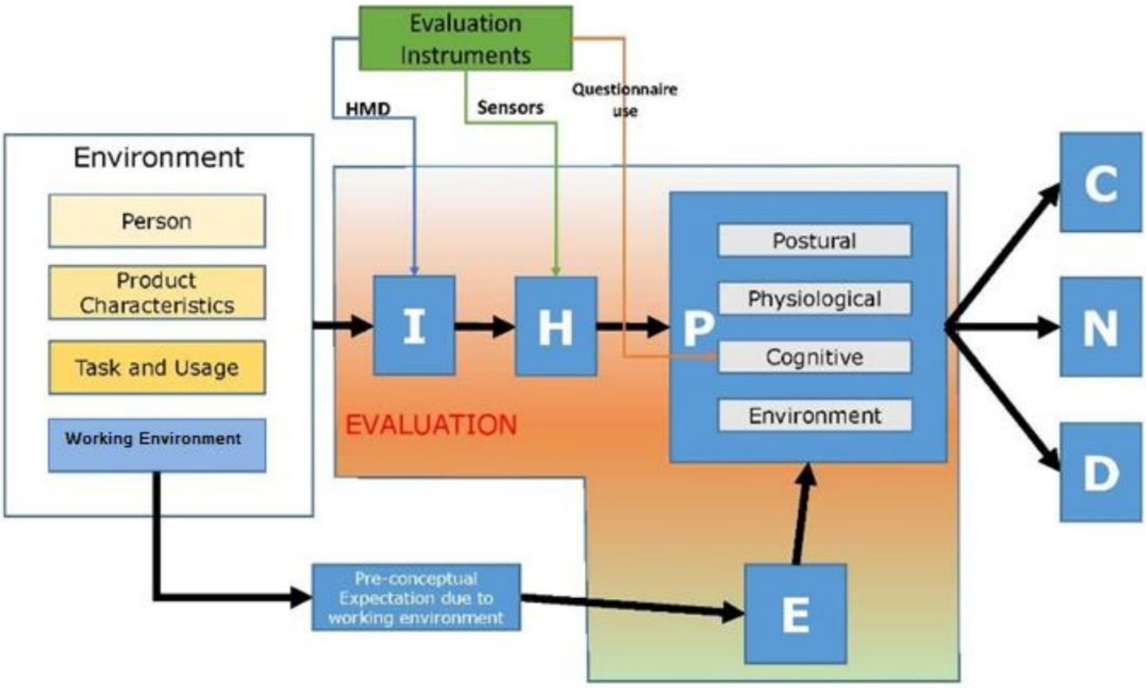


Figure 13: Naddeo et al. (2015) model of seating comfort (I- Interaction, H-Internal human body effects, P-Perceived effects, C-Comfortable, N- Nothing, D-Discomfort, E-Expectations)

With these theories of comfort there are set assumptions which describe the mechanism of the perception of discomfort, such as the 'context', 'seat', 'human' progression in De Looze et al. (2003). These assumptions cover a theoretical understanding of if a stimulus could be considered uncomfortable or not. The issue is that while these theories and models of comfort provide a descriptive explanation of comfort, they do not enable testable predictions. The model proposed in this thesis makes these testable predictions. It makes these predictions by taking considering the complex nature of comfort that previous models and theories have described and applying appropriate measures to assess this complex nature. The model proposed is based on Rolls' (1990) model of emotion. Rolls' model and the significant modifications made to adapt this model to comfort research are detailed in Chapter 5. The model uses aspects of Rolls' model which are frequent themes in comfort models and theories including the expectations of a person. The proposed model then applies them to comfort research, with the addition of assessment metrics to make testable predictions.

Chapter 3 – Experimental Assessment of Task Performance and Subjective Metrics

3.1 Introduction

This first experimental chapter revisits the literature on the sound intensity and spectral differences, impact on human comfort, and methods of measuring comfort. The chapter presents a range of task performance metrics and a new comfort questionnaire to assess their efficacy in determining the impact of sound intensity and spectral differences on comfort and task performance.

3.2 Experiments to Assess the Feasibility of Metrics in the Measurement of Comfort

This section details several experiments that were conducted in this thesis to determine the feasibility of objective metrics in the assessment of comfort. The experiments began with a pilot study to test a large battery of objective metrics. Following the pilot study, an assessment of the remaining metrics was carried out with the use of spectral differences and noise intensity.

3.2.1 Pilot Study

A pilot study was conducted to determine the effectiveness of a set of task performance metrics in their ability to measure the effect of uncomfortable stimuli. This pilot study took a battery of task performance metrics to determine which the most effective measure in this capacity was. The tests (detailed in section 3.2.1.1) included the Brief Test of Attention (BTA) (Schretlen, Bobholz, and Brant, 1994) (average completion time 4.2 minutes), the Tower of London (ToL) task (Ward and Allport, 1997) (average completion time 7.6 minutes), the Phillips Effect Paradigm (PEP) (average completion time 3.1 minutes), and the Speech Identification Task (SIT) (average completion time 4.5 minutes). (See section 3.2.1.1). The study consisted of 21 participants with an average age of 21 (female=17, male=4). These participants self-reported normal hearing.

The pilot study followed a repeated measures design, in which participants were exposed to the sound file of pulsed white noise at either 70dB or 76dB. These decibel levels were chosen as they represented both a range of sounds noted by Ivoševic, Miljković, and Krajček (2012) as being part of turbo-propeller aircraft cruise noise intensities. No 0dB condition was included as speech intelligibility does not change for low noise levels (see section 3.2.1.1.4). While the participants were exposed to these sound files they carried out the cognitive tasks mentioned above. Once the participants had completed the tasks they were given a short 5-minute break and then were asked to repeat the tests with the next sound intensity (either 70dB or 76dB). The conditions and task order were counterbalanced to avoid order and learning effects. The experiment was carried out in a soundproof booth with the sound files for the tests and the white noise played through headphones. These sound files were played through the RPvdsEx programme, which allowed for both white noise and stimuli such as the BTA or SIT to be played through the headphones at the same time (Figure 14).

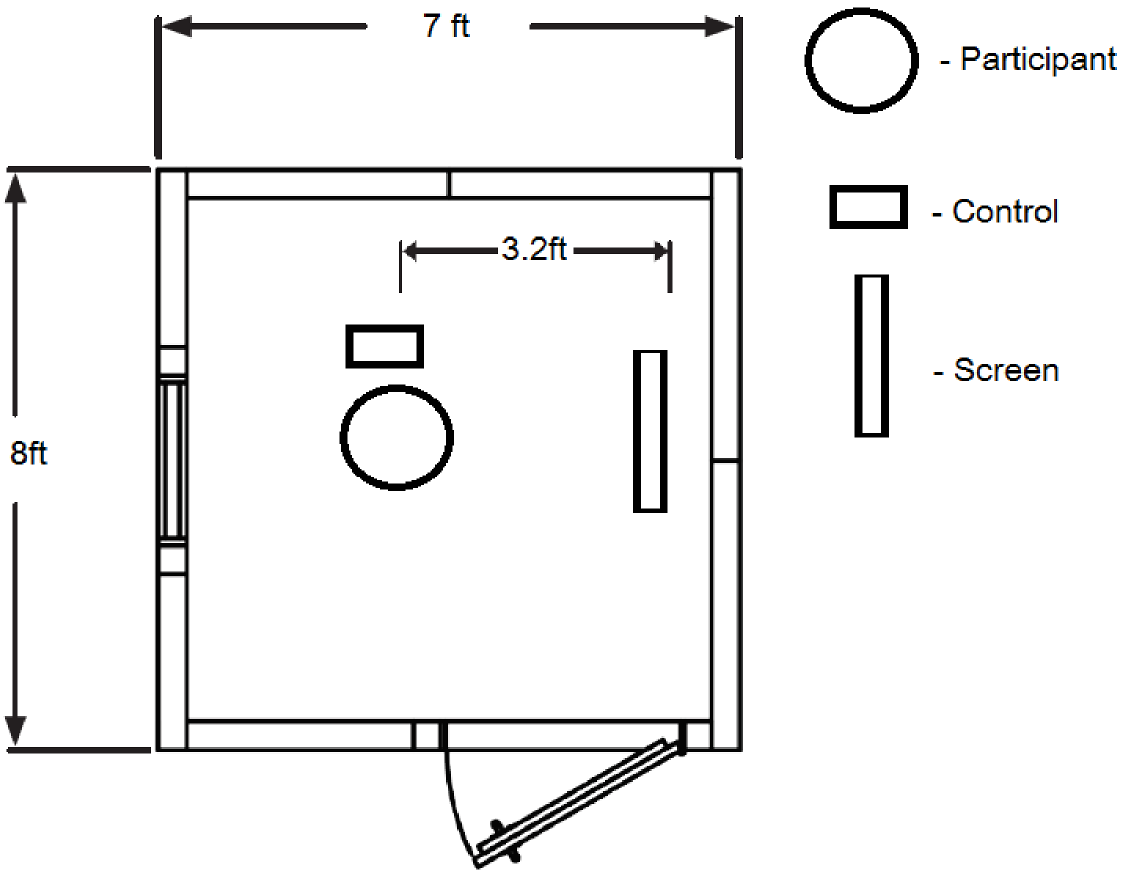


Figure 14: Experimental set-up for pilot study and Experiments 1 and 2

3.2.1.1 – Behavioural Metrics

The metrics detailed in this section were chosen due to their use of task performance as a measure of experimental manipulation. The metrics were tested on 21 participants with an average age of 21. The study was carried out in a soundproof booth to reduce any extraneous variables such as outside noise. The study used interference stimuli of pulsed white noise at 70 dB(A) and 76 dB(A), the order of which was counterbalanced to avoid order or learning effects. These decibel levels were chosen as they represented both a range of sounds noted by Ivoševic, Miljković, and Krajček (2012) as being part of turbo-propeller aircraft cruise noise intensities. These sound intensities were calibrated using a sound level meter and a synthetic ear (Fig 15) to simulate and calibrate the sound intensity the participant would be exposed to.

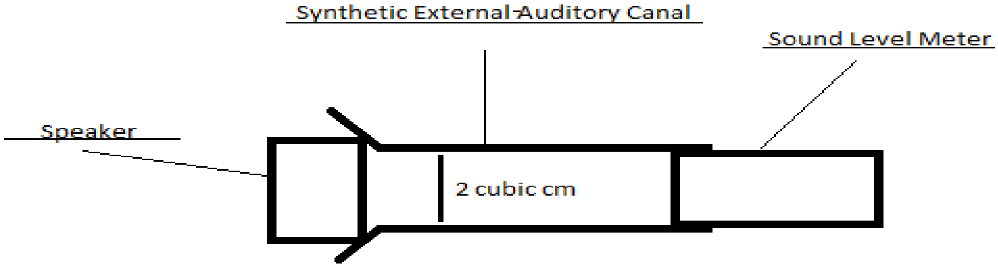


Figure 15: Diagram of synthetic ear

3.2.1.1.1 – Brief Test of Attention (BTA)

The BTA, developed by Schretlen, Bobholz, and Brant (1994) was designed to show levels of attention and the potential for maintenance of attention. The BTA as a cognition base task was identified for examination with mind towards techniques used in Iachini et al. (2012) and the ease and efficiency aspects of comfort theory put forward by Metzger (1994).

The participants were played a sound file consisting of strings of numbers and letters. These strings start at a size of 4 characters and increased to a string size of 18 characters with a total of 14 strings. The characters are read out at 1 character per second with a 5 second gap between each string. The participants are asked to count the amount of numbers in each string and note the total amount of numbers for each string down in the corresponding box.

3.2.1.1.2 – Tower of London Task (ToL)

The ToL task was developed by Ward and Allport (1997). The programme used in this study was from the Cognitive Psychology Experimental Package (CPEP: Wills, 2003). The ToL task is an assessment of complex problem-solving ability, focusing on planning abstract processes. This problem-solving and planning aspect of the ToL task was the reason for its inclusion in the pilot study. As with the BTA the cognitive performance aspects of the ToL task fit well with Metzger (1994) and their tenets of comfort, in particular ease and efficiency.

The participants were asked to move coloured shapes (Fig 16) into a predetermined pattern within a set time and movement limit. The move limits were 3, 5, 7, 9, or 13 moves. The 5 move cases were split into two categories those with sub-goals and those without sub-goals. These sub-goals were extra stages that the participant must complete for the goal pattern to be made.

Participants completion or non-completion of these tasks are recorded along with completion times, and planning times.

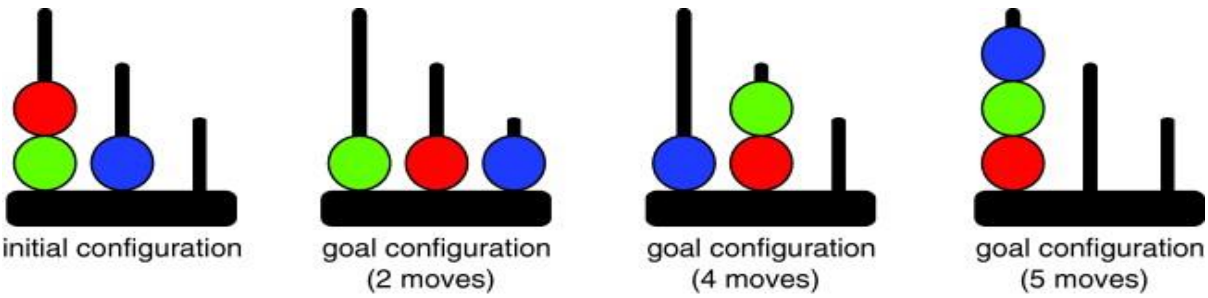


Figure 16: Example configurations of ToL task

3.2.1.1.3 Phillips Effect Paradigm

This test was developed by Phillips (1974) and is an examination of retention of information within working memory, based in the theory of the multi-store memory model (Atkinson and Shiffrin, 1968). This was chosen with reference to the ease and efficiency aspects of Metzger's (1994) theory of comfort, in which Metzger states that for comfort to be present, tasks should not be interfered with and should be considered easy. This test was taken from the Cognitive Psychology Experimental Package (CPEP: Wills, 2003).

The participant was shown a pattern of black and white squares which varied in size between 8x8 and 4x4 (Fig 17). After either a 20ms or 1000ms delay a pattern is again shown. The participant was requested to indicate through key presses on a keyboard whether the second pattern is the same or different from the previous pattern.

The number of participant's correct and incorrect answers was recorded. The number of trials used in the pilot study was 48 which was increased to 124 in experiment 1 to increase sensitivity of the test.

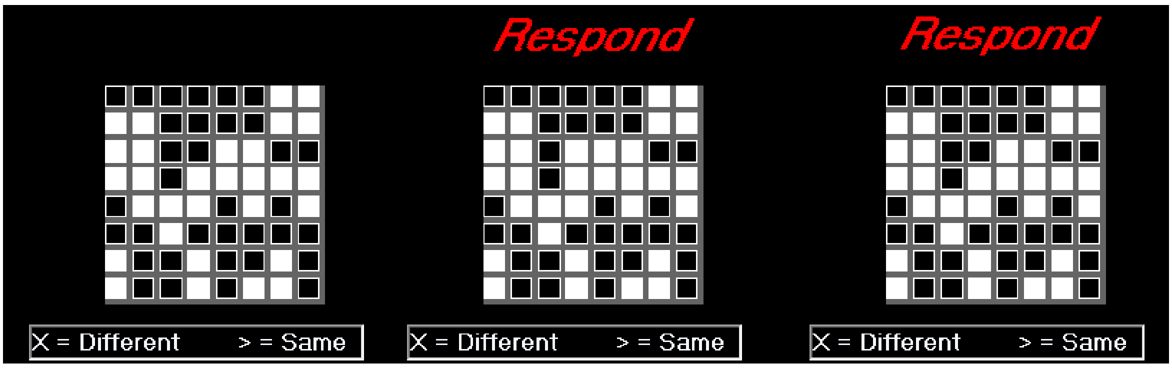


Figure 17: Phillips Effect Paradigm example, (left to right) – display, incorrect response, correct response

3.2.1.1.4 Speech Identification Task

The SIT used in this study was developed by Meyer and Morse (2003). In this research, it was used as an indicator of the difficulty in interpreting audio input. There were two main reasons for the identification of this task as a measure; speech intelligibility and comfort are both related to ambient noise level. It is argued that task difficulty is a good predictor of comfort (Iachini et al., 2012), and it is a well-established test that has previously been shown to be sensitive across a wide range of Signal to Noise Ratios (SNR) due to the masking effect shown at different SNRs (Fig 18). In Figure 18 it can be seen that as the SNR increases the consonants correct increases, meaning that as the SIT audio becomes closer in intensity to the outside interference the performance of participants increases and vice versa. It was noted in Meyer and Morse (2003) that performance on the SIT is reduced with an increase in the Signal to Noise Ratio (SNR), meaning that as the difference in dB level between the signal being heard and the other present noise (i.e. increasing the ratio) the performance on the SIT reduces due to the increased interference.

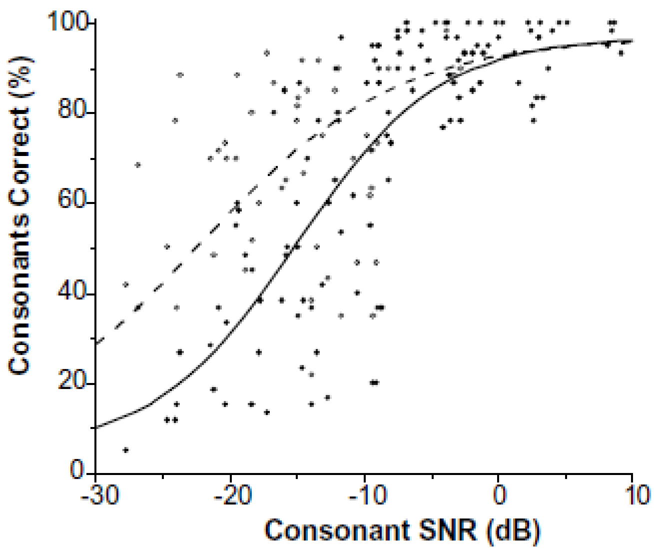


Figure 18: Depiction of consonants correct against consonant SNR (Meyer and Morse, 2002).

Figure 18, along with Meyer and Morse (2002), shows is that background noise reduces the intelligibility of speech. This can objectively be measured by scoring performance on a Speech Identification Task. Performance can be measured by the number of correct responses a participant gives. A greater SNR between the consonant sounds (such as “acha”) and the interference sound (such as pulsed white noise) causes a decreased performance. This is due to the interference with the participant’s capability to carry out the task. Therefore, this test can be used to assess the impact of a stimuli on task performance through the number of correct responses a participant provides.

The participants were played a consonant speech sound between two “a” sounds, for example “acha” where the “ch” represents the consonant speech sound. The participants were then asked to identify which item they had heard from a set of 20 similar sounds. The participants carried this out by pressing the button labelled with their chosen sound on a Graphical User Interface (GUI) (Fig 19) (Meyer and Morse, 2003). The test used 60 trials per assessment condition in the pilot study and 200 trials in Experiments 1 and 2.

The numbers of correct responses were recorded for each condition. In the current study, it was used as an indicator of a difficulty in interpreting noise input and therefore of task difficulty.

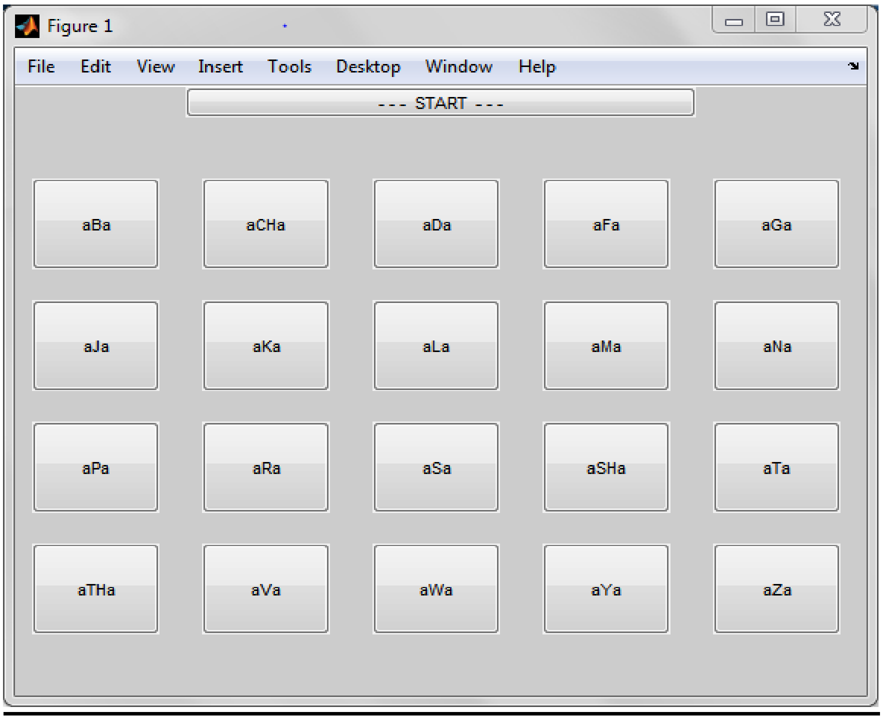


Figure 19: Speech Identification Task GUI (Meyer and Morse, 2003)

3.2.1.2 – Pilot Study Results.

During the pilot study the BTA was assessed and found not to be impacted by the change in sound intensity (t (20)=0.629, p=0.2685) (see Appendix A for more information). The reason for this non-significant result is argued to be that while there is an aspect of audio perception in the BTA, its primary focus is on short term memory and attention.

The ToL test did not show an effect of intensity in the pilot study (χ^2 (1, n=21) = 1.546, p = 0.107) (see Appendix A for more information). The ToL's lack of sensitivity to the change in dB as manipulated by the pilot study was credited to the lack of an audio aspect to the task. While this metric does provide a problem solving and planning task, the audio stimuli was unable to impact the performance of the task significantly.

The PEP showed a non-significant main effect for noise intensity ($F(1,20)=0.847$, $p=0.368$). However there was a significant main effect for pattern complexity ($F(1,20)=46.183$, $p<0.001$).

The SIT also showed a non-significant result when tested with a student paired t-test of ($t(20)=0.864$, $p=0.199$) in the case of this test it was argued that an increase in the number of trials in the SIT would significantly improve the tests effectiveness.

3.2.1.3 – Brief Discussion of Pilot Study Results

Both the ToL and the BTA were removed due to a lack of sensitivity to the change in audio stimulus. This was concluded to be due to a lack of stimuli specificity, with both the BTA and ToL providing cognitive tasks which used processes which were not impacted by noise changes. In the case of ToL this was planning and problem solving and in the BTA this was short term memory.

The pilot study also led to the number of trials in both the PEP and SIT being increased in Experiment 1 and 2. The PEP trial number was increased from 48 to 124 trials and the SIT was increased from 60 to 200 trials. This was done to improve statistical power during the analysis of these tasks as lower trial numbers had led to a lower level of reliability in the participants' responses. This approach was not taken with the BTA or the ToL as an increase in trial number was less practical due to the nature of these tasks.

3.3 Experiment 1 – Assessment for noise intensity

The aim of Experiment 1 was to assess the applicability of the PEP and SIT task performance metrics to the assessment of comfort in relation to noise intensity. This experiment also provided a test for the functionality of a comfort questionnaire. With the data from this experiment these metrics could be tested in addition to assessing the impact of noise intensity on participant comfort.

3.3.1 Subjective metric

3.3.1.1 Comfort Questionnaire

A comfort questionnaire was developed to examine a variety of aspects of comfort and experiences related to assessing the participant's perceptions of the sound in reference to comfort and audio stimuli. The comfort questionnaire is comprised of questions that were based on previously tested questionnaires, including McNulty et al. (2009), and Oborne (1975), Quehl (2001). A set of questions were chosen and put together to form this questionnaire.

The comfort questionnaire had 6 questions and followed a standard 10 point Likert scale. (Fig 20). This approach to questionnaire building has been used in the literature (Cooper, Milella, Pinto, Cant, White, Meyer 2018.). Each question was chosen to assess a separate aspect of the participant's experience efficiently. These included assessments of concentration which were used in Khan and Sunderstroms' 2004 study of vibration annoyance in relation to task difficulty. Loudness, and comfort which were used in Quehl (2001) as an aspect of determining audio comfort. Acceptability of intensity of audio which was used in the assessment of Magnetic Resonance Imaging (MRI) noise in McNulty et al. 2009. Finally annoyance which was used in assessments of noise impacts of turboprop flyover noise by McCrudy (1988).

1. How Intense was the audio?

Unacceptable intensity [1---2---3---4---5---6---7---8---9---10] Completely acceptable intensity

2. How loud was the audio?

Unacceptable level [1---2---3---4---5---6---7---8---9---10] Completely acceptable level

3. During the procedure you felt

Unbearable discomfort [1---2---3---4---5---6---7---8---9---10] Completely Relaxed

4. Did you feel any of the following, if so to what extent?

a) General discomfort

Absent [0---1---2---3---4---5---6---7---8---9---10] Present

b) Headache

Absent [0---1---2---3---4---5---6---7---8---9---10] Present

d) Difficulty concentrating

Absent [0---1---2---3---4---5---6---7---8---9---10] Present

Figure 20: Comfort questionnaire

3.3.2 – Experiment 1 – Design

This study was a repeated measures design, meaning that participants carried out all conditions. The independent variable was sound intensity of pulsed white noise which had two levels, 70 dB(A) and 76 dB(A). These sound intensities were chosen as they represented both a range of sounds noted by Ivoševic, Miljković, and Krajček (2012) as being part of turbo-propeller aircraft cruise noise intensities. The difference in sound intensity also represented as a steep part of the curve in performance found by Meyer and Morse (2003) with the SIT. This sound intensity was measured before each test with a sound level meter and a synthetic ear to simulate and calibrate the sound intensity the participant would be exposed to. The dependent variables were cognitive performance and comfort ratings. These were produced from the PEP (Fig 17) (average completion time 6.8 minutes), and SIT (Fig 19) (average completion time 6.2 minutes) when looking at cognitive performance, and scores on the comfort questionnaire (Fig 20) when examining comfort ratings. The use of the PEP in Experiment 1 was to investigate if an increased trial number would increase sensitivity to sound intensity. The order of the tests and the sound intensities were counterbalanced to avoid order effects or learning effects.

The SIT signal intensity was calibrated at 30 dB. This intensity was chosen as, while it is not the average speech level of humans (60dB), it provided a set level of difficulty in completing the task. The participants were given a short practice period for each test before recording of data to further avoid practice effects and to ensure participants understanding of the correct procedure to carry out the tasks.

The audio stimulus of pulsed white noise was then played, and the participants were given the cognitive tasks to complete. Once the participant had finished the tasks they were then requested to fill out the comfort questionnaire (Fig 20). Following the first condition the noise intensity was re-calibrated to the second intensity and the participants were given a short 5-minute break before they were to carry out the second condition. The participants were then asked to complete the tasks and questionnaire again.

The participants were predominantly students from the University of Liverpool, selected on the exclusion basis of a self-reported lack of normal hearing and/or vision through an opportunity sample. In the initial study, Experiment 1 the number of participants used was n=10. Participants had a median age of 22.

The materials used for this study included a Toshiba satellite laptop, a set of Sennheiser closed ear headphones, while Matlab was used to control the audio output and graphical user interface for the SIT (Appendix B). The audio stimulus was pulsed white noise produced through the RPvdsEx programme, this programme is a real-time processor and visual design studio which allows for simple but precise audio signal production (http://www.tdt.com/rpvdsex.html). Pulsed white noise was chosen as it provided a neutral interference audio stimulus which contained no semantic content.

3.4 Statistical analysis

3.4.1 Power Calculations

Power calculations use Cohen's d and are therefore heavily dependent on the mean and standard deviation of the pilot data. If the estimate used is incorrect, the required sample size will also be wrong. This study is particularly interested in modelling the effect of changes to the task and participant pool that further reduces the amount of observed data that is available. Bootstrapping is a well-established method to estimate the population mean and variance the basis of a small pilot samples (Kulesa, Krzywinski, Blainey, and Altman 2015). The method is used to assign measures of accuracy to sample estimates to estimate the sampling distribution of almost any statistic using only very simple methods.

The overarching concept is to estimate properties of a sampling distribution, here the mean and variance, by repeatedly sampling an approximating distribution (our pilot sample data). The empirical data of ten participants that each run two experimental runs of 200 trials is used. It is assumed in these cases that our sample of participants is representative for the range and distribution of participants that are likely to be recruited during the main experiment. This therefore allows for a bootstrap to be carried out. In a computer simulation bootstrapping (10,000 random re-samples) was used to estimate the mean and standard deviation of the sampling distribution for a range of trials and participants.

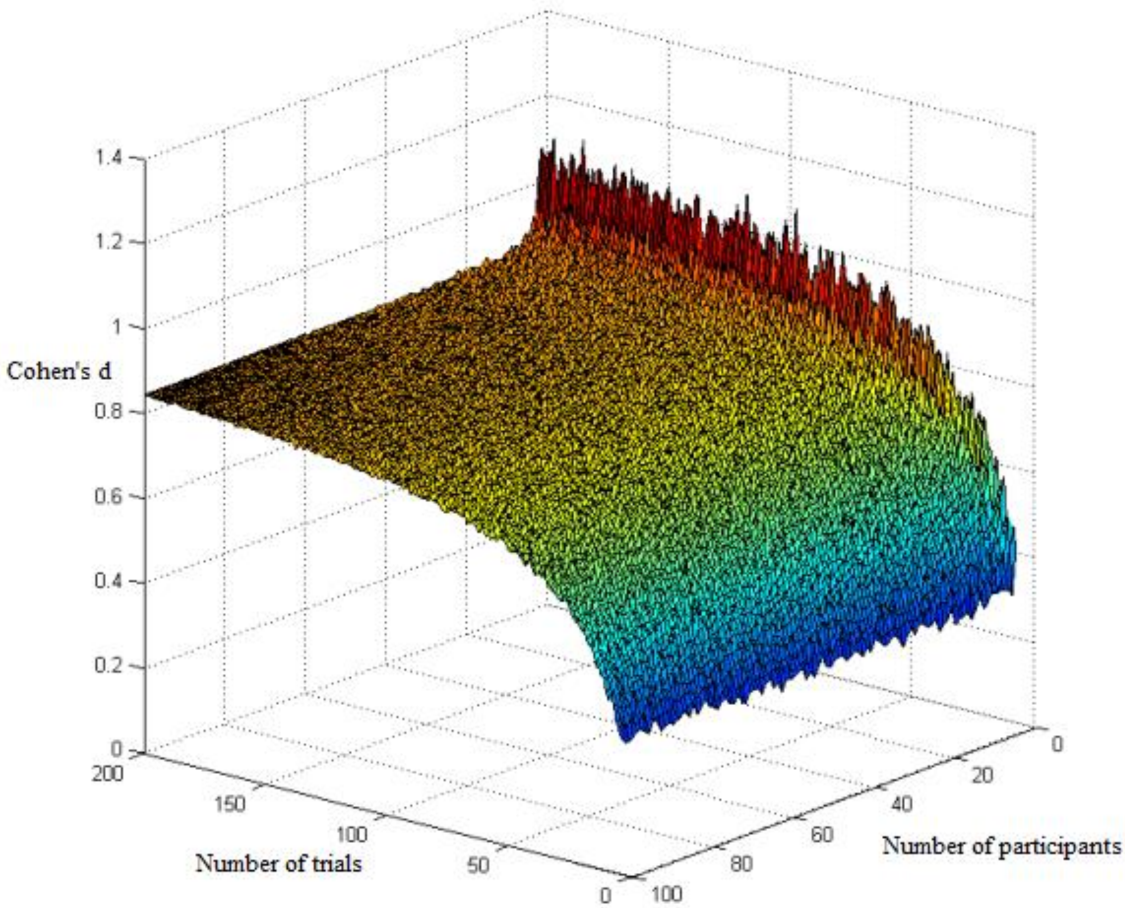


Figure 21 Cohen's d plot

Figure 21 shows Cohen's d against the number of trials (1-200) and for a range of participants (1-100). A first conclusion is that for 10 or more participants and 100 or more stimuli large (d =0.82) effect sizes can be expected. Taking the observed data for 10 participants and 200 trials the power can be estimated (=0.771). Through this the required sample size as a function of effect size a priori for the next experiment can also computed.

Figure 22 shows the statistical power of the experiment for a range of effect sizes (0.82 is an effect that is equivalent to a six dB linear level reduction) and the number of participants tested. For a power of 0.5 (equal chance of false positive and false negative outcomes) it is predicted that approximately 20 participants are needed to reliably detect an effect of half the magnitude that was tested (effect size 0.4, 3 dB) while a level change of 2dB requires 30 participants.

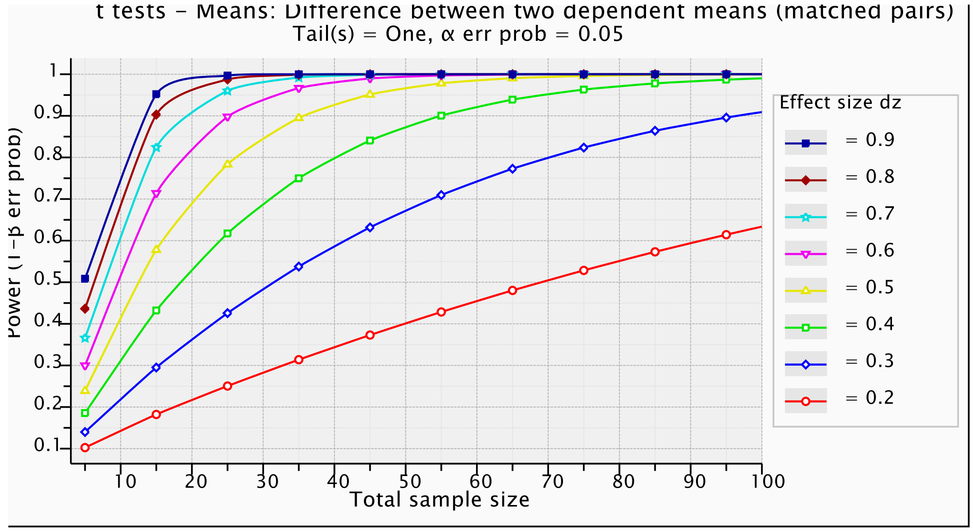


Figure 22 Experimental power as a function of effect size and participant number

3.4.2 Audio Intensity Statistical Analysis

Descriptive statistics were applied to the data to give an overview of the results, including means and standard deviations. Deeper analysis was used in the form of paired student t-tests which provided furthermore specific information on differences between levels of variables.

3.4.3 Speech Identification Task

With the analysis of the SIT the data produced was an overall performance measure of correct responses represented as a percentage, the averages of which can be seen in Figure 23. This data was primarily analysed through paired student t-tests and produced a statistically significant difference between the 70 dB(A) and 76 dB(A) conditions ($t(9) = 2.547$ $p=0.0313$) (Fig 23).

This shows that there was an effect of the increase in intensity of sound level upon participant's performance. This effect shows that higher noise intensity cause increases in identification error rates. This is displayed by the direction of the data.

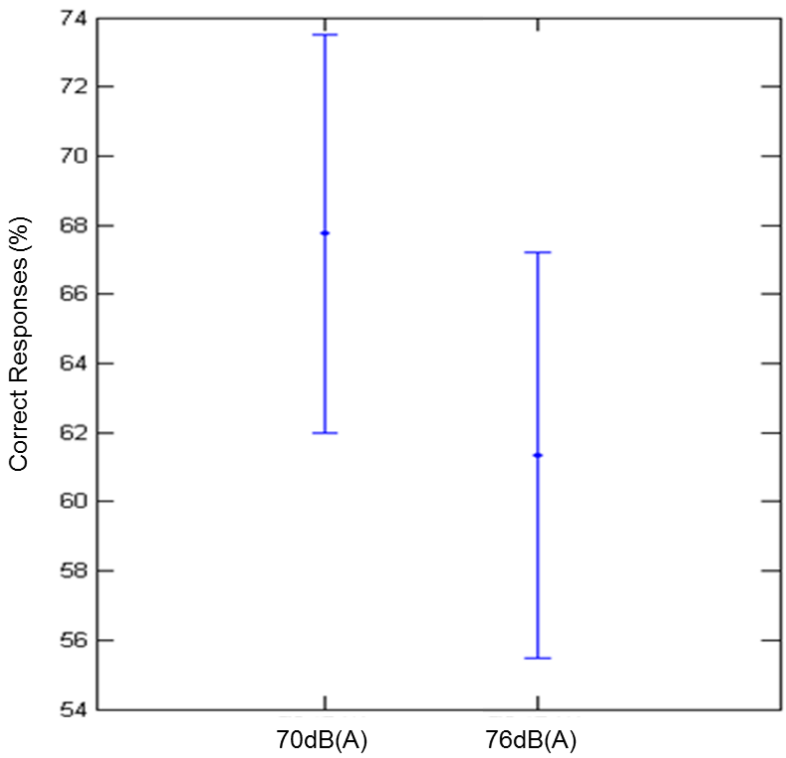


Figure 23: Difference in correct response rate between 70 dB(A) and 76 dB(A) in SIT: error bars represent standard error, (p=0.0313).

3.4.4 Phillips Effect Paradigm

The PEP showed no significant effect on performance between the two noise intensities ($t(9)=0.182$, $p=0.859$). While they did show significances between complexity and delay this lack of effect in respect to noise intensity showed that this task did not reflect the change in audio stimuli.

3.4.5 Comfort Questionnaire

The comfort questionnaire showed significant results across all questions indicating that there was a significant impact of intensity of the sound (Tables 3 and 4). These differences showed participants experienced significantly more discomfort (both sound and general) and difficulty concentrating and also noticed the change in audio stimuli between conditions as can be seen in the “Loudness/Acceptability” and “Intensity of audio” questions. These results showed a noticeable effect of intensity on participant’s comfort (Tables 3, 4 and Fig 24).

Within Figure 24 it can be seen that questions relating to “Intensity”, “Loudness/Acceptability”, and “Audio discomfort” were scored in a manner indicative of 10 being less loud or more comfortable and 0 being scored as “more loud” and less comfortable while the reverse is true of the questions relating to “General discomfort”, “Headache”, and “Difficulty concentrating” this is reflected in Figure 24 in which the vertical bisection of the figure shows the point of question reversal.

Table 3: Questionnaire student t-test results (Experiment 1)

Question	t-statistic	Degrees of Freedom	p-value
Intensity of audio	6.3959	26	>0.0006
Loudness/Acceptability	4.9301	26	>0.0006
Audio discomfort	5.4014	26	>0.0006
General discomfort	3.2731	26	>0.0006
Headache	1.0810	26	>0.0006
Difficulty concentrating	3.5375	26	>0.0006

Table 4: Means and standard error of comfort ratings

	Average rating		Standard error	
Question	70dB	76dB	70dB	76dB
Intensity of audio	7.34	6.44	0.46	0.49
Loudness/Acceptability	5.74	4.96	0.39	0.40
Audio discomfort	6.19	4.96	0.38	0.40
General discomfort	4.11	5.52	0.41	0.49
Headache	2.07	3.30	0.48	0.60
Difficulty concentrating	4.45	6.00	0.44	0.51

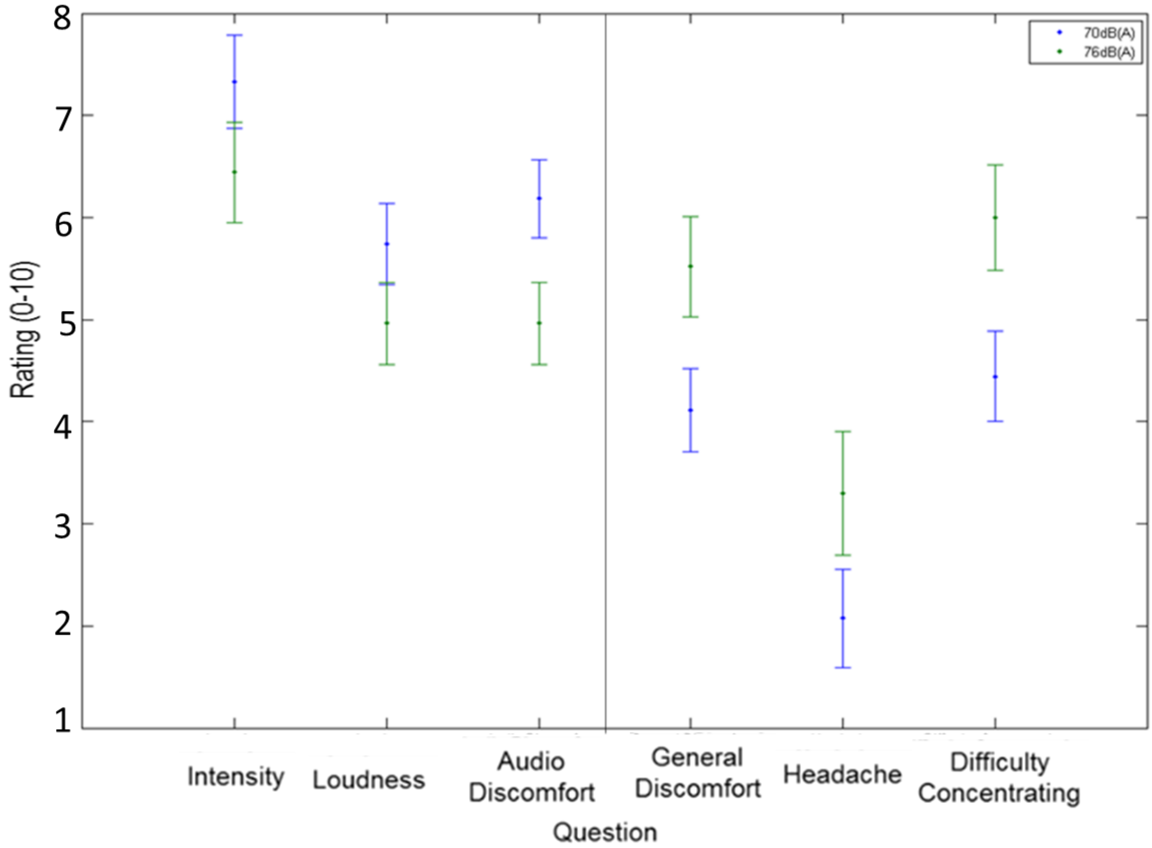


Figure 24: Displaying average ratings for each question on the comfort questionnaire at both 70 dB(A) (Blue) and 76 dB(A) (Green). The graph is bisected to indicate the reversal of ratings on the questionnaire, left being 0-Low 10-High, right being 0-High 10-Low. Error bars represent standard error. Each question showed significant differences due to the 6dB(A) change (significance can be found in Table 3).

3.5 – Brief Discussion of Experiment 1

In Experiment 1 it was found that the increases to the SIT's number of trials led to an increase in sensitivity of the measure. This allowed it to be used effectively as an objective test to determine the impact of noise intensity on participants. This study also showed that the increase in trial number for the PEP did not have the same effect and so this test was dropped from the test battery. This was attributed to the PEP not being a stimuli specific task instead assessing the impact on working memory. This led to the novel conclusion, based on the results obtained, that objective metrics which use task performance to measure comfort are required to be stimuli specific. The impact of discomfort while significant are not found to cross into other stimuli based task performances with enough impact to be reliably measured in an experimental setting.

Within this study a comfort questionnaire was also tested. This questionnaire showed it was an effective and reliable method of measuring participant discomfort (Fig 24) with consistent effects across participants in the different noise intensity conditions. Table 5 shows a summary of the tests assessed and their utility in this research.

Table 5: Summary of test selection results

Test	Pilot	Experiment 1
ToL	Rejected	N/A
BTA	Rejected	N/A
PEP	Accepted	Rejected
SIT	Accepted	Accepted
Questionnaire	N/A	Accepted

3.6 Experiment 2 – Assessment for Spectral Differences

Experiment 2 was designed to assess the use of the SIT and questionnaire when measuring participant comfort for different noise spectra. This data was used to determine the usefulness of these metrics in showing the impact of spectral differences on comfort. This experiment was carried out with use of sound files provided generated by UoL and NLR using the Dowty Propeller designs. The sound files used in this experiment are taken from a single position on the propeller and are from an early version of the sound files provided (as noted in the introduction).

3.6.1 Tonal Experiment

The design was a within participants design with two dependant variables; comfort ratings and performance on the SIT (Fig 19). The independent variable was spectral differences which had three levels; Base, Staggered, and Unequal. These spectral differences sound files were provided through work carried out by UoL CFD researchers and NLR's Human Factors group. These sound files were preliminary sound files which were from a single space on the simulated propeller and had not yet been put through a transfer function by NLR as this was what was available at the time.

The participants were students from the UoL, taken through an opportunity sample. The study used an exclusion criterion of normal hearing and vision this was self-reported. The number of participants used was n=10. Participants had a median age of 21.

The noise intensity of the three stimuli was calibrated to 70 dB (A), as this was the lower end of the range determined for turboprop aircraft cruising noise intensity by Ivoševic, Miljković, and Krajčeks (2012). The sound presentation was calibrated in the same way as it was in Experiment 1 and the pilot study, using a sound level meter and a synthetic ear. The three stimuli, Base, Staggered, Unequal, represent different spectra and were counterbalanced to avoid order effects. The SIT intensity was calibrated to 30 dB to make the SIT a challenging task. 30 dB was chosen as this tends to be where participants score around 50% correct. This is where it would be expected that the visible effects would be largest (Meyer et al. 2003). Participants ran a practice block on the task before data was recorded to minimise practice effects and to ensure the participants understood the task.

The audio stimuli were played to the participants one by one. During each audio stimuli the participants carried out one iteration of the SIT and then filled out a comfort questionnaire after each iteration of the SIT.

The background audio stimuli for this study was provided by NLR and UoL, developed though CFD; these were the first set of sound files from the ImPAcTA Project.

3.6.2 Spectral differences Statistical Analysis

The statistical analysis carried out on the data produced by this study was primarily focused on the use of paired student t-tests. These tests allowed the data to be assessed for any statistically significant differences between the sound files in either task performance, or subjective experience by the participant of the sound files.

3.6.3 Speech Identification Task

There was a significant effect of change in spectral differences upon speech identification performance; this was found through the application of paired t-tests and an examination of descriptive statistics (Fig 25).

There was a significant difference between; the Base and Unequal spectral differences (df(8), t=14.0645, p>0.003), Base and Staggered (df(8), t=20.5129, p>0.003), and Staggered and Unequal components (df(8), t=4.6417 p=0.0051). These differences showed performance to be best in the baseline condition in comparison to both the Unequal and Staggered conditions, while the Staggered condition showed a small increase in performance over the Unequal condition.

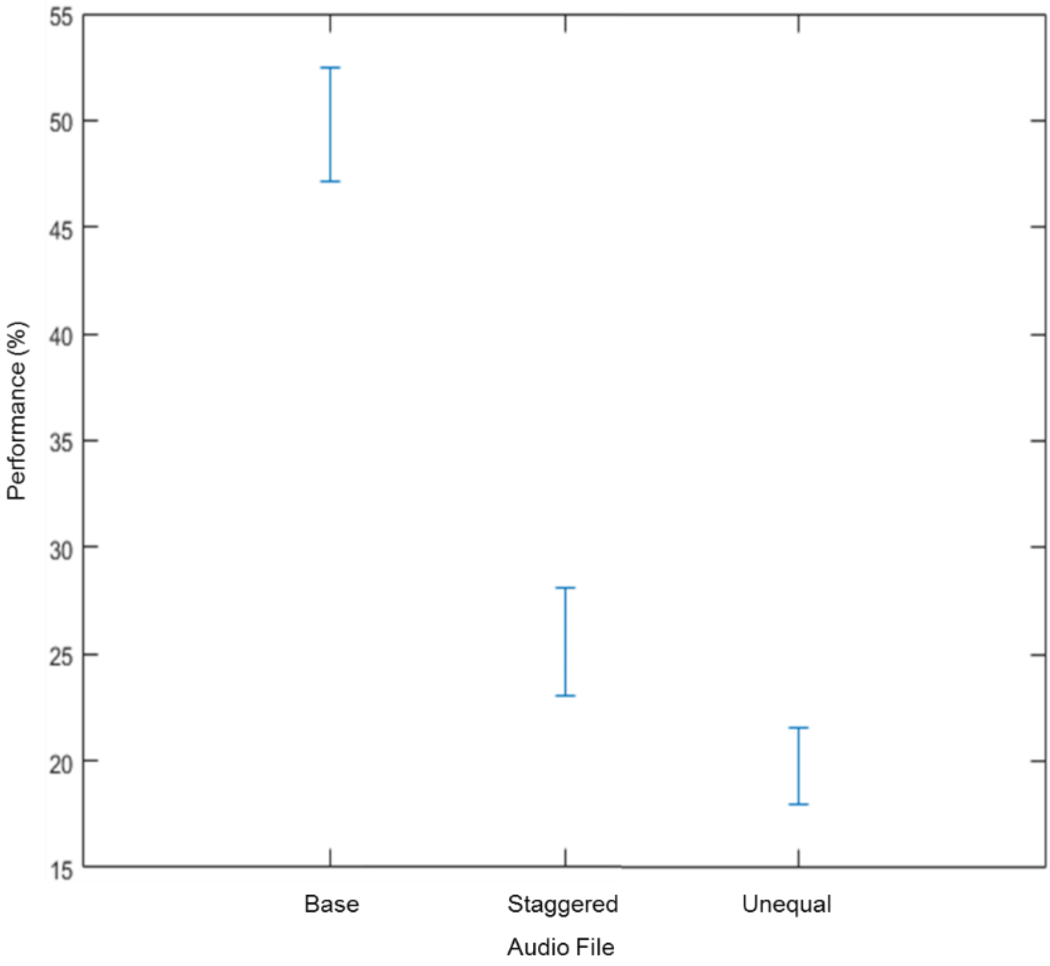


Figure 25: Displaying changes in performance on the SIT across spectral differences audio files Base, Staggered, and Unequal. Each spectral difference change had a significantly different impact on performance in comparison with the other spectral differences. Error bars represent standard error. Base/Staggered p>0.003, Base/Unequal p>0.003, Staggered/Unequal p=0.0051.

3.6.4 Spectral differences Comfort Ratings

There was a significant effect of change in spectral differences upon comfort ratings. This was found similarly through the use of paired t-tests and examination of descriptive statistics (Fig 26). Significant differences were found between the Base and Unequal (df(8), t=11.7954, p>0.003), and Base and Staggered spectral differences (df(8) , t=9.7980 , p>0.003). These differences showed comfort to be highest in the baseline condition in comparison to both the

Unequal and Staggered conditions, while the Staggered and Unequal condition showed no significant change in comfort when compared directly.

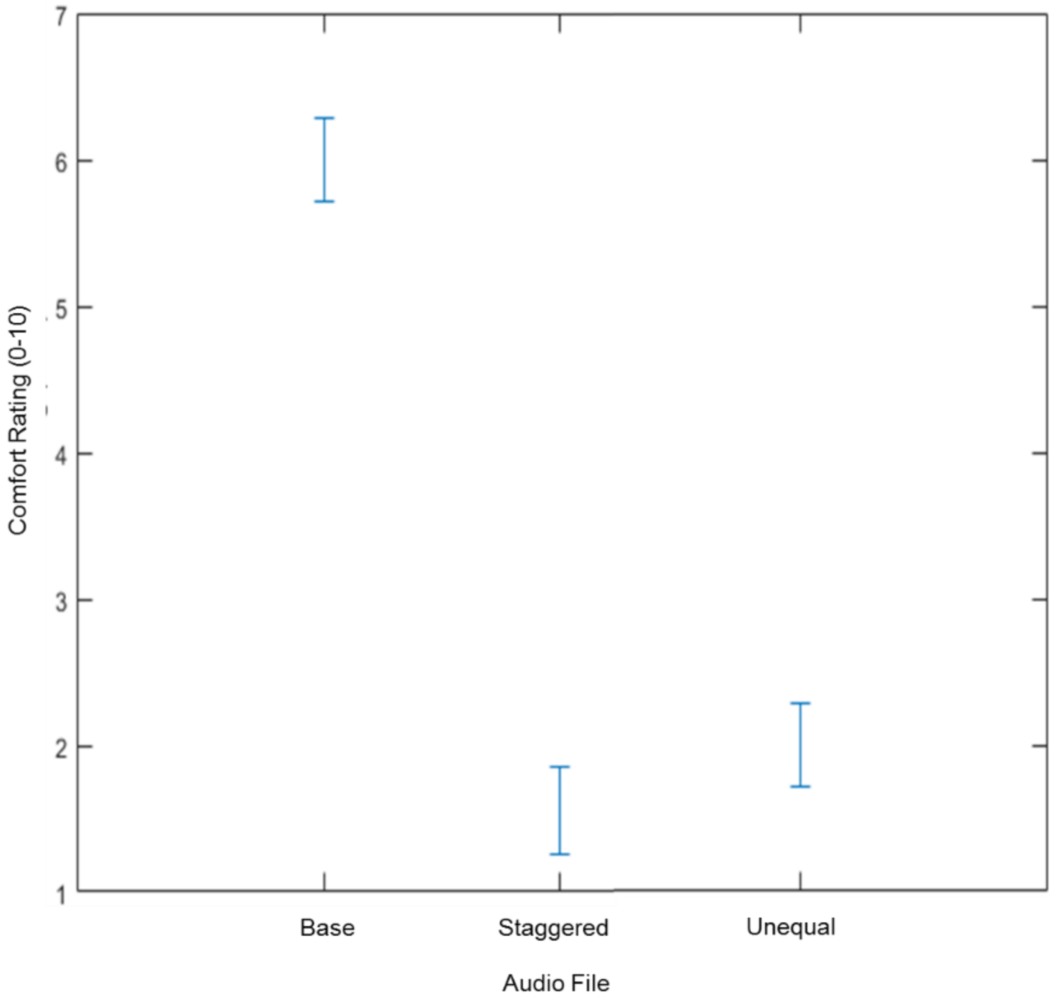


Figure 25: Displaying changes in comfort ratings on comfort questionnaire across spectral differences audio files Base, Staggered, and Unequal. The Base spectral differences were significantly different from both the Staggered and Unequal Spectral differences. Error bars represent standard error. Base/Staggered p>0.003, Base/Unequal p>0.003.

3.7 Discussion

These studies were used to assess the function of potential behavioural tests as measures of comfort via task performance. In addition to this they were used to determine the effectiveness of a simple comfort questionnaire.

The variable used in the first study was noise intensity and the change of 6dB from 70 dB(A) to 76 dB(A). The results show that the SIT captures a change in intensity as has been shown in Meyer et al. (2003) with an increased intensity leading to a significant drop in performance in identification of nonsense syllables (Fig 18). The change in intensity was also reflected in the comfort questionnaire which indicated that with an increase of 6dB participants experienced significantly more discomfort both audio specific and general as well as reporting a higher difficulty concentrating on the given task. These results both show the strong effects of a change in intensity of audio stimuli, a common factor within audio discomfort.

The decrease in participants' performance in the SIT under the 76 dB(A) condition would indicate that the process of correctly identifying the speech sounds was interfered with more than in the 70 dB(A) condition. The negative impact of noise intensity on participants' performance (Fig 25) is reflected in the comfort questionnaire with lower comfort and concentration scores being reported in the higher intensities (Fig 26). These results lead to the conclusion that interference with a task is one of the aspects of a stimuli which leads to discomfort.

The manipulated variable in the second study was spectral differences of the audio signal. The sound files used in the study had three different formats; Base, Staggered, and Unequal. The analysis displayed that comfort ratings and task performance in the SIT were impacted significantly by the changes in spectral differences.

The decreases in task performance (Fig 25) were, as with the sound intensity study, mirrored in the comfort ratings provided by the participants (Fig 26). These reductions would again indicate that there is merit in using a task performance metric in tandem with a subjective assessment of participant's comfort.

With the conclusion that interference with a stimuli specific task can lead to discomfort, it therefore follows that a useful objective metric of comfort would be a metric which is interfered with by the stimuli assessed. This is not to say that a poor score in a metric immediately indicates discomfort, instead that with controlled environments comparisons between performances in two conditions accompanied by a comfort questionnaire can show the objective impact of discomfort upon a participant, using the subjective questionnaire data as a guide. This concept fits closely with other research into comfort which use objective measures in tandem with complementary subjective metrics for example Iachini et al. (2012) and Mellert et al. (2008), who used behavioural and physiological measures respectively to provide an objective measure of the impact of discomfort in relation to changes in stimuli. These could be considered to be related to Straker el al. (1997), who put forward that stimuli that elicit a discomfort effect would also cause an impairment of cognitive performance. Additional links to cognitive performance being a useful predictor of comfort is the tenets of comfort put forward by Metzger (1994), who includes the tenets of ease of carrying out tasks and efficiency of tasks being completed. These tenets from Metzger led this research to use objective stimuli specific tasks. The results from this research show that objective tasks are important factors in the assessment of comfort.

This research finds novelty in the highlighting of the importance of stimuli specific tests to determine the impact of stimuli. This can be seen in the success of the SIT where other non-audio-based tasks showed no interference during discomfort (Table 6). The use of stimuli specific tasks has been absent from much of the literature and thus the use of the SIT is a relative novelty. The data was also used to provide a novel application of task performance as a measure of comfort which can be fitted to the model described in Chapter 5.

This conclusion was used throughout this thesis as the basis for the objective testing, using stimuli relevant objective metrics to inform the impact of stimuli on participant's comfort when guided by subjective questionnaire data. The understanding of this link between objective task performance and comfort is also further expanded upon when discussing the model of comfort proposed later in this thesis (Chapter 5).

Table 6: Updated Table 5 displaying test selection results including Experiment 2 results

Test	Pilot	Experiment 1	Experiment 2
ToL	Rejected	N/A	N/A
BTA	Rejected	N/A	N/A
PEP	Accepted	Rejected	N/A
SIT	Accepted	Accepted	Accepted
Questionnaire	N/A	Accepted	Accepted

Chapter 4: Impact of Simulator Fidelity on Participant Comfort

4.1 Introduction

Comfort measurements as tools allow for the assessment of the impact of sources of physical stimuli, which can negatively affect experience, to be undertaken, as seen in the previous chapters. However, comfort is not exclusively impacted by the physical properties of the stimuli, instead contextual factors such as environment also have a notable part to play in this interaction of how we, as humans, experience stimuli which subsequently lead to perceived comfort or discomfort. This chapter revisits literature relevant to these issues and addresses these contextual factors using varying simulation fidelity on two different fronts. The assessment of comfort in relation to the level of fidelity in a given simulation allows for a demonstration that metrics can show changes in comfort outside of those caused by primary discomfort stimuli. This also allows for a novel exploration of simulation fidelity as a significant factor in perceived comfort. The results from this study also further feeds into the proposed model of audio comfort (Chapter 5), in its development and assessment of assumptions of the model.

4.2 Experiment to Assess Impact of Simulation Fidelity on Participant Comfort

A tanker tracking task (Meyer et al., 2012) was used as a measure of task performance which was 'embedded' within a simulation. This integral link between the simulation and task allowed for an increased impact of both PI and Psi. This was useful as when focusing on the assessment of the effect of fidelity it forced participants to focus on the simulation to complete the task. This aspect follows on from the stimuli specific task relevance that was found in Chapter 3, incorporating a task such as this into the stimuli that were being manipulated. Participants were asked to track a refuelling nozzle in the HELIFLIGHT simulator (Padfield and White, 2003) (Figs 27-29) that was restricted to vertical motion with the use of a helicopter control lever (Fig 27). Performance was quantified through tracking error. This error was determined as the distance between the participant's tracking point and the target in virtual space, the performance error was then converted to a percentage (i.e. poor performance = high percentage value). This was achieved by averaging the distance the participant's reference point (gull-wing) was from the target during the task, from here a percentage to represent their performance could be calculated. This percentage was taken from each condition to provide context for level of performance. This task was chosen as to track the target closely participants had to build a model of the behaviour of the helicopter in response to collective motion and also had to learn to disambiguate target motion from self-motion. Meyer et al. (2012) had shown that sound can provide useful cues for participant to employ in the issue of disambiguation. An example of a flightpath can be seen in Figure 28 (Meyer et al. 2012).

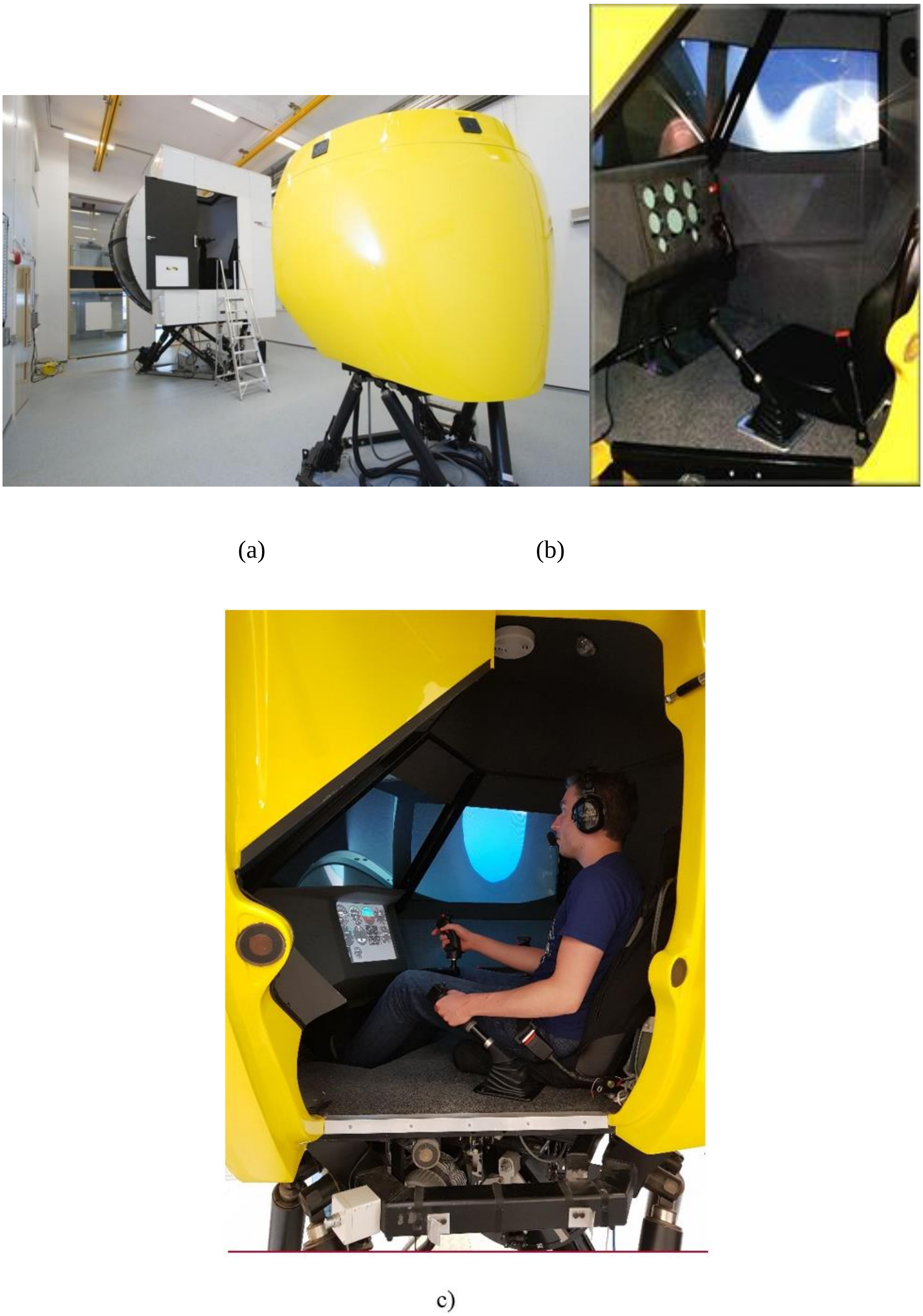


Figure 26: HELIFLIGHT simulator (a, right foreground) and flight controls (b) participant at controls (c)

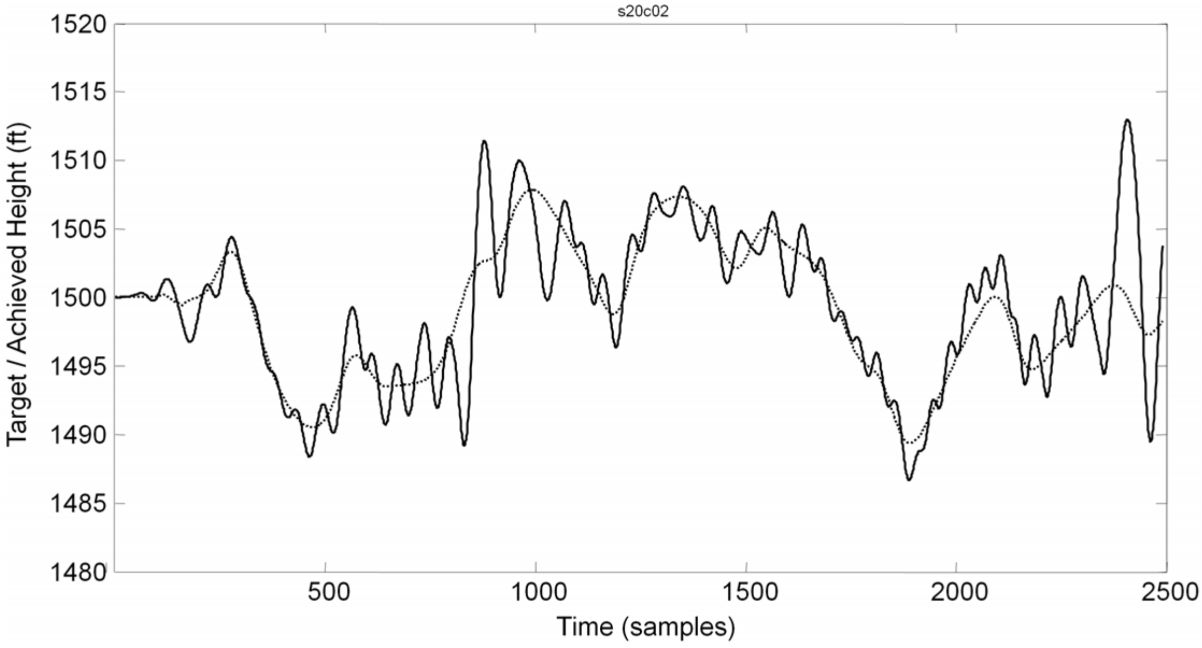
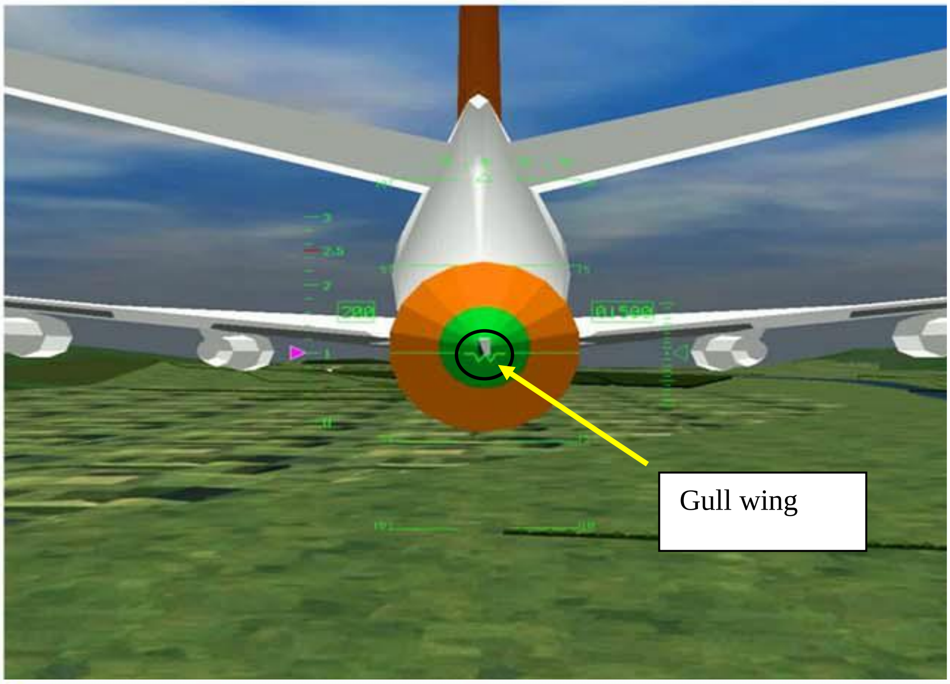


Figure 27: Sample flightpath (continuous line) and basket motion (dotted line) for a difficult flying condition. This participant oscillates around the target position. The median absolute distance between actual flying height and the target position is used as an error measure (Meyer et al. 2012)

a)



b)

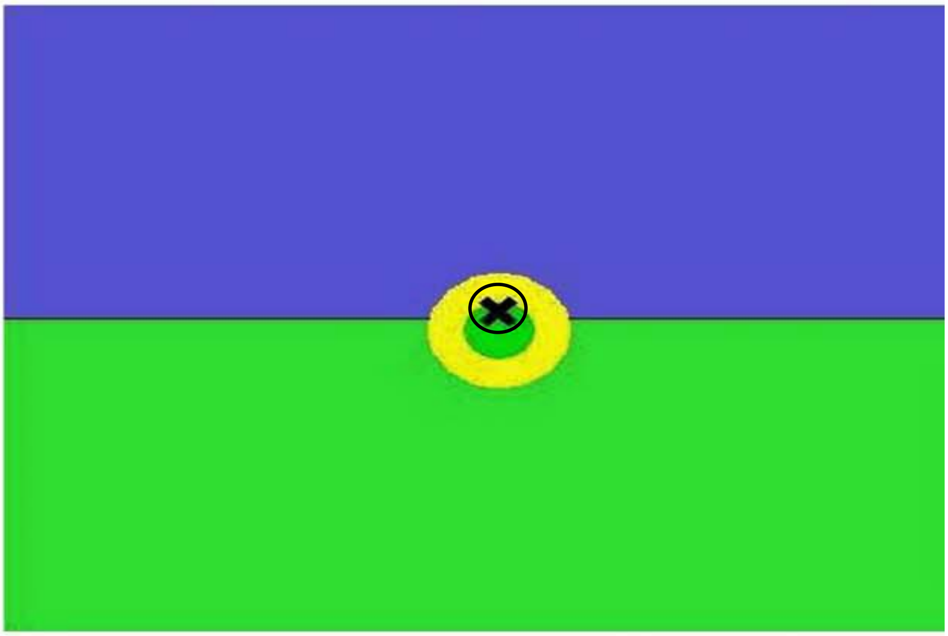


Figure 28: Visuals produced in Tanker Tracking Task: a) high fidelity, b) low fidelity. The aim of this task was to keep the gull wing (circled) (a) or X (b) as close to the centre of the circle as possible.

4.2.1 – HELIFLIGHT

The HELIFLIGHT flight simulator used in the high-fidelity environmental aspect of the study is a turnkey and re-configurable flight simulator which is used for industry relevant research projects and to support the undergraduate Aerospace teaching and research activities at the University of Liverpool. The HELIFLIGHT has five key components:

1) Selective fidelity, aircraft-specific, interchangeable flight dynamics modeling software, FLIGHTLAB with a real-time interface (PilotStation)

2) Six degree of freedom motion platform (Maxcue)

3) Four axis dynamic control loading (Loadcue)

4) A three-channel collimated visual display for forward view, plus two flat panel chin windows, providing a wide field of view visual system (Optivision), each channel running a visual database

5) A re-configurable, software-generated head-down and head up display using Engenuity Technologies VAPS software v6.3.1.

4.2.2 Questionnaire

The questionnaires detailed below were used in this study to determine the impact on comfort and presence the simulation and stimuli had upon the participants. This questionnaire combined both the presence questionnaire detailed below and a short-form version of the previously used comfort questionnaire. This questionnaire allowed for an effective measurement of both presence and comfort without having participants becoming fatigued.

4.2.2.1 Presence Questionnaire

Participants were asked to fill out a modified presence and immersion questionnaire with a 10-point Likert scale (Table 7), this questionnaire was a short form of Witmer and Singer's presence metric questionnaire (1994) (Appendix C). It consisted of three of the primary questions which assessed responsiveness, control, and realism. The reason for the inclusion of the presence questions was to ensure that it could be shown that participants were experiencing a sense of immersion and that the fidelity of the simulator was impactful. The short-form was chosen to maintain participant focus over multiple conditions. The reductions included removing questions which assessed aspects which this study was not examining such as "How closely were you able to examine objects". Other questions were excluded as they provided data at on two similar participants which were not the focus of this study such as "How compelling was your sense of objects moving through space?" and "How compelling was your sense of moving around inside the virtual environment?"

4.2.2.2 Comfort Questionnaire

The comfort and presence questionnaire (Table 7) is comprised of questions that are based on previously tested questionnaires, including McNulty et al. (2009), and Oborne (1975). The comfort aspect of this questionnaire was developed for this study to examine a variety of aspects of comfort and experiences related to the understanding and results of discomfort. The questions used assessed perceived comfort, level of annoyance, and ability to concentrate on the task.

Participants were requested to fill out the questionnaire (Table 7) and their responses were recorded. This questionnaire has 3 questions for each theme and follows a standard 10-point Likert scale format ranging from one extreme to the other.

17 participants participated in the experiment, 5 male participants and 12 female ages 19-21 years old with a mean age of 20 years old SD=2.0. The participants were students from the University of Liverpool. The participants were all screened through a self-report system for normal hearing and vision.

Table 7: Comfort and Presence Questionnaire

	Question	Assessment
1	I experienced general discomfort	Comfort
2	I experienced annoyance with the audio	
3	I found it difficult to concentrate on the task	
4	I was in control of events	Presence
5	The environment was responsive to my actions	
6	The sense of moving around inside the virtual environment was compelling	

4.2.3 Definition of Fidelity Types

Within this study the fidelity of a simulation was manipulated to assess if simulator fidelity would impact participant comfort. For this to be a consistent experimental manipulation the fidelity of the simulation was broken down into two types which are detailed below.

4.2.3.1 Audio-Functional Fidelity

Audio-functional fidelity refers to the movement of the sound which was pulsed white noise in either static or dynamic movement. In static movement, the audio stimuli remain at head level during the task within the simulation, while in the dynamic movement the audio stimuli is kept in a fixed point in the simulation and therefore changes dynamically in reference to the participant's control position within the simulation. This was achieved through the stimulus growing louder or quieter depending on the participant's in-simulation movements (see control lever in participant's left hand in Figure 27 (c)). This change in the stimulus was automatically controlled through the simulation, varying the signal as needed and was provided through the headset.

4.2.3.2 Environmental Fidelity

Environmental fidelity indicates the surrounding context in which the participant carried out the tasks in addition to the visual fidelity experienced by the participant. This variable had three levels which indicate high, mid, and low levels of fidelity (Figures 30 and 31);

Simulation (high) – Participants carried out this condition in the HELIFLIGHT simulator with high fidelity visuals identified as the highest fidelity environmental context within this study, in which there were three screens providing front and peripheral views with each screen at a size of 32" providing a total view size of 96"

Simsim (mid) – Participants were in a soundproof booth with screen size of 59" to increase FOV. This was considered the midpoint of environmental fidelity.

PC (low) – Participants were in a soundproof with a smaller screen (Size 22") this decreased FOV and was considered the lowest fidelity environmental context.

The experimental design was a 3 (environmental fidelity) x 2 (audio-functional fidelity). There were three dependant variables; comfort ratings, presence ratings, and performance on a tanker tracking task (all detailed below). The independent variables included;

Environmental fidelity with three levels; high (Simulation), mid (Simsim), and low (PC).

Audio-functional fidelity with two levels; dynamic (High), static (Low).

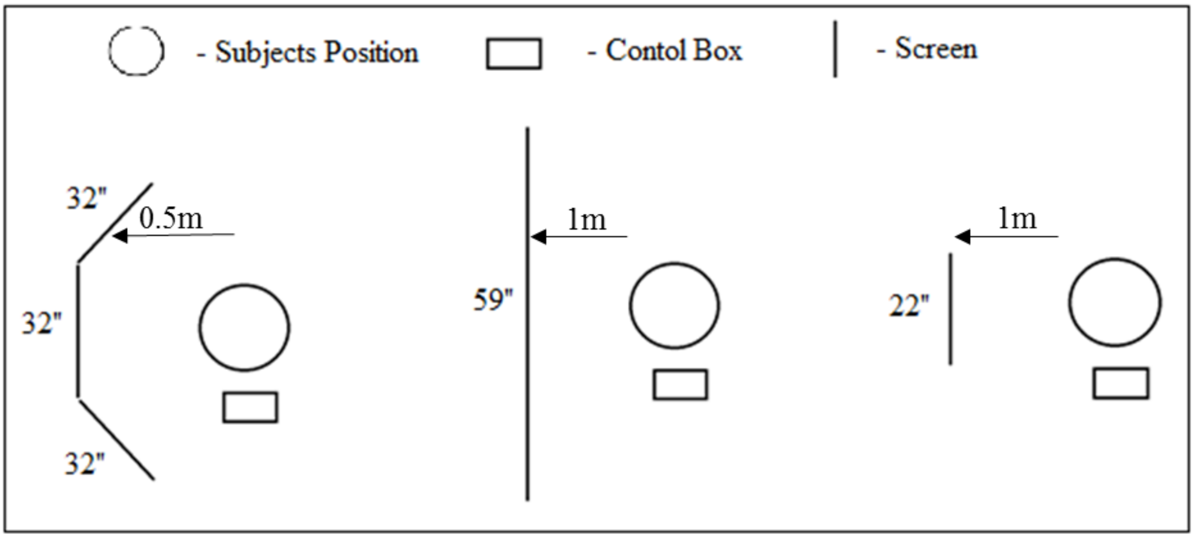
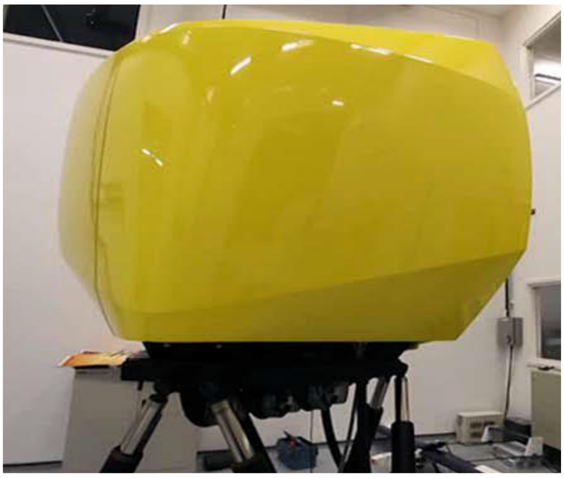


Figure 29: Representation of environmental fidelity (from left to right high, mid, low)



a)

b)

c)

Figure 30: Environmental fidelity conditions a) high, b) mid, c) low.

4.2.3.3 – Experimental Procedure

The amplitude of the audio stimulus was initially calibrated to 70 dB(A) with a TENMA 72-860A sound level meter. The participants were asked to attend three experimental sessions, one session for each environmental condition. In each session, the participants were trained on the Tanker Tracking Task to ensure that the participants were familiar with and able to carry out the task at each point. This training was carried out with the dynamic audio condition running and was undertaken for 15 minutes. The participants then carried out the tanker tracking task for 5 minutes for each noise condition in each session. After each condition, the participant was asked for fill out a comfort and presence questionnaire. Each participant's condition order was counterbalanced to avoid any learning or fatigue effects occurring in the data which would affect the analysis (Table 8).

Table 8: Participant procedure for fidelity experiment

Environmental fidelity condition (counterbalanced)	Audio-functional fidelity condition (counterbalanced)	Task	Test duration (minutes)
High fidelity	Dynamic	Training	15.0
		Tanker Tracking Task	5.0
		Questionnaire	3.0
	Static	Tanker Tracking Task	5.0
		Questionnaire	3.0
Mid fidelity	Dynamic	Training	15.0
		Tanker Tracking Task	5.0
		Questionnaire	3.0
	Static	Tanker Tracking Task	5.0
		Questionnaire	3.0
Low fidelity	Dynamic	Training	15.0
		Tanker Tracking Task	5.0
		Questionnaire	3.0
	Static	Tanker Tracking Task	5.0
		Questionnaire	3.0

4.3 Statistical analysis

The method of analysis for this data was to apply comparison of means to determine the presence of a statistically significant difference between the sets of data. This was carried out through the analysis methods of ANOVAs and paired student t-tests.

Where paired student t-tests were used to analyse the data, corrections were made to the results to take into account multiple comparisons for the tests. The correction method applied was the Bonferroni correction. This conservative correction allows for an accurate view of the data when taking into account the possibility of alpha errors, in which a true null hypothesis is rejected. The Bonferroni correction was chosen as the method of correction due to its conservative statistical nature, when a result is shown to be significant with the Bonferroni correction applied there is extra layer of confidence that the tests are indeed significant unlike other methods of statistical correction.

4.3.1 Environmental Fidelity

The analysis of the environmental fidelity was primarily using ANOVAs and then with an application of post-hoc statistical tests in the form of paired student t-tests with Bonferroni corrections applied. The comparisons are between the three levels of this condition – high fidelity, mid fidelity, and low fidelity.

4.3.1.1 Performance

Within the analysis of participants' performance, it was shown that while there was no main effect indicated however with post-hoc paired student t-tests it was found that there were strong significant differences between high and mid fidelity ($t=2.6969$ $df=16$ $p=0.0474$), and high and low fidelity ($t=2.8386$, $df=16$, $p=0.0354$) (Fig 32).

The results show that changes in environmental fidelity between high and other levels lead to significant differences in the error made by participants on the tanker tracking task. The trend in this data is that with a higher level of fidelity comes a lower error rate on the tanker tracking task, indicating a link between high fidelity and high performance.

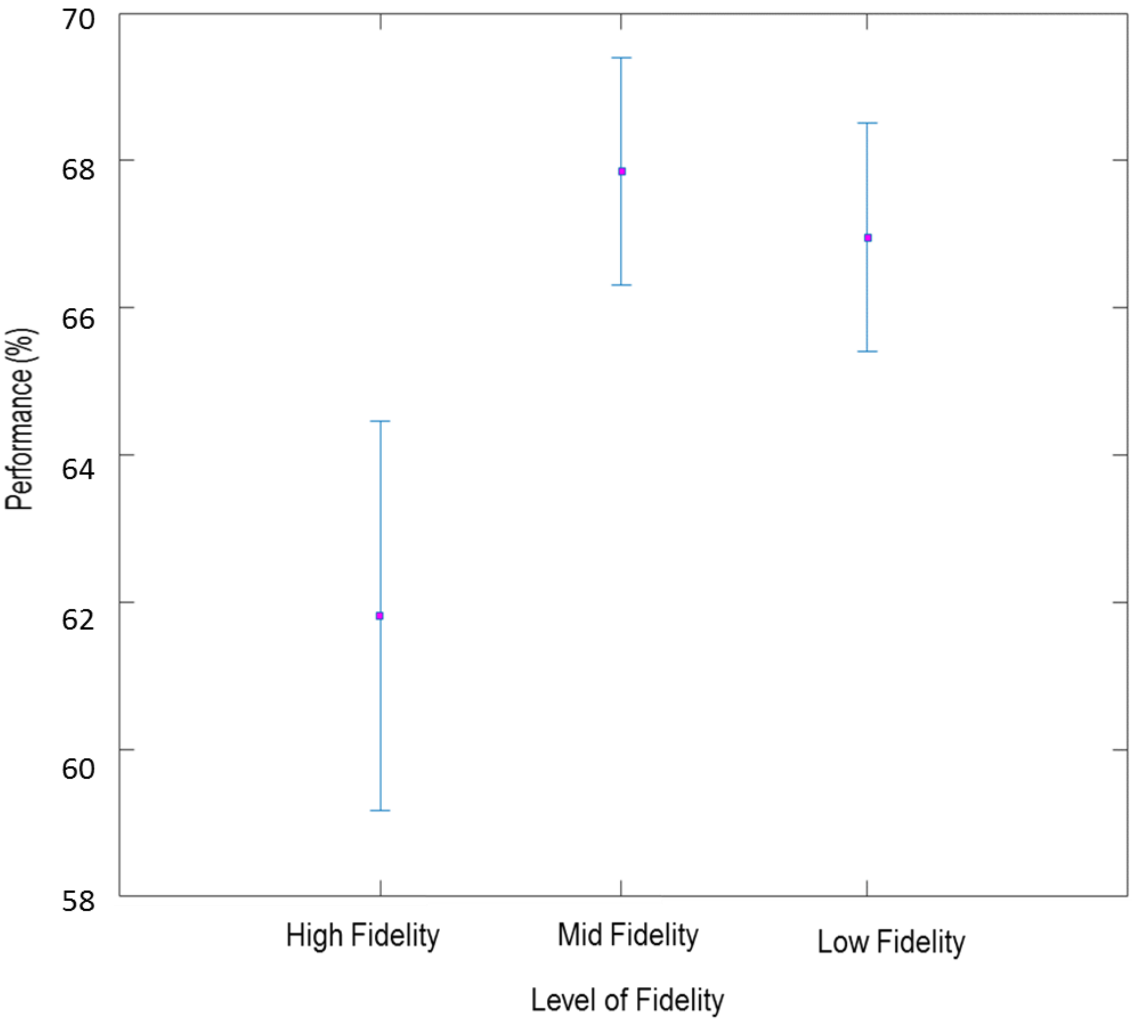


Figure 31: Effect of environmental fidelity on performance. The high fidelity condition showed significant differences in performance on the Tanker Tracking Task when compared to both the mid and low fidelity conditions. Error bars represent standard error. Comparison between high and mid fidelity conditions p=0.0474, comparison between high and low fidelity conditions p=0.0354, performance for the comparison between mid and low fidelity was not significant.

4.3.1.2 Presence

With the analysis of the environmental data within the presence metric there was a main effect found within the ANOVA and with the subsequent post-hoc student t-test analysis it was found that these effects were primarily within the comparisons between high level fidelity and the lower fidelities, with high fidelity against mid fidelity providing ($t=-7.1523$, $df=16$ $p<0.0006$), and high fidelity when compared to low fidelity showing ($t=-7.8521$, $df=16$, $p<0.0006$) (Fig 33). These results mirror those found in the performance error rates and show that, as expected, a higher level of environmental fidelity is responded to by the reporting of a higher level of presence by participants. These results confirm that the variable of environmental fidelity was being manipulated during the experiment.

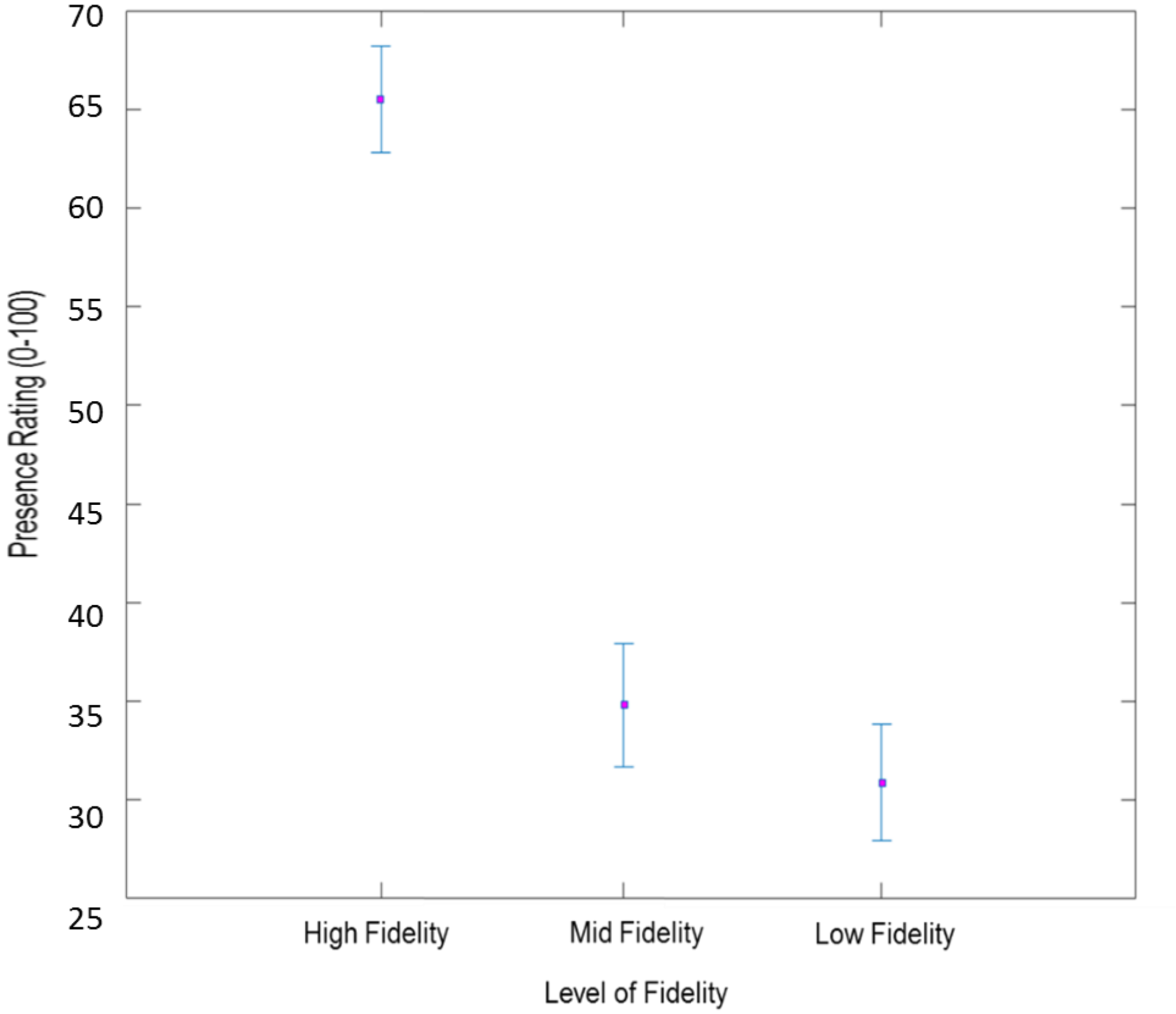


Figure 32: Impact of environmental fidelity on presence. The high fidelity condition showed significant differences in presence on the questionnaire in comparison to both the mid and low fidelity conditions. Error bars represent standard error. Comparison between high and mid fidelity conditions t=-7.1523 p<0.0006, comparison between high and low fidelity conditions t=-7.8521 p<0.0006. Presence for the comparison between mid and low fidelity was not significant.

4.3.1.3 Annoyance

With the examination of Annoyance and the changes caused by environmental fidelity it was found that there was a significant impact upon participants' annoyance ratings. These significant differences were found to be between high and mid fidelity (t=7.2185, df=16, p<0.0006) and high and low fidelity (t=6.4730, df=16, p<0.0006). These significant differences show a noticeable impact upon participants' annoyance ratings when environmental fidelity is manipulated (Fig 34). These results showed that in the high fidelity environment, annoyance was significantly lower than in either the mid or low levels of environmental fidelity.

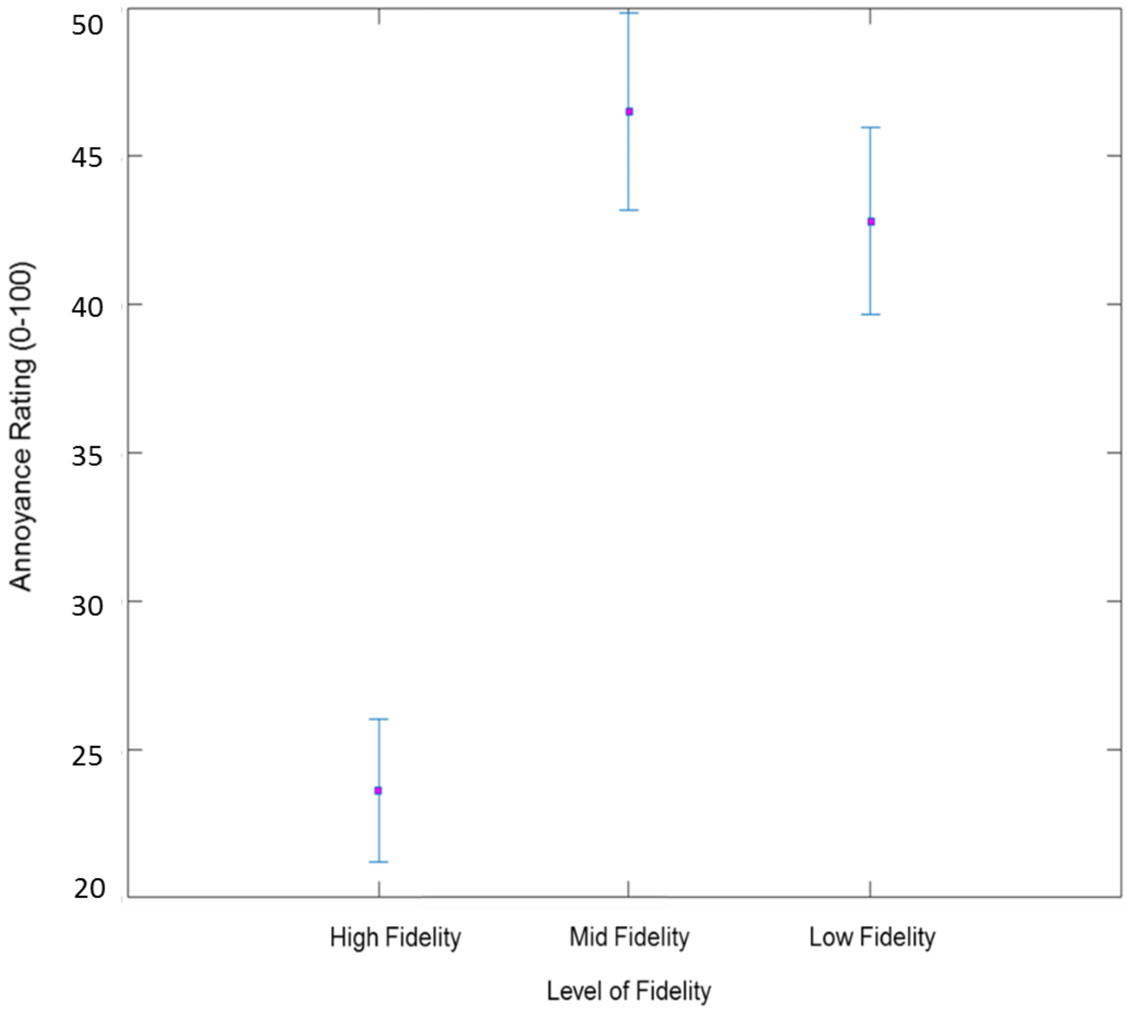


Figure 33: Impact of environmental fidelity on participant annoyance rating. The high fidelity condition showed significant differences for participant annoyance ratings on the questionnaire in comparison to both the mid and low fidelity conditions. Error bars represent standard error. Comparison between high and mid fidelity conditions $p<0.0006$, Comparison between high and low fidelity conditions $p<0.0006$. Annoyance for the comparison between mid and low fidelity was not significant.

4.3.1.4 Comfort

The final metric assessed when examining the impact of environmental fidelity was the comfort metric. The data analysed provided again a main effect which, with corrected paired student t-tests, mirrored the results seen in both the presence and performance metrics. These results show a statistically significant difference between high and mid fidelity ($t=3.7829$, $df=16$ $p<0.0006$), and a similar difference between high and low fidelity ($t=3.7553$, $df=16$, $p<0.0006$) (Fig 35). In Figure 35, 100 is representative of a high level of comfort while 0 is representative of low comfort. These results confirmed that the null hypothesis held regarding there being no effect of environmental fidelity on participants' perceived comfort could be rejected. Overall adding to the argument that with a change in simulation fidelity there is a change in perceived comfort.

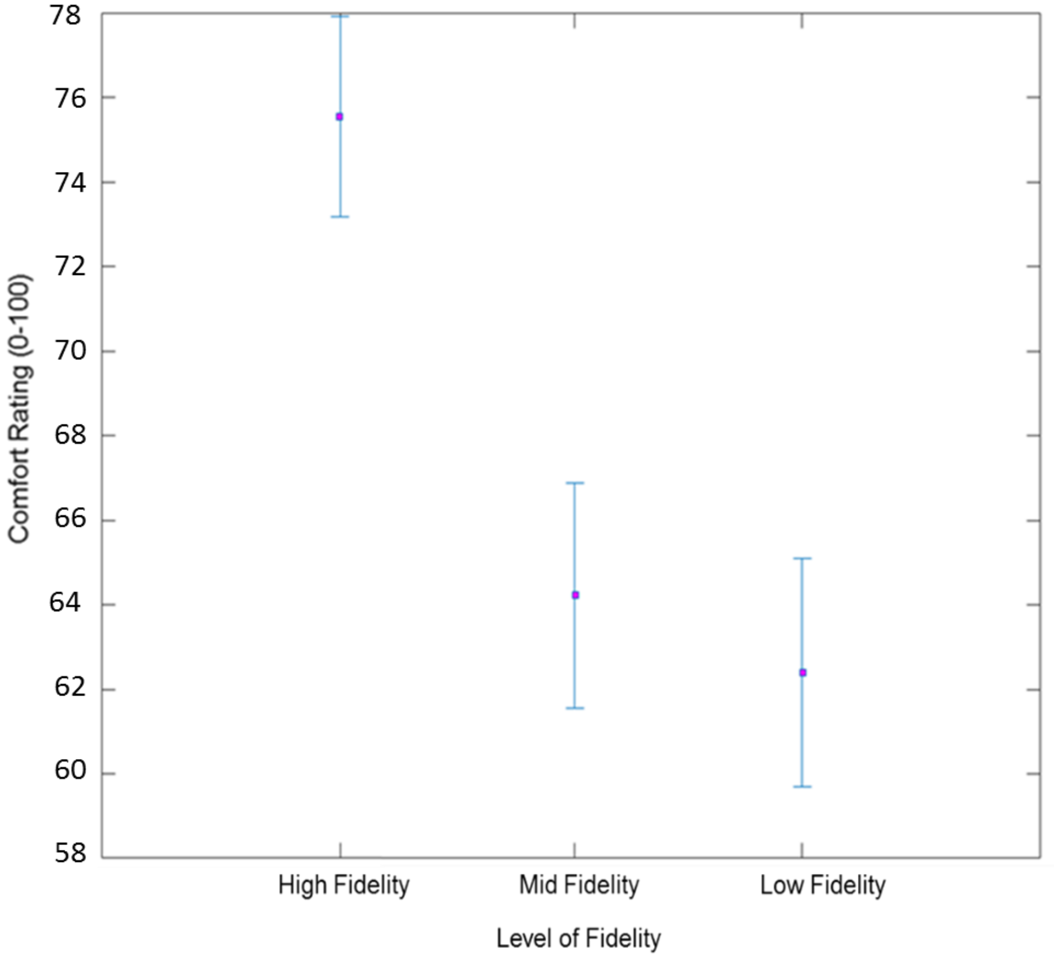


Figure 34: Effect of environmental fidelity on participant comfort ratings. The high fidelity condition showed significant differences for comfort ratings on the questionnaire in comparison to both the mid and low fidelity conditions. Error bars represent standard error. Comparison between high and mid fidelity conditions p<0.0006, comparison between high and low fidelity conditions p<0.0006. Annoyance for the comparison between mid and low fidelity was not significant.

4.3.1.5 Summary of Environmental Fidelity

Paired student t-tests show a significant difference between the high and lower fidelities in: presence, comfort, annoyance, and performance. These statistically significant differences are important to note as these significances are between the same conditions in each metric and all support the hypothesis that an increased level of fidelity has a beneficial effect; decreasing discomfort, task error rate, and increasing presence. These are reflected in the changes to presence ratings, comfort ratings, and performance on the task demonstrated above.

4.4.1 Audio-Functional Fidelity

With audio-functional fidelity the analysis method changed due to the lesser number of levels in this condition. Rather than apply ANOVAs to the data it became more appropriate to apply student t-tests immediately rather than as post-hoc tests. These tests paired the two levels of static which represented a lower fidelity condition to dynamic which represented a higher fidelity sound source more realistic and responsive to the simulation. As there was only one statistical test applied to the data per metric there was no need for a Bonferroni correction to avoid alpha errors.

4.4.1.1 Performance

There were no statistically significant differences when examining the error of the participants carrying out the tanker tracking task ($t=-0.6119$, $df=16$, $p=0.5420$). This indicates that there was no impact of the realistic movement of sound cues on the performance of participants.

4.4.1.2 Presence

Ratings on the presence scale, when compared between the static and dynamic audio conditions analysed with paired student t-tests, showed a statistically significant difference between the two conditions ($t=19.785$, $df=16$, $p>0.001$) (Fig 36). This change between conditions showed that there was an increase of presence ratings in the dynamic condition. This result follows what was expected in that more realistic sound cues which reacted as one would expect in real life leads to an increase in presence ratings.

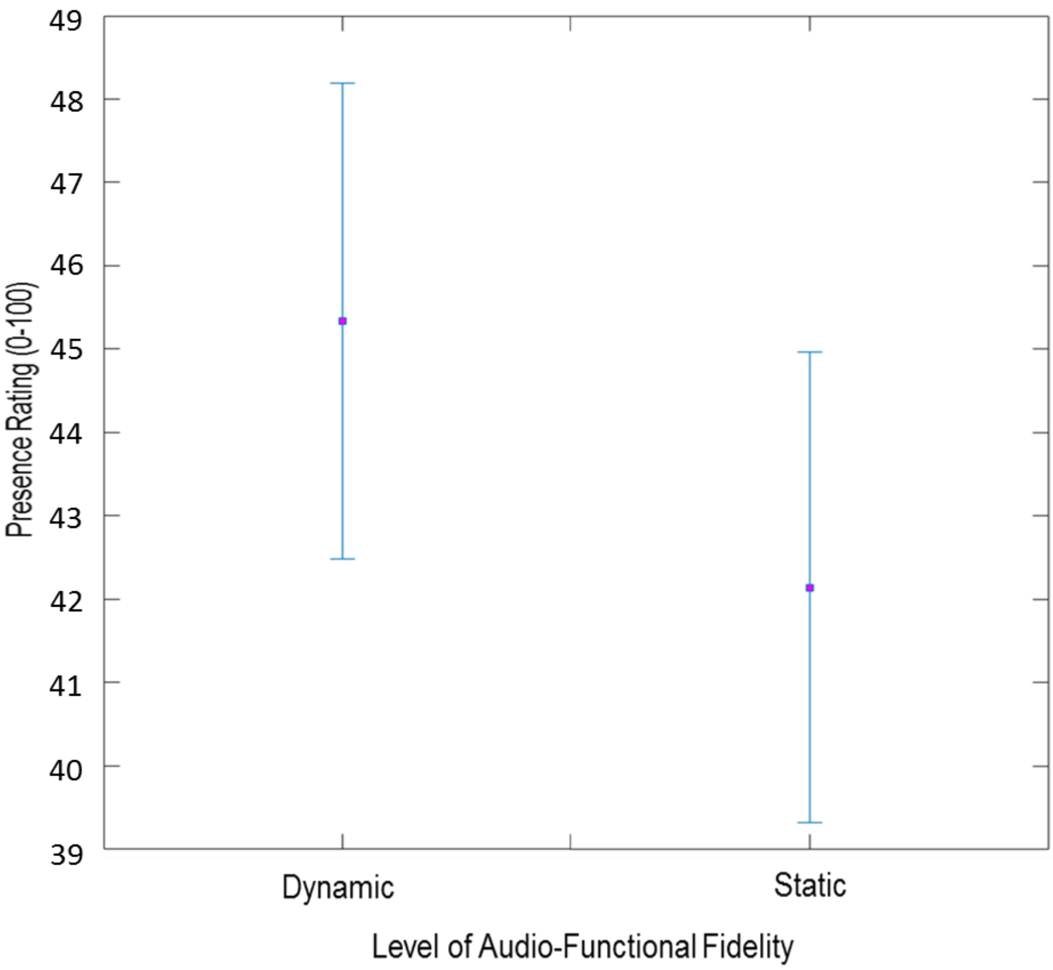


Figure 35: Effect of audio-functional fidelity on presence. There is a significant difference between the dynamic and static conditions, dynamic showing a significantly higher presence ratings. Error bars represent standard error, p>0.001.

4.4.1.3 Comfort

With the metric of comfort ratings being assessed through paired student t-tests, it was found that there was a statistically significant difference between the conditions (t=3.9638, df=16, p>0.001) (Fig 37). The direction of this difference showed a higher level of comfort ratings during the dynamic condition, as with the environmental fidelity this showed that higher comfort ratings were present during the high fidelity conditions.

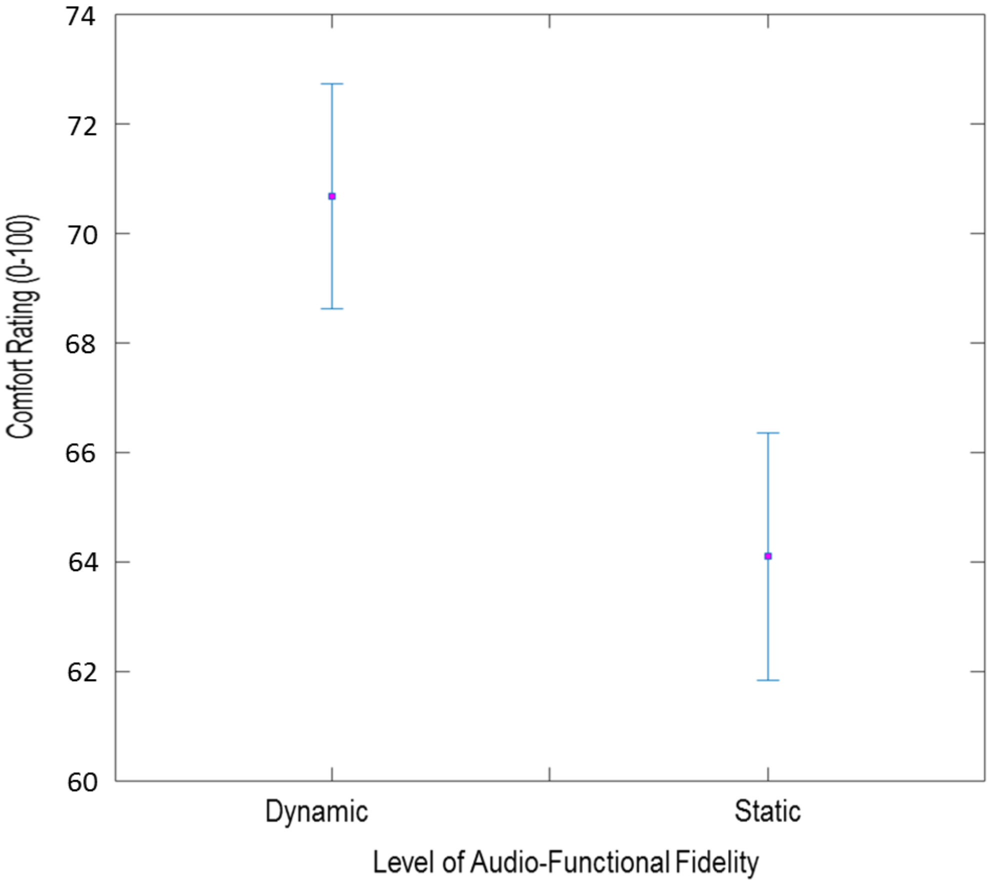


Figure 36: Effect of audio-functional fidelity on comfort ratings. The dynamic condition showed a significantly higher comfort ratings than static condition. Error bars represent standard error, p>0.001.

4.4.1.4 Annoyance

The impact of fidelity on annoyance ratings also showed a significant difference. Annoyance ratings were significantly higher in the static condition than those in the dynamic (t=4.5517, p<0.0001, df=16) (Fig 38)

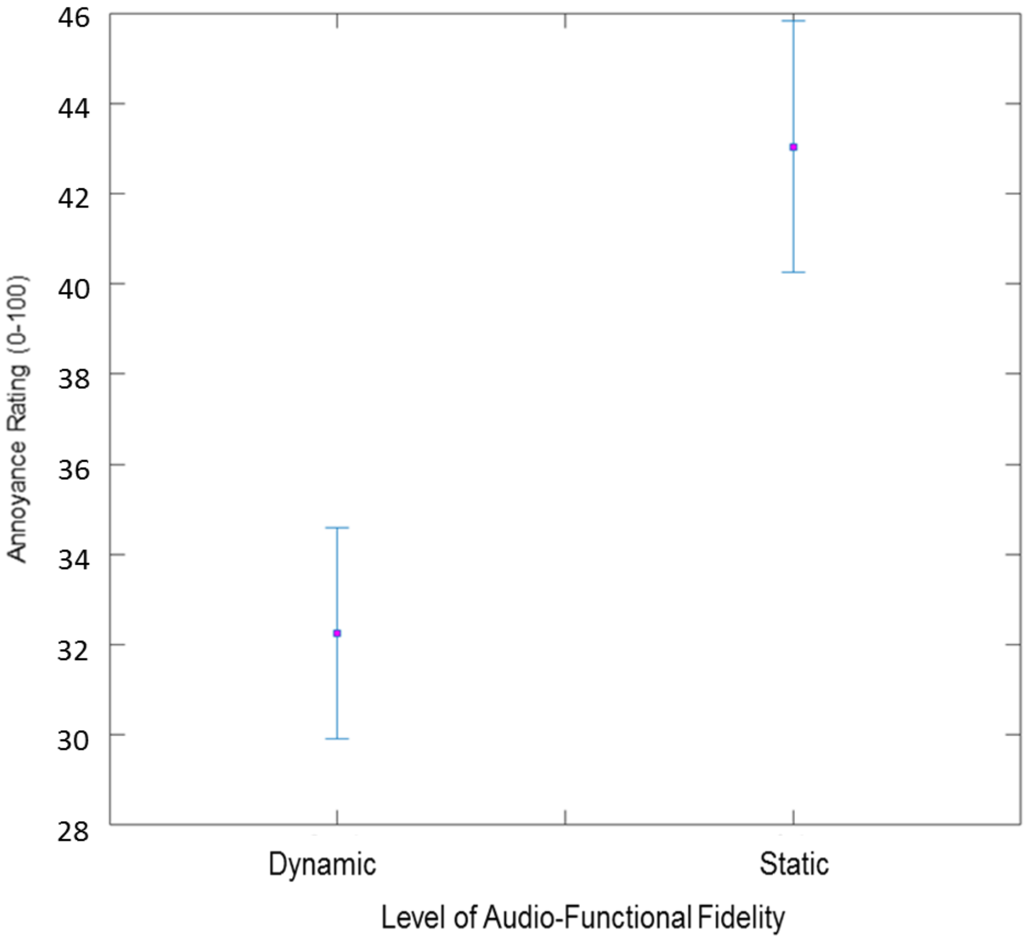


Figure 37: Effect of audio-functional fidelity on annoyance ratings. For the dynamic fidelity condition there were significantly lower annoyance ratings than in the static condition. Error bars represent standard error. p>0.001)

4.4.1.5 Summary of Audio-Functional Fidelity

During the analysis of audio-functional fidelity it was shown that while there was no effect seen in relation to task performance there was a difference shown in both presence ratings and comfort ratings. These differences were not only statistically significant but also followed the trend seen in the environmental fidelity conditions showing that in higher fidelity levels (in this case the dynamic audio condition) there were increased levels of presence and comfort.

4.5 Comfort and Presence Correlation

A linear correlation was carried out on the data between comfort ratings and presence ratings. This linear correlation showed that there was a weak but significant correlation between presence and comfort ratings $r=0.36$, $p=0.0012$. This correlation showed that as presence ratings increased so did participants' comfort ratings.

This correlation further supports the statement that with a higher level of fidelity experienced by the participant there is a higher level of comfort reported by the participants, solidifying the link between comfort and fidelity.

4.6 Discussion

The study assessed the impact of simulation fidelity upon comfort ratings and task performance across changes in both environmental and audio-functional fidelity. The results show that perceived simulation fidelity is affected by both the environment in which the simulation is displayed, and the accuracy of sound cues to the simulated events. These changes in simulation fidelity effected task performance particularly when environmental fidelity was manipulated. Participant comfort was shown to be negatively affected by a decrease in simulator fidelity. Overall the data reflected that when a decrease in simulation fidelity was reported by the participants there was a similar decrease in participants' task related performance and their comfort ratings.

The negative effect of lowered simulation fidelity on task performance is already noted within the literature (Meyer, 2012) and this study supports the link between the two. The addition of changes in presence and comfort being reported by participants and following similar trends illustrates a further link. This relationship between comfort and presence can allow for the interpretation that simulation fidelity may have a significant impact upon participants' perception and interpretation of sound cues. This change in interpretation can be attributed to both the environment in which the cues were received and the method in which they were presented (environmental and audio-functional fidelity respectively), not just the physical qualities of the cues. This effect is attributed to the previously mentioned concepts of PI and Psi by Slater (2009), which explains realistic responses to virtual reality through the sensation of presence (PI) and the realistic reactions of the simulation (Psi), when these two illusions are enacted Slater claims that participants will react to a simulated environment and events realistically. In the case of this data it can be seen that with higher fidelity simulations (confirmed by higher presence ratings) comes higher comfort ratings. This is potentially due to participants expecting the sound cues in higher fidelity settings. This caused the participants to treat them as an inherent part of the simulation and so reacting to them realistically leading to less task disruption resulting in higher comfort ratings.

In the case of this data it is important to note that with the audio-functional fidelity variable the dynamic audio setting was found to be the highest fidelity, as expected when compared to the static audio setting. This condition provided more information to the participants on the task at hand indicating acceleration and direction. All of this would come well under the umbrella of Psi. However, there was no significant change in the task performance, indicating that the participants, while aware of change to the cue and representing the increased level of fidelity with increased presence ratings, were unable to use this information to become better at the task without training (Meyer, 2012), though there was still a statistically significant increase in comfort ratings between the static and dynamic conditions. This confirms that this effect on comfort was not due to an increase in ease of task with addition of helpful cues but was instead due to the action of PI and Psi.

The set-up and development of this study provided insight into the efficacy of fidelity as a force which can impact comfort. This insight was used in the development of a large-scale study (section 5.5.1). The development primarily led to the changes made to the HELIFLIGHT-R simulator reported later.

The novelty of this chapter is shown in the assessment of different types of fidelity, both the environmental and audio-functional. This assessment provided a deeper view into the impact on comfort of specific methods of manipulating fidelity than had been done previously. In addition to this novelty the data produced here is used to show a novel test of aspects of Slater's Plausibility and Place Illusions. This data, when applied to the model, provided information which confirmed the link between annoyance, performance, and comfort. This link becomes one of the cornerstones of the novel model of audio comfort.

To conclude, this chapter has demonstrated that higher simulation fidelity leads to an increase in comfort due to the effects of both the Plausibility and Place Illusions. This chapter also demonstrates that higher simulation fidelity does not necessarily need to be cue specific to enact an increase in comfort for an audio modality, instead the entire sensory experience is used in experiencing the illusions.

Chapter 5: A New Model of Audio Comfort

This chapter revisits literature relevant to the models of audio comfort to re-enforce the information provided earlier. The review then goes on to discuss additional new literature relevant to the model. Following the literature this chapter goes on to detail the proposed model of audio comfort, providing an explanation of the theoretical grounding of the model along with its assumptions. With the model’s assumptions laid out, data from Chapter 4 is examined to test the model’s validity. From here a new large-scale study is introduced with the aim of determining the accuracy of the proposed model’s assumptions.

5.1 Comfort

Comfort is often seen as a subjective aspect of the human experience (Woszcyk, Bech and Hansen, 1995). However, high intensity noise (Nor et al., 2008) and vibrations are universally considered as uncomfortable experiences to the extent in which, at high levels, they can cause pain (De Looze, Kuijt-Evers, and Van Dieen, 2003). The elements of a stimulus causing universal discomfort are varied in aspect. Two such universal discomforts have been identified as high intensity noises and high intensity vibrations (Dempsey and Leatherwood, 1976, Bhiwarpurkar, Saran, Goel, Mansfield, and Berg, 2009). Furthermore, the claims from Canadas et al. (2005) imply that it may be possible to split these further, allowing the identification of aspects within these universal discomfort variables. With sound, this could potentially include pitch or other spectral differences. Supporting this claim McCrudy (1986) showed that the sound spectrum, in particular high frequency components, make a strong contribution to discomfort. Spectral differences in the aircraft environment have also been shown to increase discomfort more than their level alone would predict (Leatherwood, Clevenson, Dempsey, 1978), which is in-line with more general models of sound annoyance. The impact that different sound spectra have on comfort has also been displayed in research presented in Chapter 3 of this thesis. Dempsey and Leatherwood (1976), for example modelled aircraft passenger discomfort on the basis of sound and vibration levels and showed that the two aspects show an interaction effect and together are often considered related factors. These findings are corroborated by studies such as Khan and Sunderström (2006), who concluded that, as with noise intensity, vibration has a significant effect upon passengers' comfort, annoyance, and ability to carry out tasks. Mellert et al. (2008) also presented data showing significant effects of vibrations and noise intensity as two major factors in passenger comfort and staff performance.

5.2 Rolls' Theory of Emotion

The cumulation of this research and the literature described in this chapter led to the production of a theory of comfort based on Rolls' (1990) theory of emotion. Rolls' model states that there are 9 functions of emotion (Rolls and Treves 1988).

1. To elicit an autonomic and endocrine response
2. To learn flexible behavioural responses
3. To motivate to action
4. To cause communication expressing an emotional state
5. To promote social bonding
6. To impact the cognitive evaluation of events
7. To store memories more effectively
8. To recollect memories

These functions of emotion are various evolutionary traits which have been developed to promote survival, as well as promote rewards and avoid punishments.

Emotion is described by Rolls as the states caused by stimuli, which are reinforcers. Some stimuli, named primary reinforcers are innate or natural reinforcers, for example, pain caused by vibrations or sound at high intensities. Secondary reinforcers are learnt associations with primary reinforcers, for example money, which can buy food. The process in which these secondary reinforcers are acquired is called stimulus-reinforcement association via classical conditioning (Rolls 1990).

The model also puts forward that emotions are impacted by the expectations of the participant. This is the case where a punishment or reward is unexpectedly either present or omitted. In these cases, the participant's expectations of a situation have been subverted and this causes emotional modulation.

Rolls' theory differentiates between positive and negative reinforcers and denotes their impact represented by the position on a scale between Ecstasy and Terror. The impact of expectations on emotion is separated from this and with the unexpected omission or presentation of positive or negative reinforcers. This leads to a position on a scale between Relief and Rage.

Rolls (1990) represents these interactions on a scale (Fig 39) with Ecstasy/Terror on the Y axis and Relief/Rage on the X axis. This scale shows that with more intensity of the emotion, and therefore the effect of the reinforcers (their presence or their removal), the larger the distance from the 0 point.

The Y axis describes the presentation of either positive (increased value of the Y axis) or negative (negative value of the Y axis) reinforcers, while the X axis shows the termination of a positive reinforcer and/or the omission of a positive reinforcer, which leads to negative values on the X axis. Termination of a negative reinforcer and/or omission of a negative reinforcer leads to positive values on the X axis.

Rolls explains that, while these effects are consistent, the impact of possible behaviours available to the participant due to the environment (i.e. if the participant has agency to act on either the presentation or omission of a reinforcer) can lead to an alternate emotional response. This response would be altered if the participant lacked agency and instead could only passively experience the reinforcer or lack thereof. Rolls uses the example of a response of anger to the omission of a positive reinforcer while the participant has agency. He states this would instead become more akin to sadness when the participant is only able to passively experience the situation.

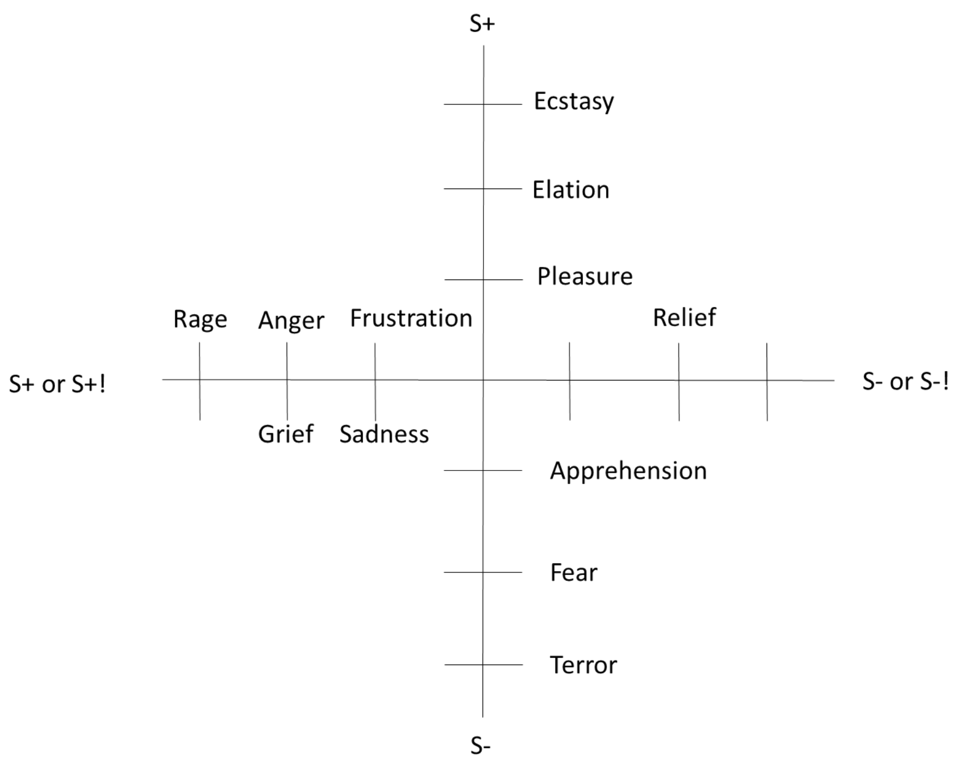


Figure 38: Rolls (1990) Theory of Emotion axis

5.3 A New Model of Audio Comfort

The key aspect of Rolls' model of emotion is that he considers the primary function of emotion to be to 'motivate behaviour'. The further the emotional state is displaced from the 'origin' the stronger the emotion, but also the requirement to 'take action'. The idea underlying the model of comfort proposed here is that comfort is a state that does not motivate action, in other words that comfort and emotion are opposites. The model represents deviations from comfort in the same coordinate system as Rolls' model (Fig 40).

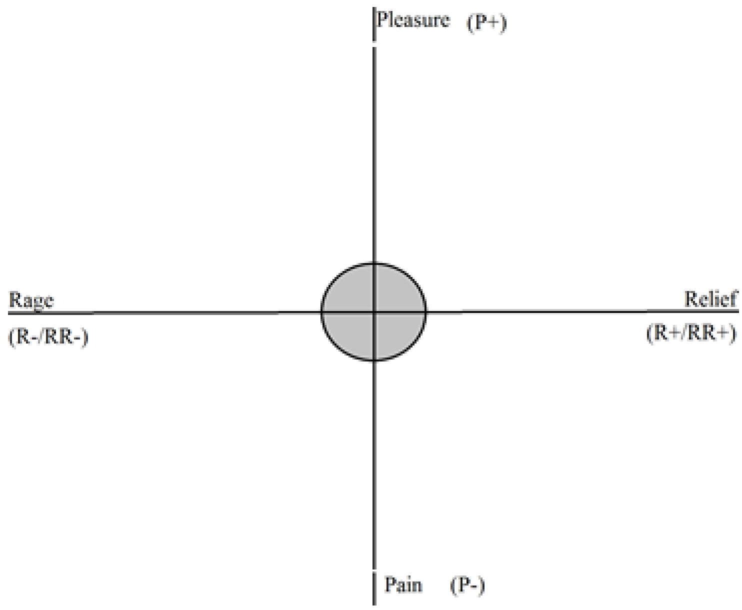


Figure 39: Representation of model of audio comfort axis with centre comfort space

In its most basic terms, the model states that when a participant experiences an uncomfortable stimulus (such as high intensity sounds) there is movement on the Y axis towards pain. When this uncomfortable stimulus is reduced or removed there is movement on the Y axis towards pleasure. This movement on the Y axis is characterised by changes in task performance as uncomfortable stimuli (such as high intensity sounds) have been shown to impact upon task performance (Clark and Sörqvist, 2012, Khan et al., 2004, Iachini et al., 2012, and Mellert et al., 2008).

When a participant experiences the unexpected removal/introduction of a helpful (or unhelpful) stimulus, such as the removal or provision of a stimulus that provides task cues or the introduction or removal of a stimuli that masks task cues, there is movement on the X axis (i.e. towards rage or relief). This movement on the X axis is characterised by changes in reported annoyance. Essentially, performance drives the Y axis and the expectations of the participant drive the X axis.

This model posits that movement on the Y axis is caused by a participant experiencing stimuli which would normally be considered to cause discomfort e.g. high intensity noises, vibrations etc. The amount of movement in this model is increased with the intensity of the stimuli, for example a higher intensity sound (dB) leading to larger movements on the Y axis moving towards pain (Fig 41). This movement is noted as P-.

Movement in the alternate direction (towards pleasure) due to the introduction of pleasurable stimuli would be noted as P+. As this model is a model assessing comfort, rather than pain or pleasure, these movements should head towards these points but not reach them. Instead they should move away from the zero point (0,0) on the axis and if the discomfort caused is strong enough leave the ‘comfort space’ it would indicate that the stimulus applied is causing discomfort.

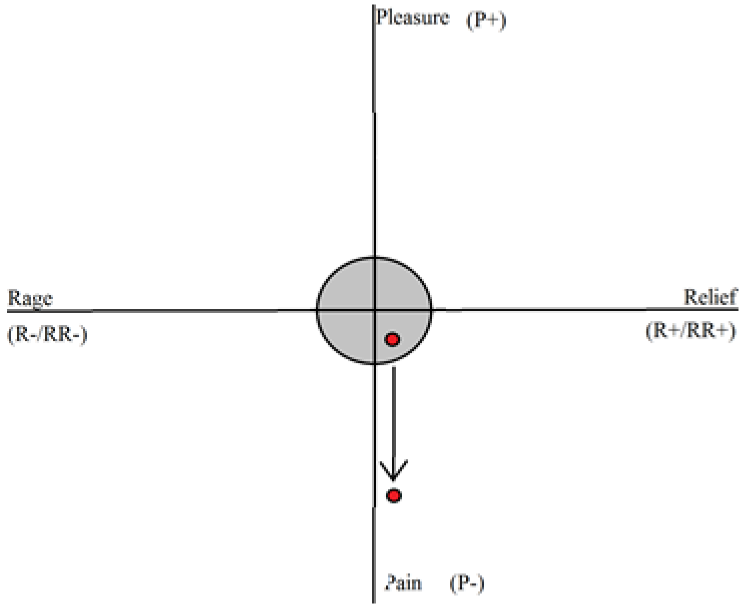


Figure 40: Example of audio comfort model Y axis movement due to experience of uncomfortable stimulus.

Movement on the X axis is reliant on the participant's experience and expectations of the presence of stimulus. With the introduction of a useful stimulus (R+) or the removal of an unhelpful stimulus (RR+) there is movement towards Relief. The introduction of an unhelpful stimulus (R-) or removal of a helpful stimulus (RR-) leads to movement towards rage.

These movements on the axis are based on alterations to what the participant expects from the environments. It is important to note that while introduction of unhelpful stimuli or removal of helpful stimuli will change the participant's perception of their experience of the task, it does not necessarily mean that there will be a subsequent tangible effect on performance. Instead there will be a perceivable impact upon the participants' level of frustration or annoyance with the experience (Fig 42).

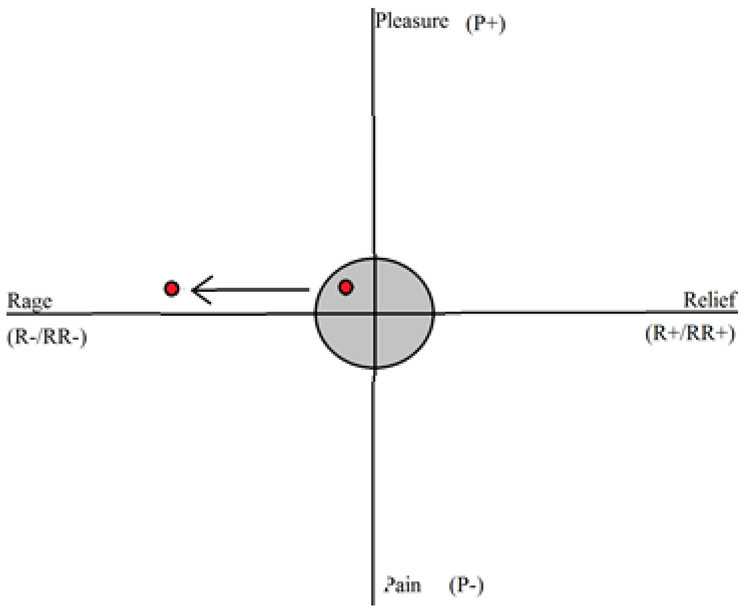


Figure 41: Example of audio comfort model X axis movement due to the removal of a helpful stimulus or inclusion of an unhelpful stimulus.

The concept of the comfort space is based upon the first 4 functions of emotion set out by Rolls et al. (1988). These functions are similar to when a participant is not comfortable and therefore is outside the comfort space.

1. The elicitation of autonomic and endocrine responses, such as an increase in heart rate and adrenaline, is a straightforward indication that a person is not comfortable, as this is part of a physiological response to stress.
2. The requirement of a participant to learn new and flexible behavioural responses to avoid the discomfort they are experiencing or attain a reward they are not currently presented with.
3. Being motivated to take action, in this case taking action to reduce discomfort.
4. Facilitating communication. When an individual is in a situation where they must communicate discomfort.

The comfort space in this context is a novel concept which is a boundary on the model's axis around the zero point. Within this minimal to no discomfort is reported, along with minimal to no annoyance and impact upon performance. Outside of this space discomfort and annoyance are reported and there is a visible impact upon performance. The further outside the comfort space an environment is, the larger the discomfort experienced and reported by the participant.

The movement outside of the comfort space is, as the axis would suggest, two dimensional and therefore the movement can be lateral indicated by annoyance or vertical indicated by an impact on performance. The movement of the environment is not restricted to one or the other and can freely move along both axis with the correct manipulations to the environment. For example: in a situation with audio stimuli with an increase in only the dB level of the noise to uncomfortable levels, vertical movement would be predicted as the primary reinforcer becomes more impactful. If an unhelpful stimulus were to be included, horizontal movement would be expected as a secondary reinforcer becomes salient. If both were to be introduced it would be expected to see both horizontal and vertical movement resulting in a diagonal movement away from the comfort space.

The movement and magnitude of movement along the axes is determined through assessments of the participant. In the case of the Y axis the impact and movement is determined through the participant's performance on a stimuli specific task (for example a SIT while assessing the impact of audio stimuli on comfort levels). The stimuli specific task directly interfered with by the stimuli being assessed provides the model with information on the physical impact of the discomfort stimuli (i.e. how negatively it affects participant's ability to act as they would without the discomfort stimuli (P+/-)). The X axis is assessed differently, with the measure for this being a subjective questionnaire which inquires on the level of annoyance the participant experienced while carrying out the stimuli specific task (R+/- and RR+/-). This question feeds information pertaining to the participant's expectations and experience of the stimuli and environment into the model.

The proposed model provides a novel way of using task performance and annoyance ratings to predict comfort. While the model can use abstract space to define movement along the axes, there is also the opportunity to apply data to these axes. Part of the novelty of this model is that, with the use of data shown below, comparisons can be made between stimuli to determine if there is a difference in their impact upon comfort. This novel application of data can be either examining shifts across a population, or in a single participant over multiple trials. This aspect of the model provides a great amount of versatility in what it can be applied to. Though the proposed model is a progression of the Rolls (1990) model of emotions, there are key differences. Both the addition of a comfort space and statistical analysis methods allowing for the application of real-world data provide a novel aspect to the proposed model. With these additions the model fills a gap in the literature displaying a model which diverges from De Looze et al. (2003) and Metzger (1998). This model instead provides not only a novel explanation of the process in which discomfort is enacted, but also enables the application of research data to this framework. This application of data can be visually represented on the model and analysed through techniques specified in the model's development.

5.4 Testing the Assumptions of the Model of Audio Comfort

Data from the experiment detailed in Chapter 4 of this thesis has been applied to the model by examining the impacts upon comfort, performance, and annoyance. It can be seen in the segments below that the data fits the model showing movement related to both annoyance (X axis movement) and performance (Y axis movement) correspond to a similar change in comfort ratings from the participant. These shifts show that the model provides quantifiable data which will be tested in the following sections.

5.4.1 Examining Shifts on the Model

Examining the data in the context of the model, there are notable shifts when manipulations, such as changes to the fidelity of a simulated environment (Chapter 4), are applied. These manipulations manifest in significant differences in both annoyance ratings and task performance. The model can predict these significant differences. For example, the 21-point increase in annoyance due to reduction in environmental fidelity (Fig 43) can easily be represented on the proposed model as shifts away from the central comfort space along the X axis. In the case of the simulator fidelity study the model predicts a higher level of fidelity leading to a shift towards the comfort space. Due to the use of data from Chapter 4 several relevant figures from the chapter are repeated to re-illustrate the data before the data is applied to the model.

In the case of environmental fidelity, both the annoyance (Fig 43) and performance (Fig 44) metrics follow the same pattern. There are no significant shifts amongst the lower environmental Fidelities, leaving a high level of annoyance and participant performance error. There is only a significant shift when a high level of environmental fidelity is introduced. This can be represented on the model as shown in Figures 45 and 46 by plotting on the model participant error on the Y axis and reported annoyance on the X axis. As fidelity increases there is a shift towards the comfort space due to the reduction in participant error (on the Y axis) and reported annoyance (on the X axis).

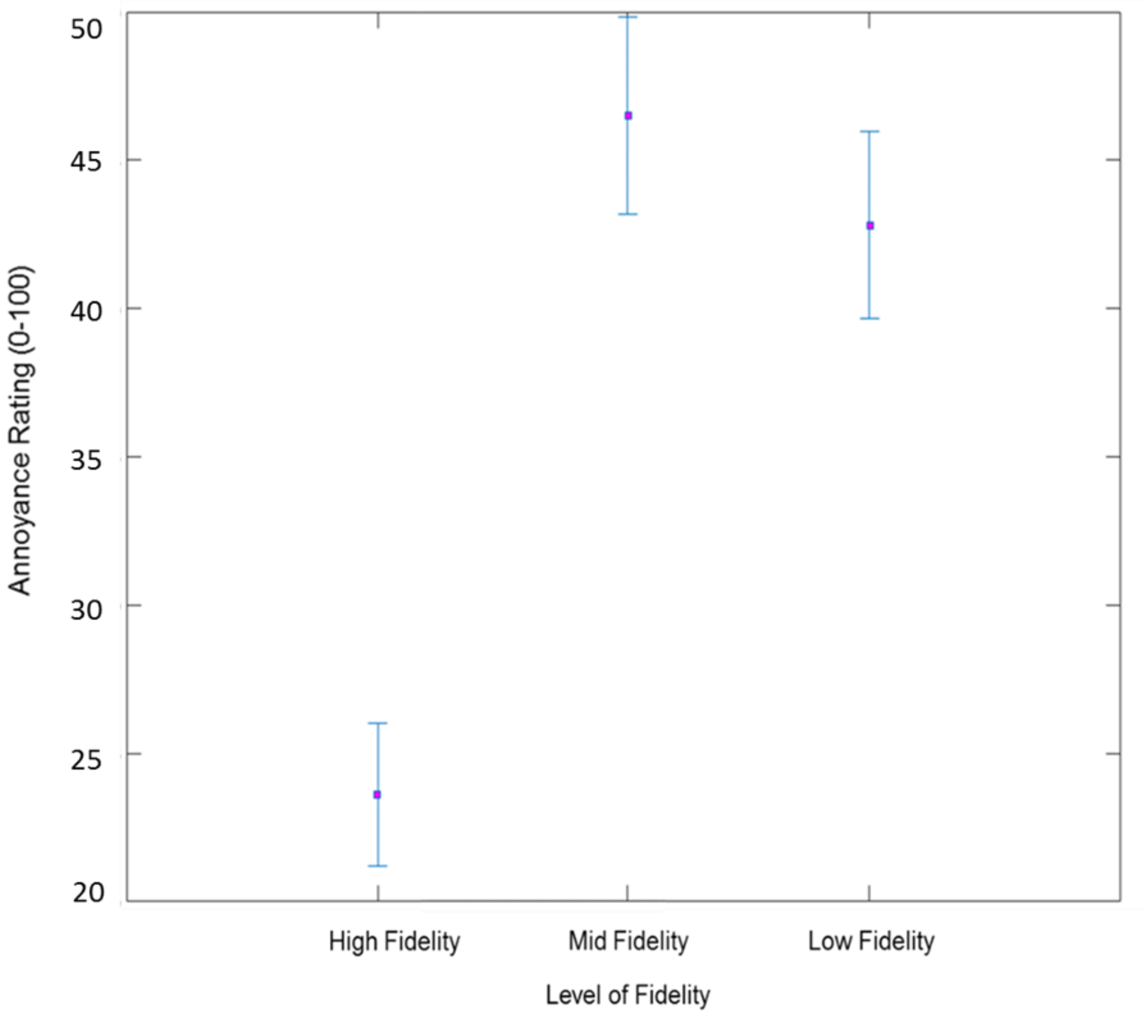


Figure 42: Impact of environmental fidelity on participant annoyance rating. The high fidelity condition showed significant differences for participant annoyance ratings on the questionnaire in comparison to both the mid and low fidelity conditions. Error bars represent standard error. Comparison between high and mid fidelity conditions $p<0.0006$, comparison between high and low fidelity conditions $p<0.0006$. Annoyance for the comparison between mid and low fidelity was not significant.

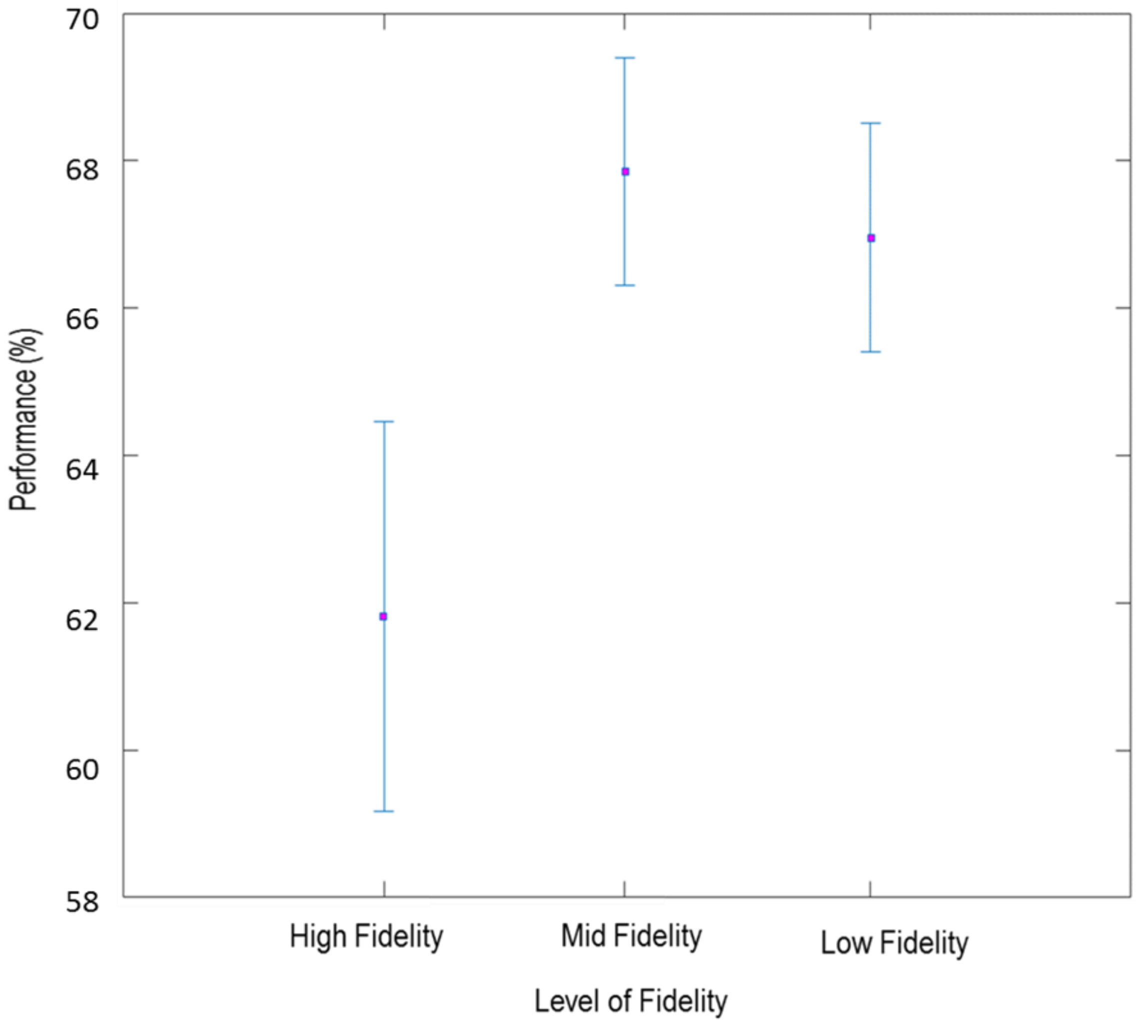


Figure 43: Effect of environmental fidelity on performance. The high fidelity condition showed significant differences in performance on the Tanker Tracking Task when compared to both the mid and low fidelity conditions. Error bars represent standard error. Comparison between high and mid fidelity conditions p=0.0474, comparison between high and low fidelity conditions p=0.0354, performance for the comparison between mid and low fidelity was not significant.

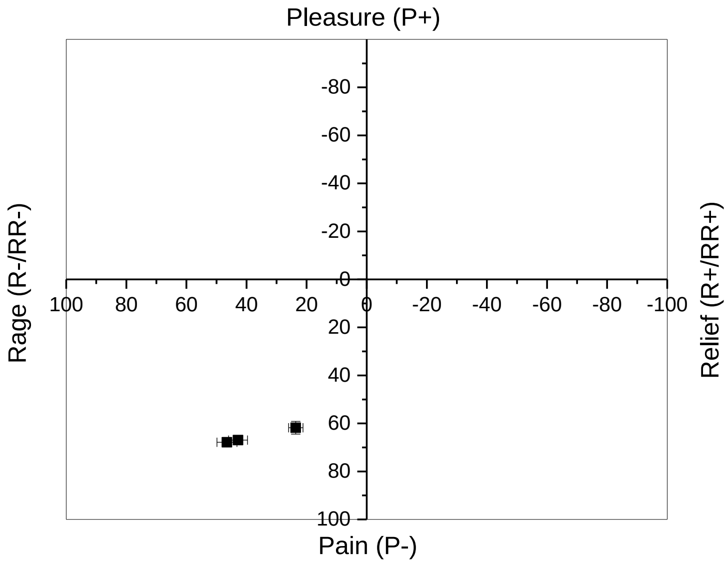


Figure 44: Environmental fidelity data applied to the proposed model of audio comfort. Error bars represent standard error. The data used for the Y axis is task performance. The data for the X axis are participant annoyance ratings.

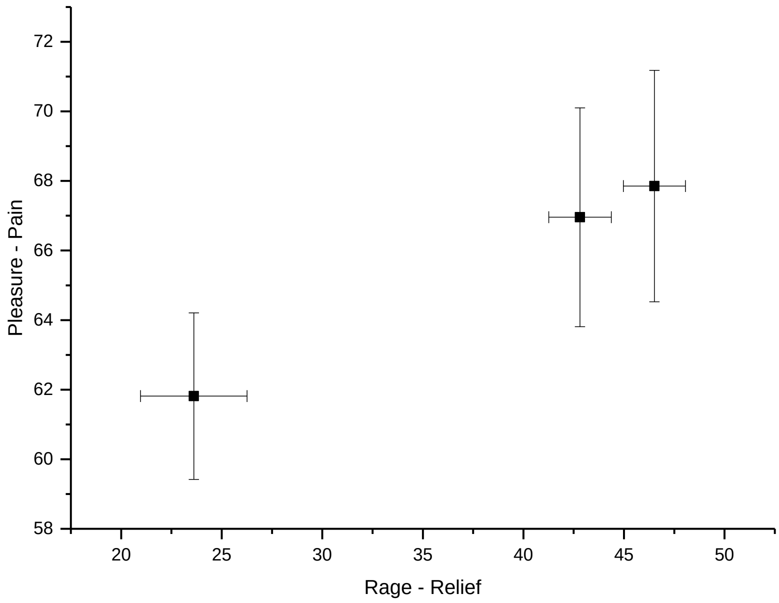


Figure 45: Close in view of environmental fidelity data applied to the proposed model of audio comfort (bottom left quadrant). Error bars represent standard error. The data used for the Y axis is task performance. The data for the X axis are participant annoyance ratings.

The data fits the model's predictions. The data shows that in conditions with low environmental fidelity, high levels of reported annoyance and performance error are displayed. When environmental fidelity is increased, there is a reduction of reported annoyance and performance error. This is in line with the model's predictions with changes to environmental fidelity causing a shift on both the x-axis and y-axis. These manipulations to increase/decrease environmental fidelity decrease/increase the participants' distance from the comfort space.

In the case of audio-functional fidelity there is no significant impact upon the performance error (Fig 47), this is likely due to the need for significant training of participants before they can reliably use high fidelity audio as a cue for improved performance (Wong, Meyer, White, Perfect, 2013). However, there is an impact upon participants' annoyance ratings. These ratings show a significant increase in annoyance whilst there is low fidelity audio rather than high fidelity audio (Fig 48). This indicates that a higher fidelity setting can lead to a lowered annoyance level. It is posited by the model that this therefore would also lead to lowered discomfort in the participants (Figs 49 and 50). This is represented on the model by plotting participant error on the Y axis and reported annoyance on the X axis. These changes in annoyance and discomfort, but not in performance levels, confirm the model's assumptions that when a participant's expectations of a stimulus or environment are not met, the participant's tolerance for discomfort becomes lower. The participant is then more likely to report discomfort due to changes in expectations of the environment (which would be represented as annoyance), despite no change to the audio stimulus' physical properties being made (which would be represented as a change in performance).

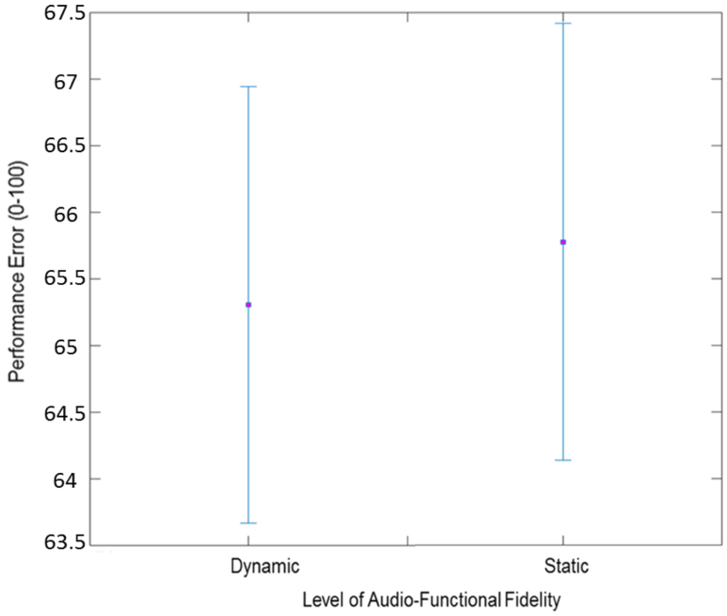


Figure 46: Effect of audio-functional fidelity on error rate (%). The dynamic condition did not show a significant difference in performance ratings when compared to the static condition. Error bars represent standard error.

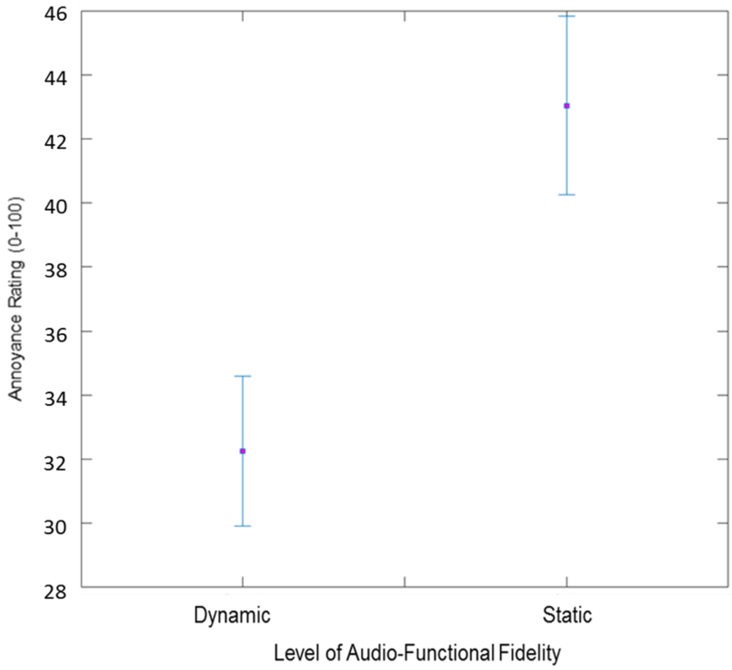


Figure 47: Effect of audio-functional fidelity on participants' annoyance ratings. The dynamic condition showed significantly lowered annoyance ratings in comparison to the static condition ($p<0.001$). Error bars represent standard error.

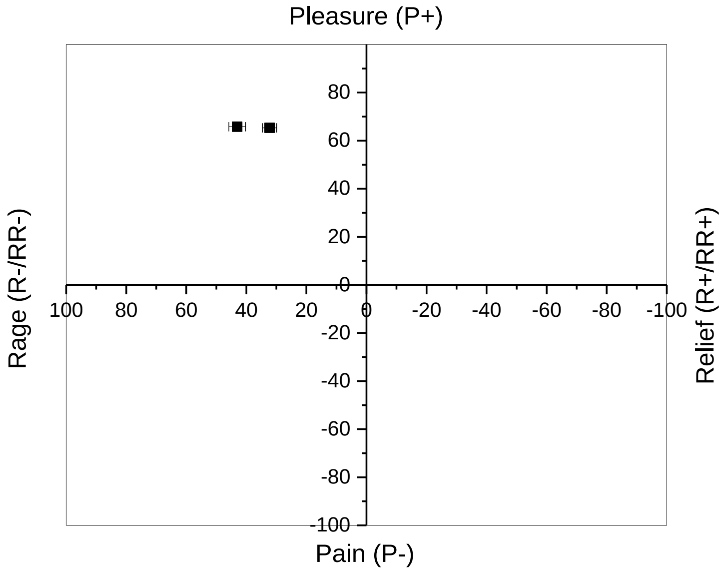


Figure 48: Audio-functional fidelity data applied to the proposed model of audio comfort. Error bars represent standard error. The data used for the Y axis is task performance. The data for the X axis are participant annoyance ratings.

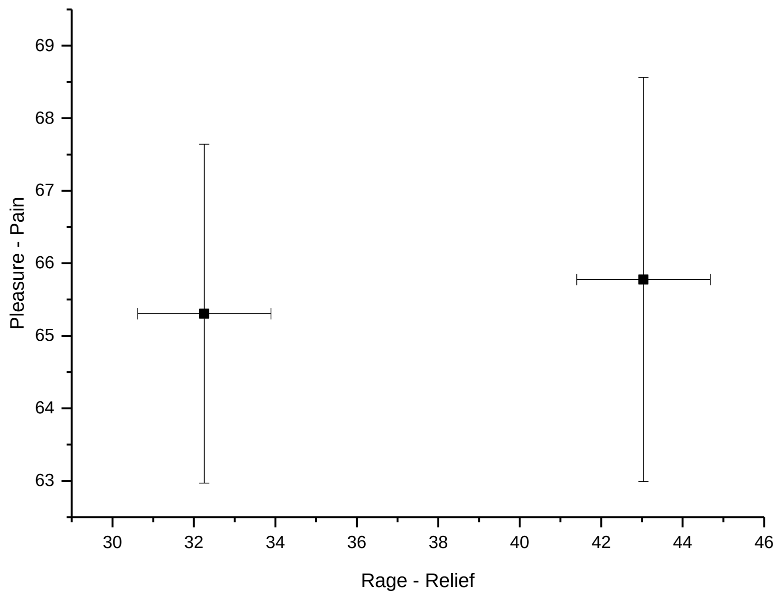


Figure 49: Close in view of audio-functional fidelity data applied to the proposed model of audio comfort (top left quadrant). Error bars represent standard error. The data used for the Y axis is task performance. The data for the X axis are participant annoyance ratings.

5.5 Assessment of Model's Assumptions through Correlations

This segment of the thesis will discuss a large-scale comfort study in which the assumptions of the model are further tested. The justification for this large-scale study was to provide additional data to be used in the analysis of the effectiveness of the assumptions of the model. The data from the fidelity study (section 5.4) displayed the model's capability to represent shifts in participant comfort using annoyance and performance as markers. With this demonstration of the model's effectiveness the assumptions needed to be rigorously tested to prove that they were reliable and consistent. These assumptions are tested through the use of two methods of analysis. With these methods, the model is found to be able to accurately display and predict comfort responses from task performance and annoyance ratings data.

5.5.1 Large Scale Comfort Study

The aim of this study was to use three different sound files (post-transfer function) to assess the assumptions of the model. This data would be used in tandem with the data from the study on the impact of fidelity on audio comfort (see section 5.4). Together these studies provided two sets of data which could be applied to the model to test its assumptions.

The study was a repeated measures design in which participants would carry out both the SIT and comfort questionnaire in each of the three conditions. These conditions were counterbalanced to reduce any fatigue, training or order effects. The independent variable in this study had three levels which were represented by the three different sound files; Base, Staggered, and Unequal (Figures 6 and 9). In this process the participant was exposed to one of the three the sound files while they completed the SIT, they would then complete their digital questionnaire. This process was completed for each of the three sound files which comprised the three conditions.

These sound files were provided by collaborative work between the NLR and UoL's Engineering department, providing the transfer function and CFD respectively. Through this work the sound files were representative of the audio output experienced within the cabin from three different propeller configurations designed for turbo-propeller aircraft. These sound files are distinct from those used in the tonal component experiment in Chapter 3, with application of a transfer function from NLR. To provide a high-fidelity simulation the HELIFLIGHT-R simulator was reconfigured to represent an aircraft cabin (Figs 51 - 55) (section 5.5.1.1). During the study, the sound files were calibrated to 70 dB(A) and the SIT was calibrated to -30 dB(A) SNR. The participants (n=26) for this study were recruited through opportunity sampling and had an average age of 23, these participants had self-reported normal hearing. During the testing six participants were tested per testing scenario and were instructed to stay silent and not interact with the other participants to avoid any cross-participant effects, the participants were closely monitored to ensure that they followed these instructions. The noise for both the turboprop stimuli and the SIT stimuli were played to the participants through a set of closed back Sennheiser headphones.

5.5.1.1 – HELIFLIGHT-R Reconfiguration

To provide a high-fidelity simulation of an aircraft cabin environment to match the turbo-propeller sound files used as stimuli the HELIFLIGHT–R flight simulator was used. HELIFLIGHT-R is a re-configurable flight simulator with an interchangeable cockpit. The projection dome provides a 220 degree (horizontal) by 70 degree (vertical) field of view. It has a 6 degree of freedom motion platform that was not used in this study.

Figure 50: HELIFLIGHT-R (left foreground)

For this study, a new floor arrangement was designed and integrated into the simulator. A new sub-frame was designed and built (Figs 52-54) to provide a platform on which to install the three rows of A320 airline seats (Fig 55).

Figure 51: HELIFLIGHT-R original full cabin

Figure 52: Base frame for new sub-floor for HELIFLIGHT-R

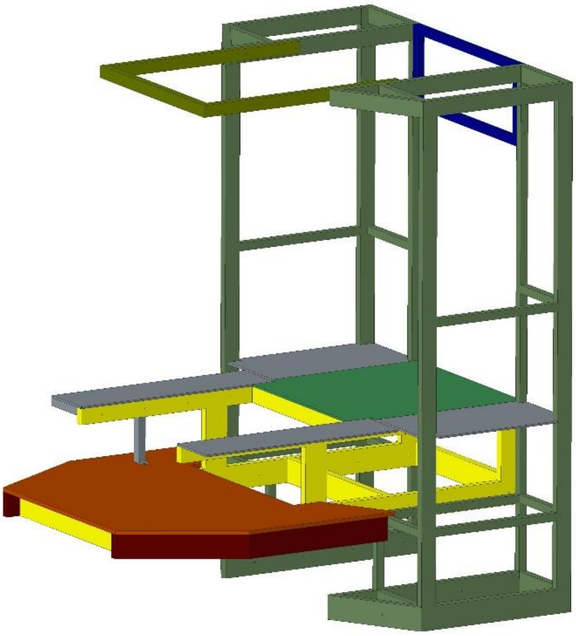


Figure 53: HELIFLIGHT-R basic frame adaptations

Figure 54: Cabin simulation development used in the HELIFLIGHT-R Simulator

A 3-D visual model of an A320 airline cabin was also integrated into the simulator's visual scene and projected onto the internal dome of simulator in front of the airline seats.

5.5.1.2 Application of Large-Scale Comfort Study Data to Model of Audio Comfort.

The data produced from the large-scale comfort study was used to assess the basic assumption of the model. This assumption was that participant comfort can be predicted through task performance and annoyance ratings. The analysis to determine this was a basic linear correlation, this analysis assessed the correlation between annoyance and performance, annoyance and comfort, performance and comfort, and then the correlation between the distance from the origin of plotted points between annoyance and performance against comfort.

The underlying assertion for this analysis was that if comfort correlated more highly with distance, than either of the components separately then their impact as predictors of comfort is stronger together as than separately. This would therefore confirm the model's assumption that both annoyance ratings and task performance together predict comfort.

With this method of analysis several correlations were carried out. These were the correlations between comfort and annoyance, comfort and performance, and then a correlation between comfort and the distance from the origin when annoyance and performance are plotted against each other. These correlations were carried out three times once for each condition. The conditions in this study were three different spectral differences named; Base, Staggered, and Unequal.

With the Base condition when distance was correlated with comfort against annoyance ($r=0.46$, $p=0.0155$) (Fig 56), comfort against performance ($r=0.27$, $p=0.1755$) (Fig 57), and comfort vs distance ($r=0.48$, $n=26$, $p=0.0122$) (Fig 58).

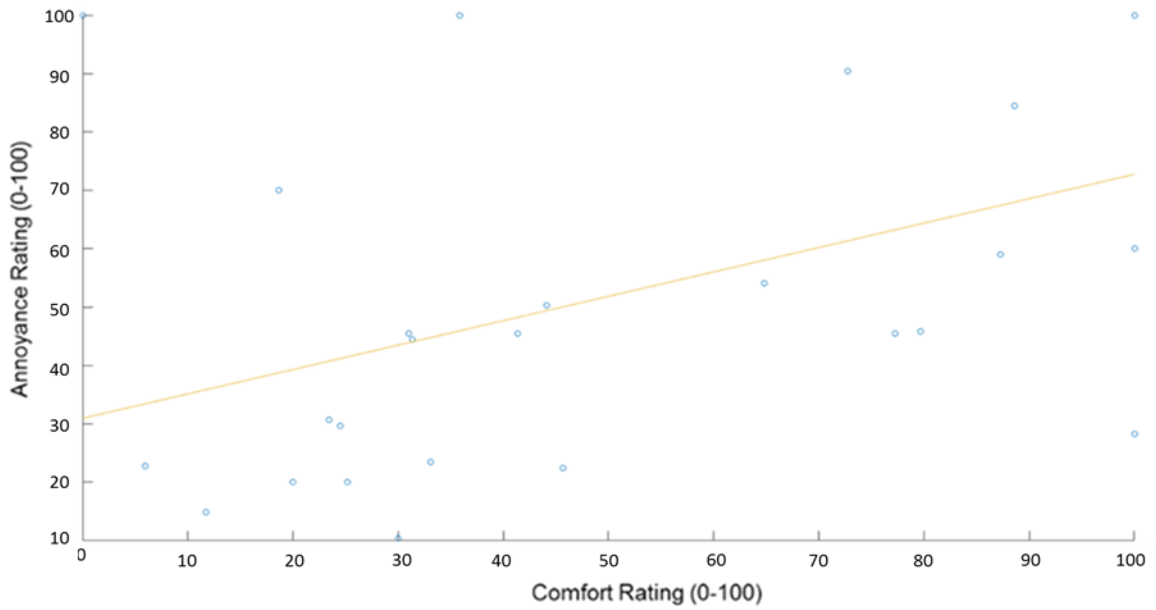


Figure 55: Base audio file, annoyance ratings against comfort ratings. This test showed significant (p=0.0155) correlation(r=0.46).

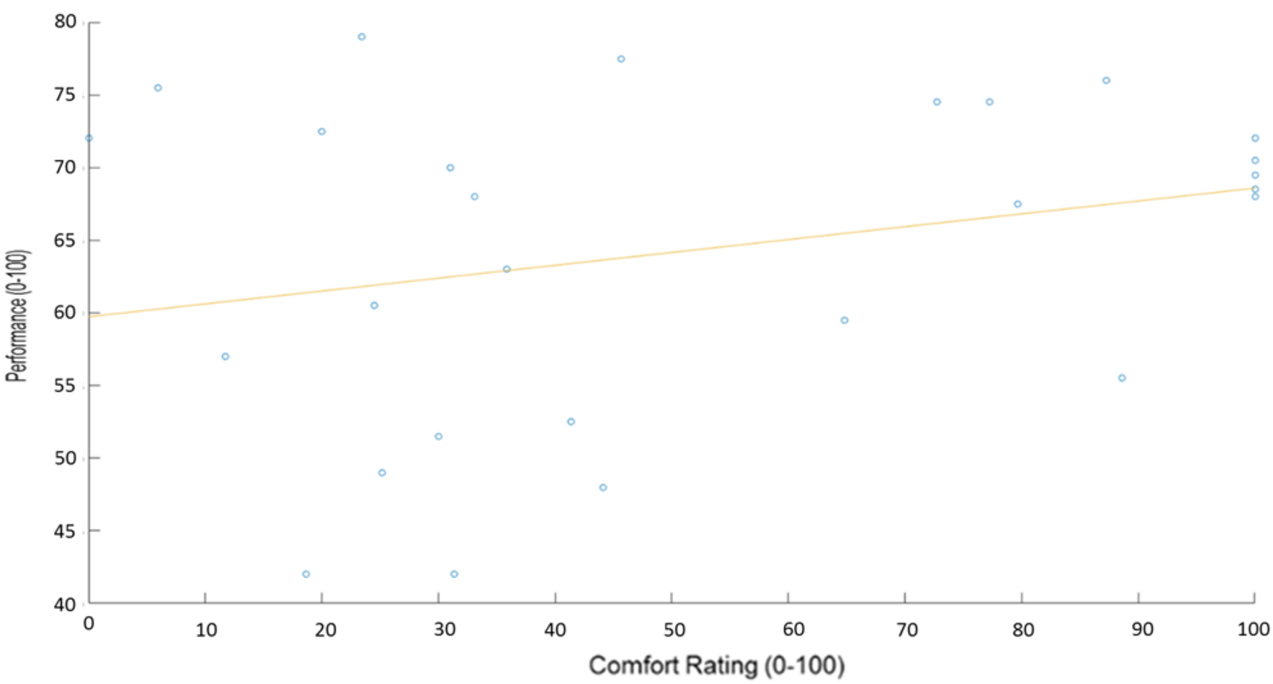


Figure 56: Base audio file, performance against comfort rating. This test showed a non-significant correlation (r=0.27).

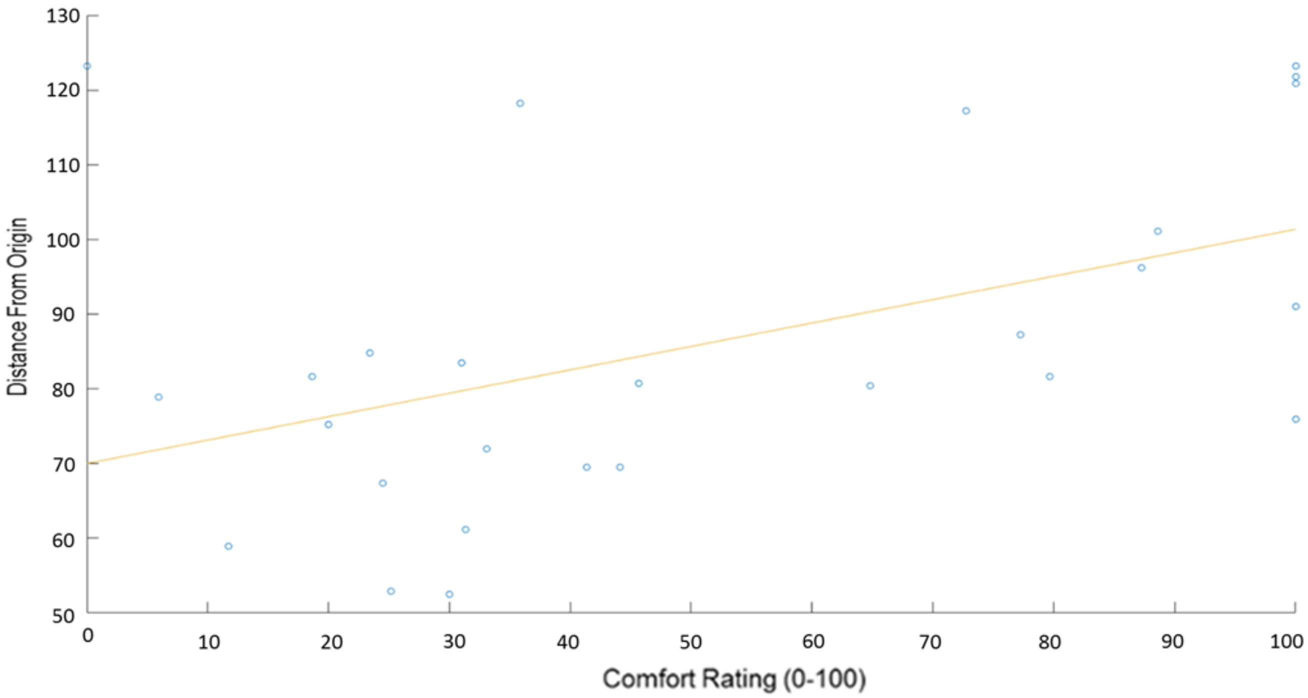


Figure 57: Base audio file, distance from origin against comfort ratings. This test showed that distance from the origin is more strongly correlated (p=0.0122) than previous correlations (r=0.48).

To assess if the difference between these r-values were enough to indicate the usefulness of the model. The highest correlation comparison of annoyance or performance was compared against the r-value of the comfort/distance correlation. To do this the correlation comparison was bootstrapped 10,000 times as was the comfort/distance correlation. Bootstrapping is a process in which the data is randomly sampled a large number of times with the assumption that the participants are ones that reflect the general population, this allows the data to be sampled and re-sampled. With this random re-sampling multiple different data sets can be made from one single data set. With the multiple data sets produced by bootstrapping the original data tests can be carried out numerous times, in this case with the intent of increasing the number of correlations that could be carried out to allow for post-hoc tests that would impossible without the application of bootstrapping. This provided 10,000 r-values for both comparisons these were then compared in a t-test. In the case of the Baseline sound file the results were significant (t=-40.7944, p<0.0001, df=26), the means and standard errors (SE) from the bootstrapping of each sound file can be seen in Table 9.

Table 9: Descriptive statistics for correlation bootstrapping

	Base		Staggered		Unequal	
	Mean	SE	Mean	SE	Mean	SE
Annoyance/Comfort	0.4491	0.0020	0.5351	0.0017	0.2902	0.0021
Performance/Comfort	0.2947	0.0016	0.3763	0.0014	0.3964	0.0015
Distance/Comfort	0.4741	0.0019	0.5936	0.0014	0.4566	0.0018

This significance carried across to the other sound files. Staggered providing correlations of comfort against annoyance (r=0.55, p=0.0031) (Fig 59), comfort against performance (r=0.37, p=0.05) (Fig 60), and comfort vs distance (r=0.61, p<0.0001) (Fig 61). In this case the comparison t-test was carried out between comfort/annoyance and comfort/distance, and produced results of (t=-81.5895, p<0.00001, df=26).

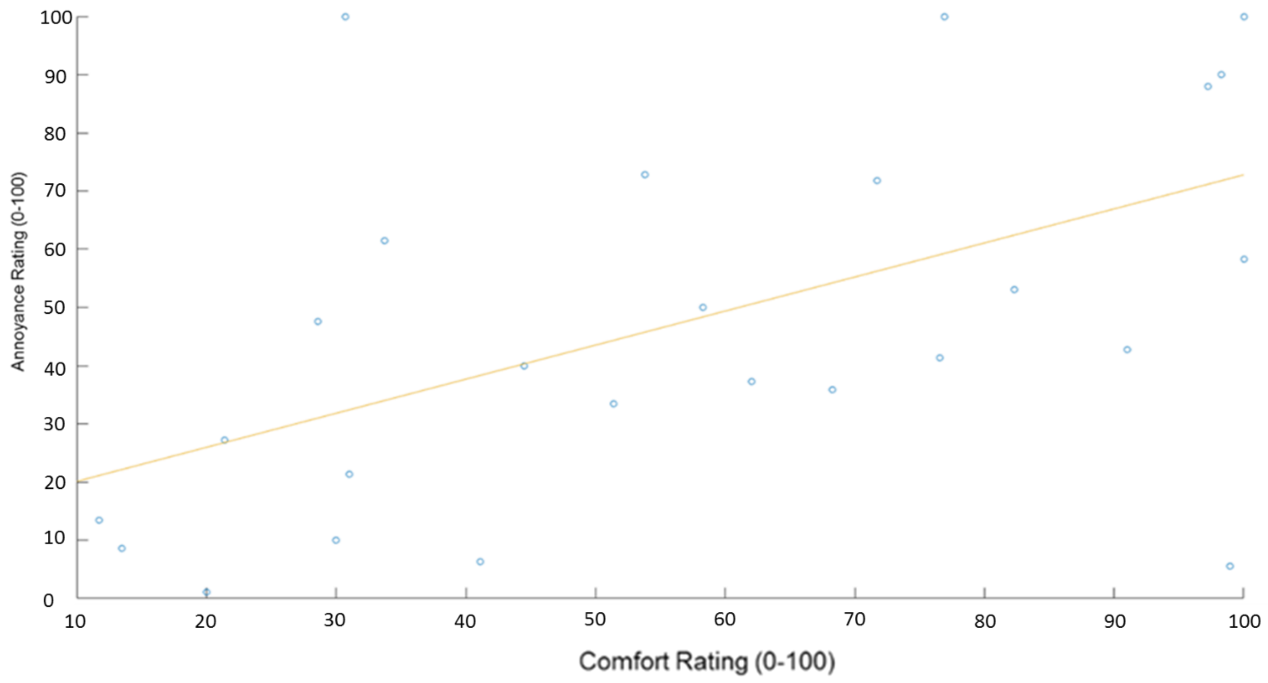


Figure 58: Staggered audio file, annoyance rating against comfort rating. This test showed significant (p=0.0031) correlation(r=0.55).

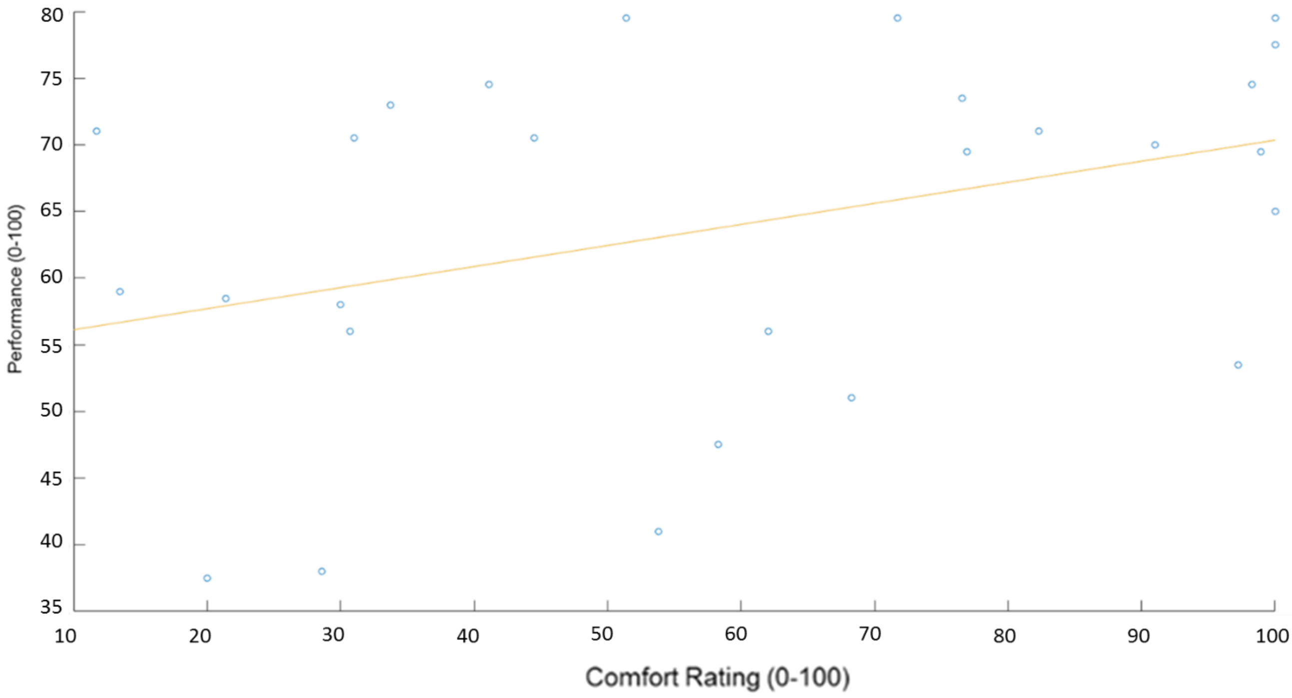


Figure 59: Staggered audio file, performance against comfort rating. This test showed significant (p=0.05) correlation(r=0.38).

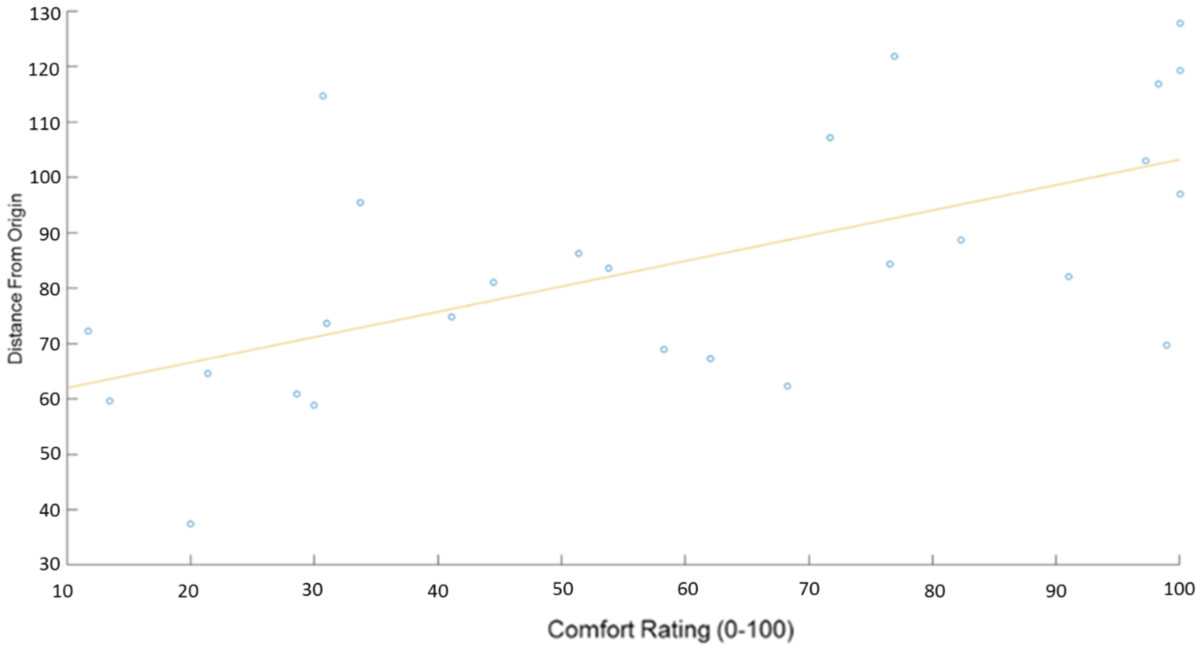


Figure 60: Staggered audio file, distance from origin against comfort rating, This test showed a stronger significant (p<0.0001) correlation(r=0.61).

Unequal providing comfort against annoyance (r=0.29, p=0.1364) (Fig 62), comfort against performance (r=0.40, p=0.0367) (Fig 63), and comfort vs distance (r=0.47, p=0.0132) (Fig 64). In this final case, the comparisons were examined between comfort/performance and comfort/distance showing significant results of (t=-38.0416, p< 0.00001, df=26)

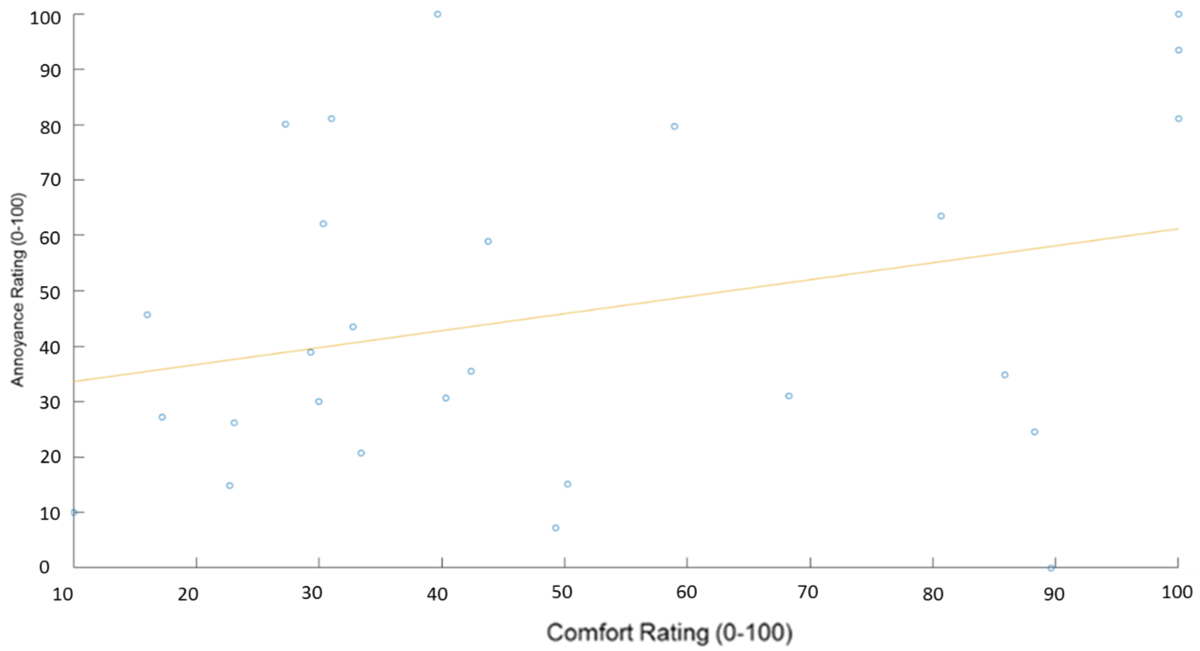


Figure 61: Unequal audio file, annoyance rating against comfort rating. This test showed a non-significant (p=0.1364) correlation(r=0.29).

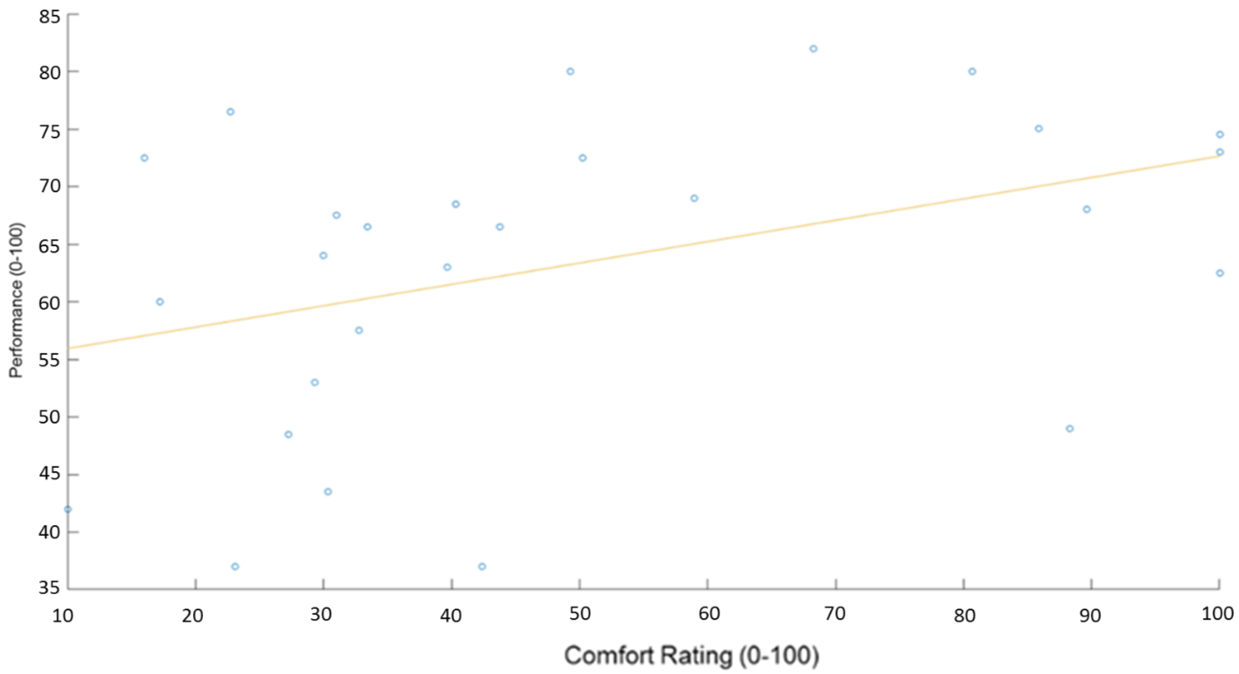


Figure 62: Unequal audio file, performance against comfort rating. This test showed significant (p=0.0367) correlation(r=0.40).

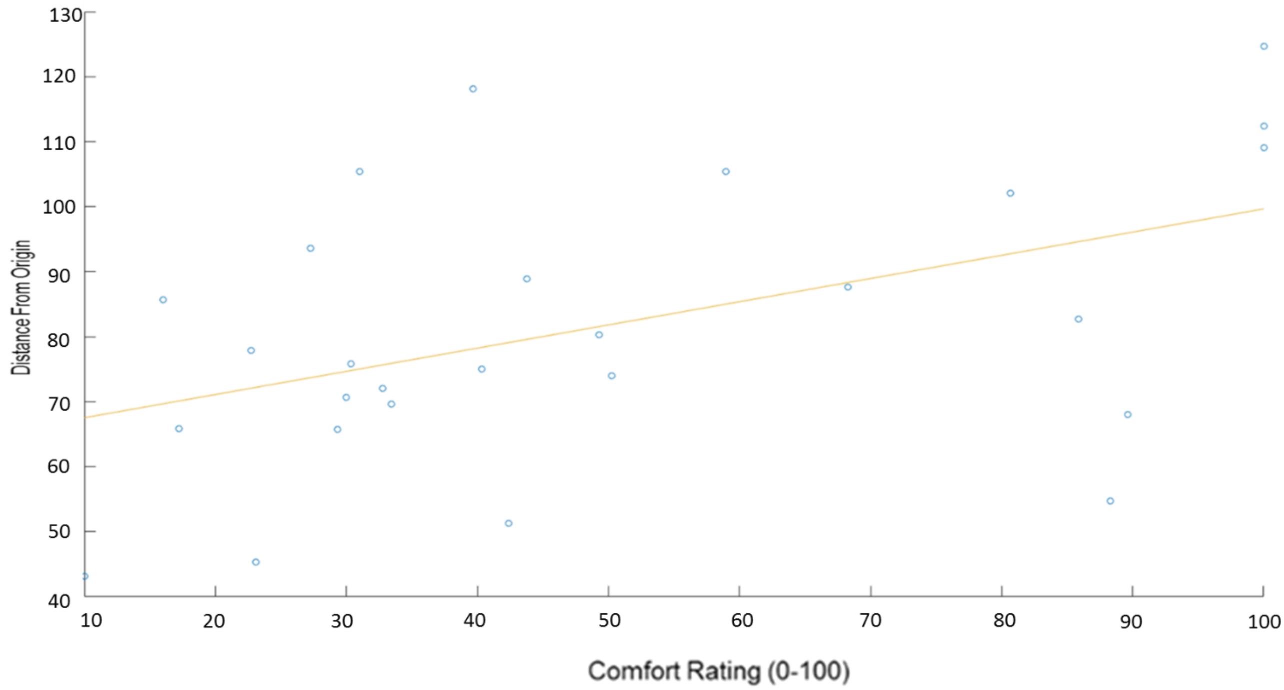


Figure 63: Unequal audio file, distance from origin against comfort rating. This test showed a stronger significant (p=0.0132) correlation(r=0.47).

5.6 Examining Shifts Within Correlations

The analysis for the large-scale comfort study was relatively simplistic due to the straightforward nature of the assumption being assessed. The next question however was a more complex one, how to represent data produced from several within participant conditions to fit to the model in one simple representation? Cooper et al. (2018) explain that correlation analysis of data that contains performance data, therefore has a common and objectively defined scale, and subjective evaluation data, which does not have a common scale or baseline, is difficult. She argues that the correlation of pooled means of the data across conditions is an appropriate analysis. This approach controls for systematic and idiosyncratic differences between participants' internal rating scales. Therefore, this data was analysed using correlated averages in the same manner as Cooper et al. (2018)."

The justification for the use of correlated averages when applying data to the model in the case of representing shifts is that our interest is not at the individual level, assessing each participant's reactions. Instead to make the model as reliable and applicable to any population the model should be able to fit to the "average participant". Therefore, with that in mind the correlated averages method was used which reduces the individual differences to their minimum while still retaining the information caused by a variable change and resulting in a shift in the data.

This method negates the possibility that correlations stem from participant's result correlating with other participant's data, rather than their own providing a false positive. In addition to this the method removes the issue of cross-participant variation which have been noted to occur in these research scenarios (Hood 1968, and Sanchez et al., 2016). These cross-participant variations stem from the idea that while a stimulus may be universally more uncomfortable than another, the internal schema for comfort is likely to be different across participants. Therefore, with each participant comes a new understanding of where the limits of comfort lie. Instead, this method forces the analysis to consider the change in conditions as the sole driver of the shifts in values of annoyance, performance, and comfort and ignores the individual differences produced by participants' own personal comfort schema. This change in the perspective of the data allows the model to provide a more straightforward representation of manipulations made and their impact upon participant comfort.

However, it is usual that this method provides an inflated r-value. This is due to the averaging across participants which reduces the variance in the data by consolidating each condition to a single data-point. This means that the results from this method of analysis, while valid, should be confirmed with analysis on the raw data. This simple additional confirmation would reduce the possibility that the use of correlated averages was creating a correlation where one did not previously exist. Instead these r-values can be used to highlight correlations for within-participant cross condition studies. Therefore, in this section both the raw data and the correlated averages data will be displayed, the correlated averages to provide an insight into how the model functions with correlated data shifts due to experimental manipulations, and the raw data will be shown to provide context.

The next question for the model is while these variables show significant differences and in the correct direction, do they correlate? To answer this question several correlations were carried out, the initial correlations were between participants' comfort ratings and the two other measures annoyance ($r=0.63$, $p<0.0001$) (Fig 65) and performance error ($r=0.33$, $p<0.0001$) (Fig 66). With these tests showing directional correlation in line with the predicted trends the next question was in relation to the model. As the model uses annoyance and performance as predictors of comfort, does the model allow for these predictions? This was ascertained by taking the distance from the zero point (the direct distance from 0,0 to x, y) when annoyance and performance error were plotted against each other and then these numbers were correlated with the comfort ratings ($r=0.66$, $p<0.0001$) (Fig 67). This assessment showed a stronger correlation than either of the previous correlations on their own. The reliability and significance of the difference that this result shows was then tested by bootstrapping the data. This was done by randomly sampling the data to create 10,000 data sets of $n=17$, and then running 10,000 correlations on each of the combinations. These r-values were then compared in a paired sample student t-test which showed the difference between the correlations to be significant: annoyance and comfort against distance and comfort ($t=-128.8807$, $p<0.00001$, $df=16$) performance and comfort against distance and comfort ($t=-495.7952$, $p<0.0001$, $df=16$). This shows that the model can assume that the combination of annoyance and performance to be a stronger predictor than either of the two alone.

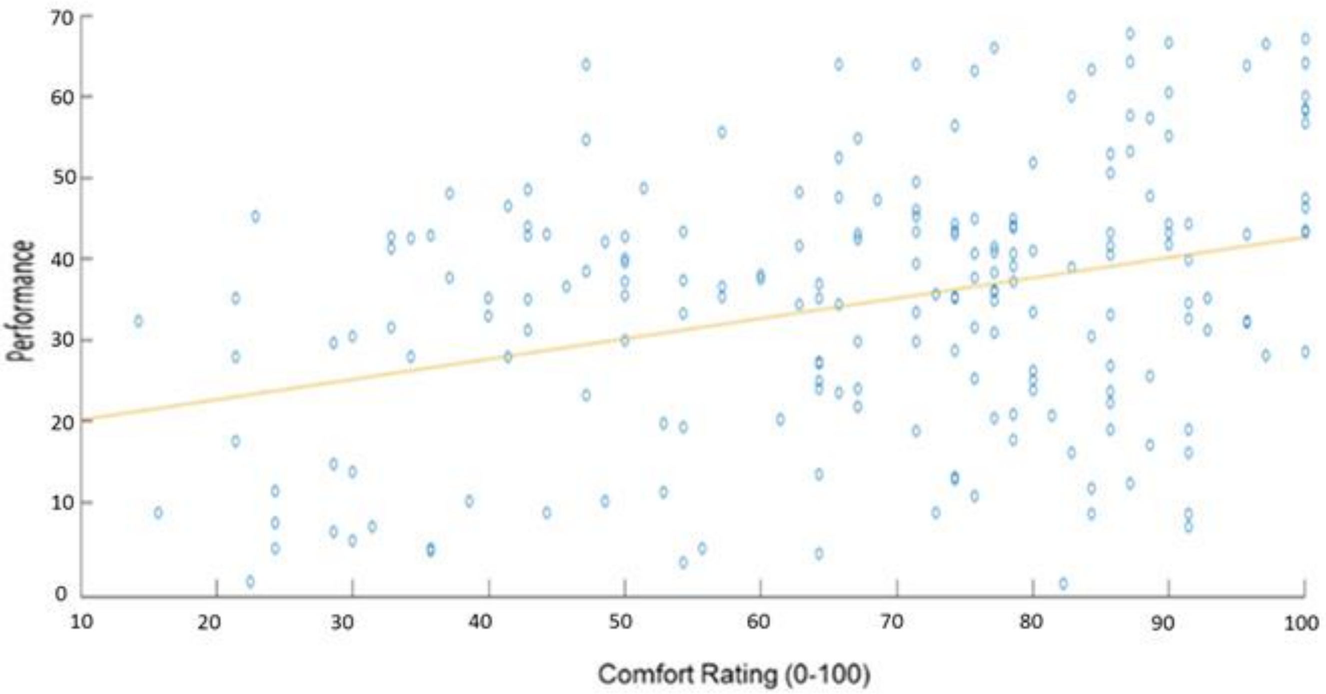


Figure 64: Comfort and performance, r=0.63, p<0.0001

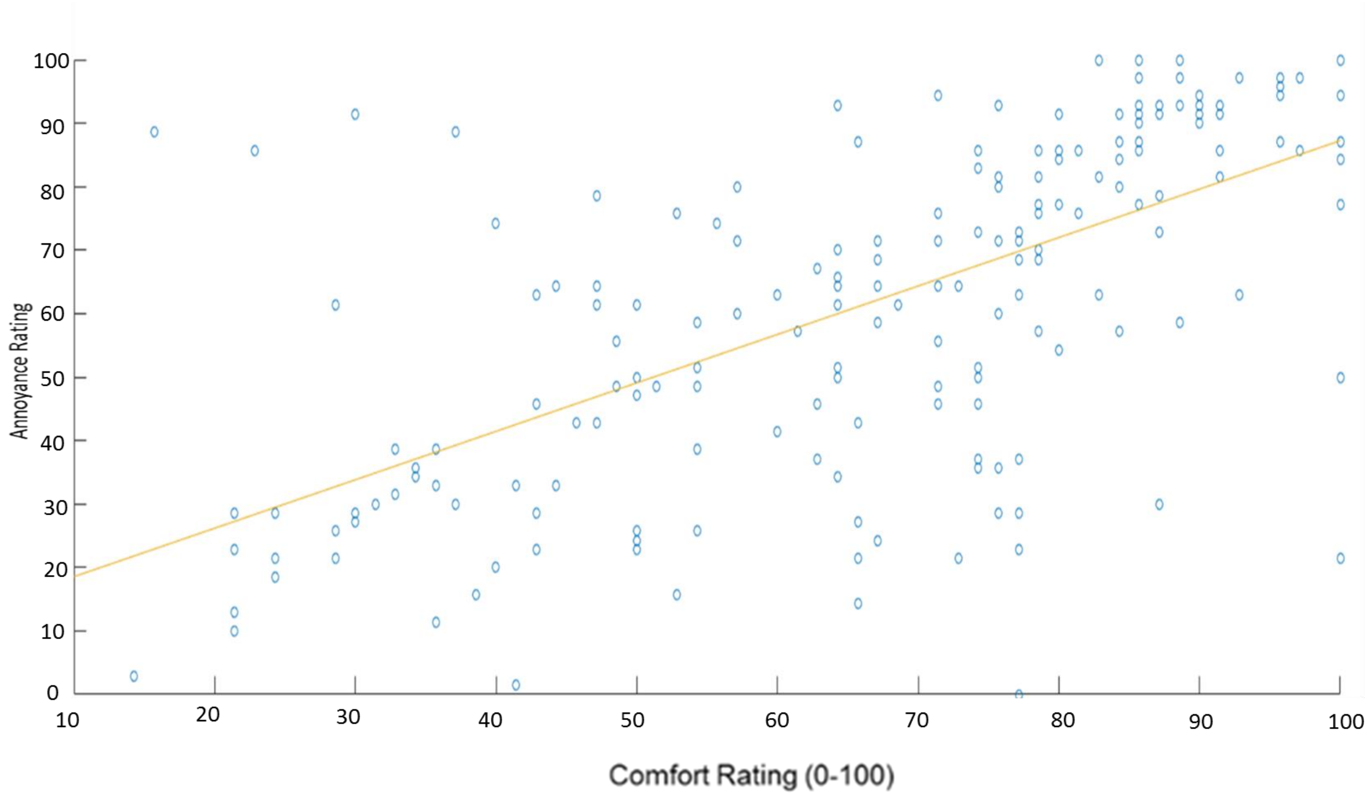


Figure 65: Comfort and annoyance, r=0.33, p<0.0001

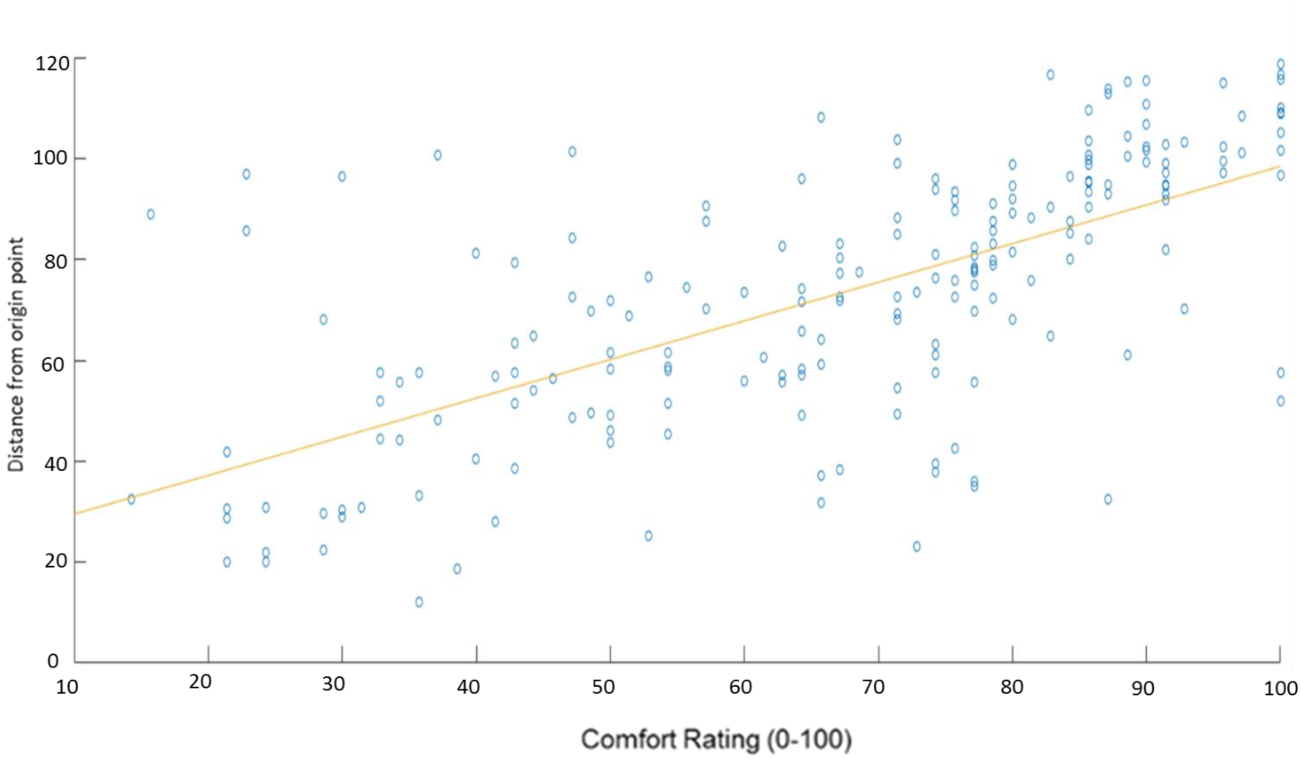


Figure 66: Comfort and distance from the origin, r=0.66, p<0.0001

These correlations have the possibility of correlation across participants while having no correlation within participants. This is why the use of correlated averages was also employed. The correlated averages method of assessment does not have the issue which hampers the previous analysis, and therefore can be used as a useful additional analysis step to clarify the conclusion drawn. The correlations carried out with this data show similar trends with increased r-values, as mentioned previously this is due to the reduction in data noise.

These correlated averages showed relationships between comfort and the other two metrics; annoyance (r=0.9177, p<0.0001) (Fig 68) and performance (r=0.7796, p=0.0028) (Fig 69). Following the method of analysis used in the previous studies the distance from the origin point was also averaged and correlated with comfort showing a similar increase to that of the raw data (r=0.9195, p<0.0001) (Fig 70).

When these correlations are viewed in the context of the model it is clear that the model's assumptions hold true with both methods of analysis. Furthermore, this data displays that these assumptions are held when showing shifts in the data caused by experimental manipulations within participants.

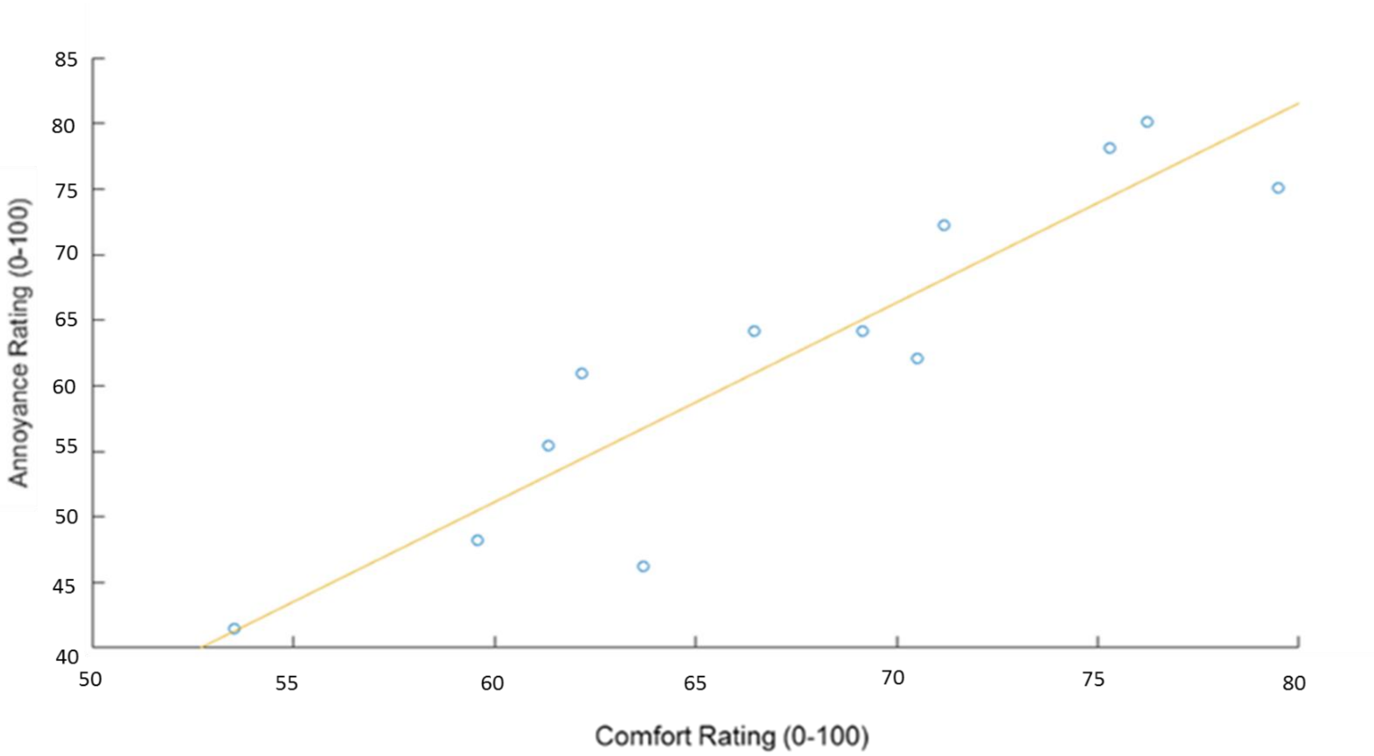


Figure 67: Average data from annoyance rating and comfort rating correlated for each condition. R=0.9177, p<0.0001.

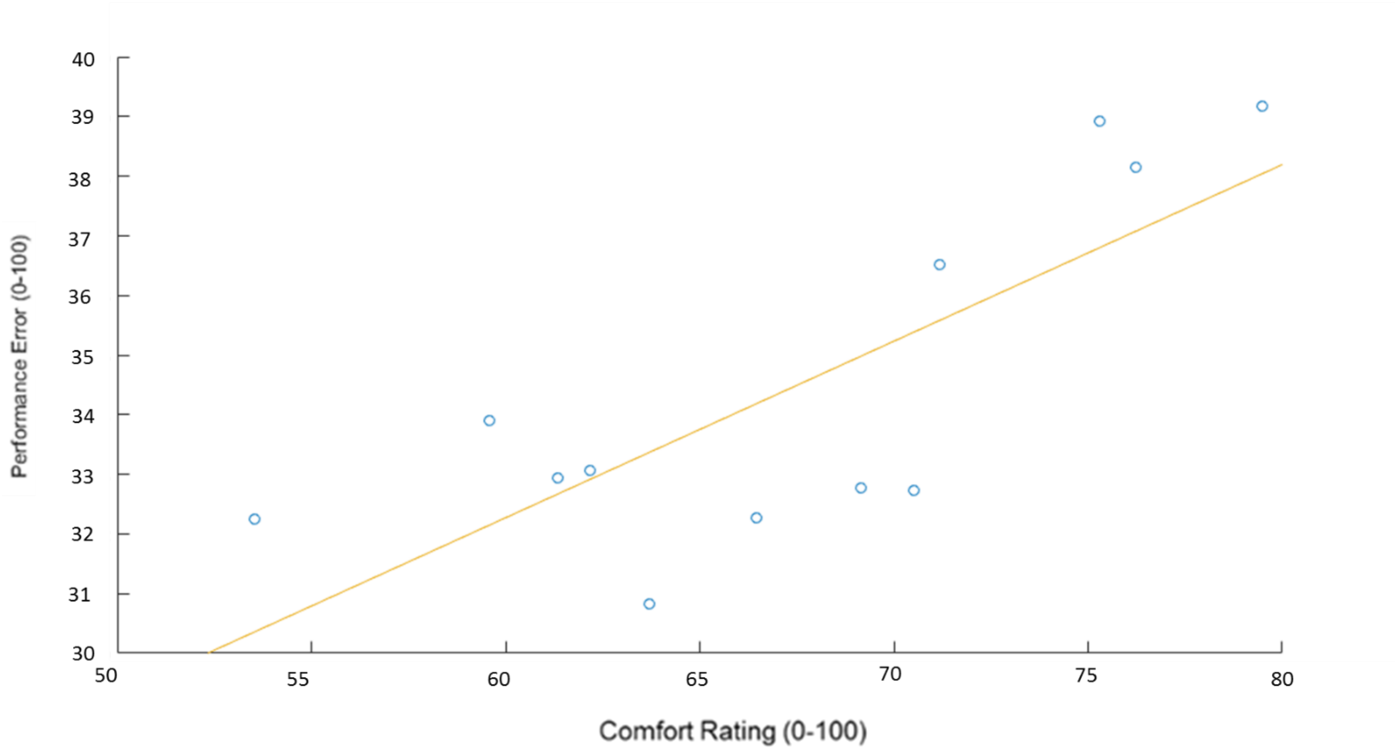


Figure 68: Average data from performance error and comfort rating correlated for each condition. R=0.7796, p=0.0028.

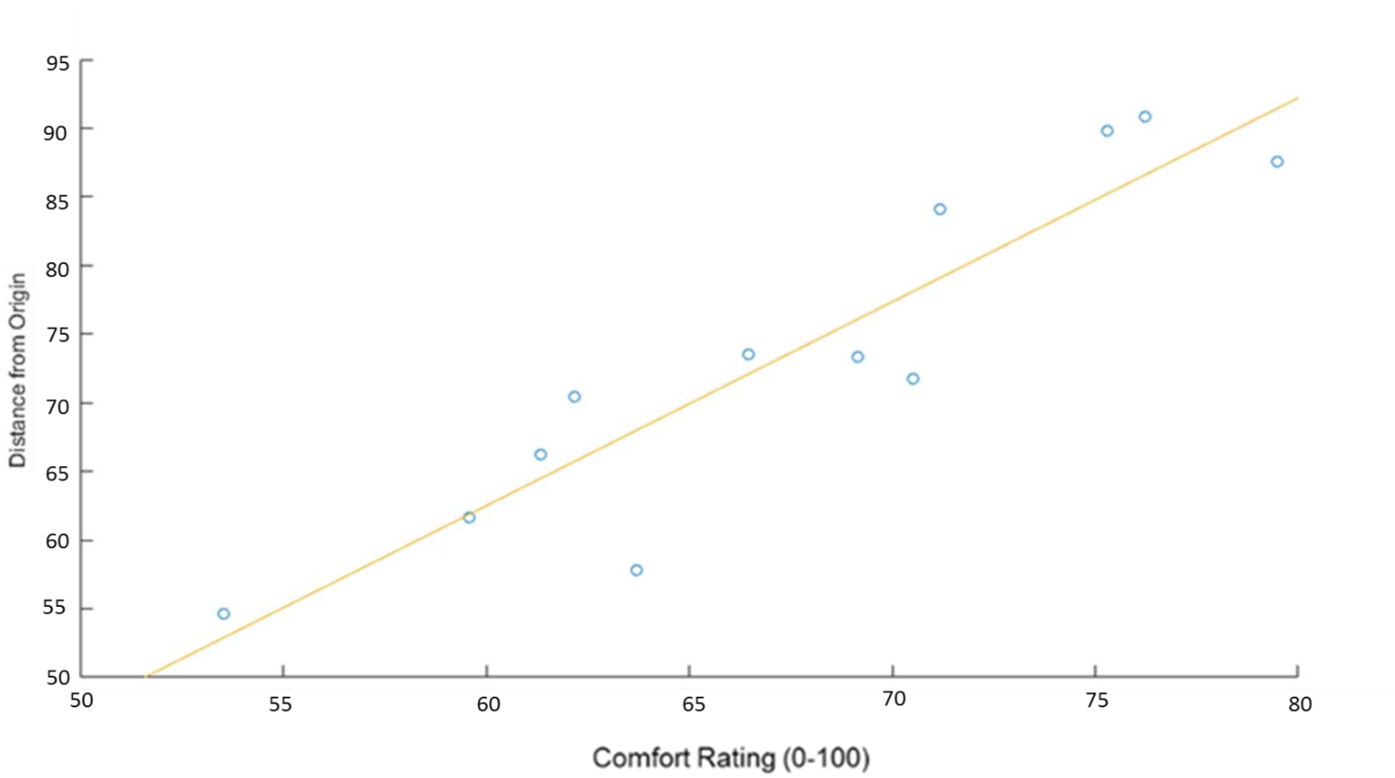


Figure 69: Average data from distance from origin and comfort rating correlated for each condition. R=0.9195, p<0.0001.

5.7 Implementation of this Model

The model of comfort proposed in this thesis, together with the data presented, shows that it can rely on its assumption that together annoyance ratings and task performance are a good predictor of comfort.

This model can be used with one of two methods. The first is to treat the model as participant specific – applying individual participant's data across multiple conditions to the model. This method would allow for a representation of the impact of changes in stimuli upon a single participant. The second method is to average across the conditions, taking an average of all the participants in one condition and comparing them to the average of the next condition. Again, this method allows for the representation of shifts in comfort caused by the change in condition.

The reason for the use of these methods is that, as has been mentioned previously, subjective rating scales are subject to cross-participant variability issues. This, in particular for comfort and annoyance, could well be due to differences in tolerance thresholds and internal schemas of comfort. Therefore, the methods of implementing the model are designed to negate this issue. The first method uses only a single participant at a time, therefore having one consistent tolerance level and internal schema. The second method instead reduces this issue by creating an average schema and tolerance level across all participants by using the averaged data.

With these methods, the proposed model of audio comfort can be implemented in both research and industry. The model has been utilised for the transformed Dowty propeller configurations (Chapter 5), with participants showing no preference for any of the particular designs. Delays in the ImPAcTA project due to issues with wind tunnel testing prevented validation of the numerically generated sound files. Dowty, however, have continued innovation in the development of propellers and have produced a new configuration which, at the time of writing, is currently undergoing CFD evaluations. The model developed in this thesis would be an effective method of assessing this and other new configurations in comparison with previous developments. The model's capability to provide a comparative impact statement for new stimuli, such as the new propeller configuration, makes it a useful tool in an industrial setting. The comparative nature of the model also lends itself well to a research setting in which it can be determined if there is a significant difference between presented stimuli.

Chapter 6: Discussion

In the development of a new model of comfort it was required to develop the model in terms of its assumptions and measurable effects which could then be interpreted and represented on the model. These assumptions and methods of determining and measuring effects were formulated through several studies which provided the empirical groundwork from which the theoretical aspect of the model could be developed.

Chapter 3 of this thesis discusses the groundwork that was carried out for the understanding of the proposed model. This included testing of a battery of metrics used to assess comfort through task performance in a pilot study. Through this pilot study the ToL (χ^2 (1, n=21) = 1.546, p = 0.107) and BTA (t (20)=0.629, p=0.2685) performance metrics were eliminated from the battery. This was due to their lack of sensitivity to the change in audio stimuli, which in turn was concluded to be due to a lack of stimuli specificity using cognitive tasks not impacted by changes in sound. The pilot study also provided methodological changes to increase the sensitivity of the remaining SIT and PEP test, which was done through increasing the number of trials for each task. In the subsequent experiment (Experiment 1) it was found that the increase in trial number did not impact the sensitivity of the PEP ($t(9)=0.182$, $p=0.857$). This again was concluded to be due to the lack of use of stimuli, specificity using cognitive tasks not impacted by auditory changes in the PEP. The increase of trails in the SIT however did produce a rise in sensitivity of the task ($t\ (9) = 2.547$ $p=0.0313$). This allowed it to be used effectively as an objective test to determine the impact of noise intensity on participants. Experiment 1 led to the conclusion that objective metrics which use task performance to measure comfort are required to be stimuli specific. The impact of discomfort while significant are not found to cross into other stimuli-based task performances with enough impact to be reliably measured in an experimental setting. Within this study a comfort questionnaire was also tested. This questionnaire showed it was an effective and reliable method of measuring participant discomfort with consistent effects across participants in the different noise intensity conditions.

In Experiment 2 the SIT and comfort questionnaire were used to assess the impact of spectral differences. The data from this study once again showed the sensitivity of the SIT to auditory discomfort with significant comparisons between each sound file;

Base turboprop design and Unequal turboprop design (df(8), $t=14.0645$, $p=0.003$)

Base turboprop design and Staggered turboprop design (df(8), $t=20.5129$, $p=0.003$)

Staggered turboprop design and Unequal turboprop design (df(8), $t=4.6417$ $p=0.0051$)

Similar sensitivities were found in the comfort questionnaire;

Base turboprop design and Unequal turboprop design (df(8), t=11.7954, p=0.003),

Base turboprop design and Staggered turboprop design (df(8) , t=9.7980 , p=0.003).

These results indicated that the tasks chosen were sensitive to the impact of spectral differences as well as the previously found sensitivity to sound intensity. These results confirm that stimuli that elicit a discomfort effect can also cause an impairment of cognitive performance. The results, however, go on to highlight the importance of stimuli specific cognitive tasks to determine the impact of the stimuli rather than simply cognitive performance. This novel highlight also provided a theoretical assumption for the model, the assumption being that a discomfort stimulus has an additional impact which is impaired task performance on a stimulus specific task. This allowed for the model to use task performance as one of the metrics which could be used to predict comfort ratings.

In the examination of the impact of fidelity on comfort it was found that comfort is significantly impacted by levels of fidelity. This impact is not only determined by the fidelity level of the environment but also the level of fidelity enacted by the response of cues within the environment such as movement of noise. The difference in these fidelity types in the study carried out were designated as environmental fidelity and audio-functional fidelity respectively. Fidelity changes were shown to have an impact upon comfort and annoyance ratings and in the case of environmental fidelity also upon task performance. In the case of higher levels of fidelity participants reported higher comfort, lower annoyance, and produced less errors when completing a task.

Differentiating between the two types of fidelity used in the study is integral as it reflects Slater's (2009) distinction between Psi and PI. This division between fidelity types not only represents the various methods of changing perceived fidelity in an experimental setting, but also provides a useful context for understanding the two axes of the proposed comfort model. In the case of Psi, audio-functional fidelity is best reflected with low fidelity in this case leading to a disparity between participants' expectations of cue reaction and the simulations reactions. The disparity caused is mirrored in the Rage/Relief axis of the model which as mentioned previously comes into effect when a task or environment is varied from the expectations of the participant, leading to a discomfort response. When PI is taken into account this is best represented by the environmental fidelity changes. These changes are in the physical and virtual space the participants are exposed to.

This study showed that it was possible not only to determine that a stimulus or experience was having an impact, but also that with the application of correct metrics the impact of the environment in which the stimulus was experienced and its presentation method could also be measured. This allowed for this understanding to be incorporated into the model.

The result of these pieces of information is a novel model of comfort which can provide a prediction of participant comfort from assessment of annoyance and stimulus-specific task performance. These assessments map onto two axes, each representing a continuum of reactions to discomfort stimuli. The mapping of annoyance and performance on the model was justified through the comparison of correlation strengths of the assessments against comfort (sections 5.5.2 and 5.6). The model reliably predicts the impact of discomfort stimuli on participants' comfort ratings, and in addition is capable of detecting changes in the method of presentation or environmental factors involved in the discomfort reaction. From the statistical analysis of the data, when applied to the model, two viable methods of analysis have been proposed. The first, which separates participants treating them as individuals, applying individual participant's data across multiple conditions to the model. This method has the advantage of identifying anomalous participants and maintaining the integrity of each individual participant's internal schema of comfort and tolerance to discomfort. Using participant's data in relation only to themselves do, however, come with some difficulties. When examining data in this way it must be taken into account that each participant will have a specific schema and tolerance for comfort, this makes cross-participant comparison more difficult with this model. Therefore, this method is best suited to representing changes in stimuli for each participant.

The second method of analysis uses an averaging approach which incorporates every participant into one usage of the model. This method of analysis provides the advantage of using participants' data of the impact of discomfort changes on an 'average' member of the population. This method does also come with one primary drawback which is that when the data is averaged in this way what is represented on the model is an average for the population. Hence due to the reduction in detail of the data this method is best used to represent overarching changes in a population, such as being used on an industry level. However, while these methods have different advantages and drawbacks, they both deal with the issue of cross-participant variations which as previously mentioned are often present in methods of measuring and modelling comfort. The issue of cross-participant variations is one of importance as it can skew the understanding of data acquired due to differences between participants in their tolerances and schemas.

The model of audio comfort has fulfilled the primary aim of this thesis. The model is capable of being applied to a set of audio stimuli or different contexts in which stimuli are presented and accurately display their impacts upon perceived comfort. This model is functional in both research and industry. In the case of industry this model was developed during the ImPAcTA project with Dowty Propellers and could easily be used in future industry projects to determine the human factors impact of new technology of audio comfort. This applicability is particularly useful in an ever-expanding and evolving industry such as aircraft design, however the model is not limit to this industry. Due to the nature of the assessments used in the function of this model (performance and annoyance) this model is not restricted to the aviation industry, it could feasibly be used in any context in which comparison of stimuli which may cause a difference in auditory comfort might arise such as discomfort from new developments in car engines or to assess impact of areas with high levels of noise pollution. The research presented in also Chapter 4 clearly shows the capability for this model to be used in the case of VR and beyond.

7 Conclusion and Recommendations.

To conclude, this thesis has put forward studies which test the assumptions of the developed model of comfort. The studies detailed in Chapter 3 tested the applicability of task performance measures as a method of assessing comfort levels, along with piloting the first iteration of the comfort questionnaire. This chapter also provided the theoretical consideration that interference with a task can be used to measure discomfort.

Chapter 3 Findings and Conclusions:

- The BTA, ToL, and PEP were found to lack sensitivity due to the cognitive focus of these tasks being short term memory, planning, and working memory respectively rather than sound based.

- The SIT was found to be a sensitive task to noise intensity changes and spectral differences changes due to its audio specific task nature
- The comfort questionnaire was also found to be a sensitive metric in the case of sound intensity changes and spectral differences changes.
- The comfort questionnaire and SIT had mirrored responses in the experiments showing as discomfort increased so did errors on the SIT.
- Experiments 1 and 2 allowed for the conclusion that interference with stimuli specific tasks (i.e. in this case, a sound-based task) can lead to discomfort.
- Following from this it allowed for the conclusion that useful objective measures of comfort would be task performance metrics that are interfered with by the stimuli assessed

The studies in Chapter 4 expand on the feasibility of using annoyance as another metric of comfort alongside task performance. This study explores the impact of simulator fidelity upon participant's comfort and provides a novel view into the application of measures of audio comfort to stimuli outside of the physical properties of sound, such as environment or presentation method.

Chapter 4 Findings and Conclusions:

- The fidelity of a simulation has a significant impact upon comfort and annoyance ratings
- The fidelity of a simulation has a significant impact upon task performance
- The impact on task performance is more specific to environmental fidelity than audio-functional fidelity without extensive training.

- In environmental fidelity, the impact on comfort and annoyance ratings is mirrored by task performance, as annoyance increases and comfort decreases errors on task performance increase.
- This difference between audio-functional fidelity and environmental fidelity shows that they have distinct impacts upon participants.
- These distinct impacts show that there are different methods or pathways of creating discomfort or comfort.
- The data from this study produced the conclusion that participants in high fidelity simulations react realistically to stimuli and events
- The reaction of participants is shown to extend to comfort responses explaining changes in comfort, annoyance, and task performance

The data from Chapter 4 is further discussed with additional data from a subsequent large-scale comfort study detailed in Chapter 5, in which both are applied to the model of audio comfort's framework as a test of the efficacy of the model itself. This analysis showed strong correlations with the two presented methods of interpretation and confirmed the model's key assumption. From the data presented the novel model of audio comfort can predict the impact of direct audio stimuli as well as the impact of the environment or presentation method of such stimuli. The model also allows for the comparison between multiple stimuli across a population or within a singular participant across multiple trails without the issue of cross-participant variability. The model of audio comfort carries out these functions though application of task performance and annoyance ratings to determine effects.

Chapter 5 Findings and Conclusions:

- The model of audio comfort states that audio comfort is predicted by task performance and annoyance ratings

- The model posits that this can be represented on a four-quadrant graph
- Movement on the Y axis of this graph is primarily though not exclusively driven by stimuli experience, this movement is measured by task performance
- Movement on the X axis of this graph is primarily though not exclusively driven by participants' expectations, this movement is measured through annoyance ratings
- The data from Chapter 4 when viewed in the context of this model follows these assertions
- Data from the large-scale study detailed in Chapter 5 was used to confirm the model's assumption that annoyance and task performance can be used to predict comfort
- The model's assumptions were concluded to be accurate and applicable to research data

To summarise the research within this thesis has provided a reliable application for task performance in measuring audio comfort. Carried out studies which provide a novel view into the environmental and presentation methods of stimuli experience and how these can be assessed for impacts upon audio comfort. This thesis has also produced a novel model of audio comfort which reduces the impact of cross-participant variability. The applications of this model of audio comfort are viable both in academic comfort research and industry product development.

7.1 Recommendations for Further Work

This thesis has developed a novel model for assessing comfort which has been tested in regard to sound stimuli and participants' assessment of sound stimuli when exposed to various environmental stimuli. However, this model of comfort would benefit from additional testing such as use of threshold testing to examine an average shift point for participants. In this research it would be suggested that standard stepwise assessments would be carried out to determine a participant's tolerance to change in a stimulus and its potential to impact comfort.

Further research would also be recommended to elaborate on the novel "comfort space" proposed in this thesis. This concept in the context of the model requires further research to define its shape and how it might vary from participant to participant. This once again would most likely benefit from threshold testing to determine the measurable point at which a participant leaves their comfort space. This as the literature in section 2.2 discusses does have the issue that participant's tolerance and reactions to discomfort vary and so a "universal comfort space" may be elusive.

While the testing of this model's assumptions has been carried out with sound stimuli it is possible that the assumptions this model holds would function for other sensory aspects. Therefore, further research with this model could include sensory modalities that could be involved in performance measures such as tactile comfort and visual comfort.

Human vibration sensitivity is almost a mirror image of the hearing system: while our hearing is best around 1,000-4,000 Hz, and humans are not very sensitive below 100 Hz, we are most sensitive to vibrations between 1-4 Hz, and do not feel vibrations much above 100 Hz. This means that the energy peaks at around 50 Hz are barely audible but are likely to be felt as vibration. A major difference between the two senses is the dynamic range: humans tolerate sound intensity differences of 80-90 dB above hearing threshold before suffering significant discomfort, while for vibrations this dynamic range is only about 20-30 dB (Fig 71).

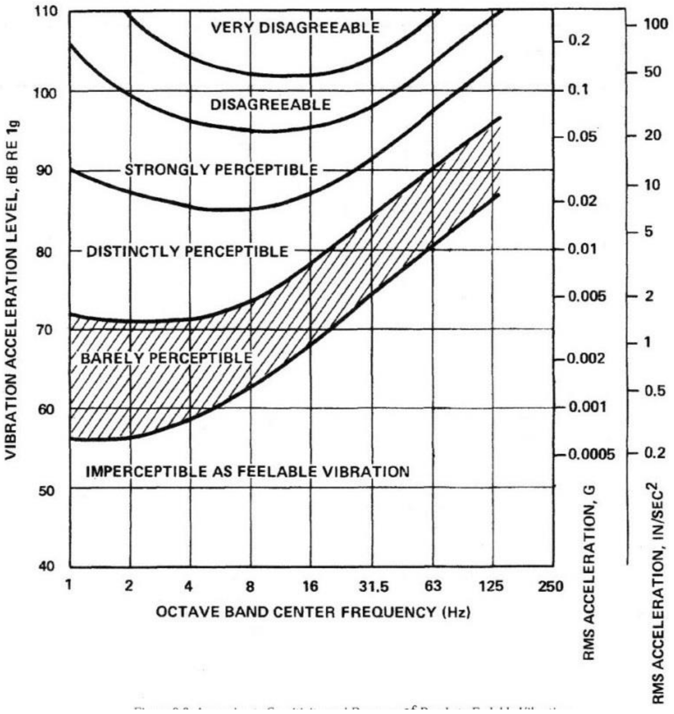


Figure 70 Vibration detection and response threshold for vibration. From: anon (1995) Noise and Vibration Control Manual, Department of the Army (No. 5-805-4) and the Air Force (No. 88-37)

Whether vibration affects comfort depends on the absolute levels, given that the frequencies in question are relatively low, it may be possible to use structural, passive, or active vibration attenuation measures to ensure levels stay below the detection threshold. It is, however, worth bearing in mind that there are only about 20dB level difference between ‘barely perceptible’ and ‘disagreeable’. The fact that the model does not take into acount the impact of vibration on comfort when assessing auditory comfort can be considered to be a limitation of the model. This limitation of the model may be noteable as there is a relationship between vibration and auditory sensitivity in human percpetion as well as the tendancy of both to be present in cases where auditory discomfort occurs. Future research could evaluate the impact of vibration and the interaction of vibration and noise on comfort. This could be incoporated into the model and would extend the model’s ability to provide comfort assessments.

Due to delays in the wind tunnel test campaign, it was not possible for Dowty to use the model with the acoustic results from the tests. The implications for design process are that with CFD simulated sound files such as were used in section 5.5.1 can be tested using this model, therefore informing the impact upon passengers' comfort and so determining if the design would have a more negative effect upon passenger comfort than the previous design. This would allow the industry to maintain high standards of passenger comfort, this not only has the effect of increasing customer satisfaction but also provides marketing value.

References

Allen, T., and Tarr, A., (2005). "Driving simulators for commercial truck drivers – Humans in the loop", *Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, pp.335-341, University of Iowa, USA*

Atkinson, R.C. and Shiffrin, R.M. (1968). "Human memory: A proposed system and its control processes", In K.W. Spence (Ed.). The psychology of learning and motivation: advances in research and theory Vol. 2(pp.89-195). New York: Academic Press

Barakos, G. N., Johnson, C.S. (2016). "Acoustic comparison of propellers". *International Journal of Aeroacoustics,* Vol 15, Issue 6-7, pp. 575 – 594

Benton, S. (1997b). "Low-frequency noise and the impact upon an individual's quality of life: Case study reports". *Journal Low-frequency Noise Vibration, 16, 203-208*

Bhiwapurkar, M. K., Saran, V. H., Goel, V. K., Mansfield, N. N., and Berg, M. M. (2009). "Study of human comfort under thermal and vibratory environment using physiological indices". *International Congress on Sound and Vibration, 7*4376-4383.

Candas, V., and Dufour, A. (2005). "Thermal Comfort: Multisensory Interactions?" *Journal of Physiological anthropology and Applied Human Science*, *24*(1), 33-36. doi:10.2114/jpa.24.33

Clark. C., Sorqvist. P., (2012). "A 3 year update on the influence of noise on performance and behaviour". *Noise and Health. Vol. 14:61.*

Clevenson, S. A., Dempsey, T. K., and Leatherwood, J. D. (1978). "Effect of vibration duration on human discomfort*". Sherman A. Clevenson, Thomas K. Dempsey, and Jack D. Leatherwood, Langley Research Center*. *[Washington]: National Aeronautics and Space Administration, Scientific and Technical Information Office; Springfield,*

Cooper, N., Milella, F., Pinto, C., Cant, I., White, M., Meyer, G. (2018). "The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment" *Plos one, 13(2): e0191846.https://doi.org/10.1371/journal.pone.0191846*

Cox, M, R., Alexander, C, G/. Taylor, M, I., and Gray, A, G., (1997). "The contour test of loudness perception". *Ear Hear, 18(5):388-400.*

Dempsey, T. K., Leatherwood, J. D., and Clevenson, S. A. (1976b). “Noise and vibration ride comfort criteria”. *Journal of The Acoustical Society of America*, *60*(S1), S122. doi:10.1121/1.2003157

Dempsey, T. K., Leatherwood, J. D., (1975). “Vibration Simulator Studies for The Development of Passenger Ride Comfort Criteria”. *NASA Langley Research Centre. NASA-TM-X-3295. doi: 10.1177/154193127602001401.*

Dowty. (2015). “The Aerodays 2015 innovation event spotlights Dowty Propellers’ research efforts.” http://dowty.com/feature-stories/the-aerodays-2015-innovation-event-spotlights-dowty-propellers-research-efforts/#. (Accessed 06/03/17).

Environmental Research and Consultancy Department. (2016). “Aircraft noise and health effects: Recent findings”. Civil Aviation Authority. https://www.caa.co.uk/Consumers/Environment/Noise/Noise/ (Accessed 07/03/17)

Fuller, C. R. (1998). “Structural Acoustic Characteristics of Aircraft and Active Control of Interior Noise”. Research Report on NASA Grant NAG-1-1681.

Harrison, J. K. (2013, December 2nd). “The Quiet Cabin: No Simple Solution”. Retrieved from http://www.ainonline.com/aviation-news/aviation-international-news/2013-12-02/quiet-cabin-no-simple-solution (Accessed 07/03/17)

Hood, D, J., (1968). “Observations upon the Relationship of Loudness Discomfort Level and Auditory Fatigue to Sound-Pressure Level and Sensation Level”. *The Journal of the Acoustical Society of America. 44(4):959-64*

Ivosevic, J., Miljkovic, D., Krajcek, K., (2012). “Comparative Interior Noise Measurements in a Large Transport Aircraft – Turboprops vs Turbofans”. *Congress of Alps-Adria Acoustics Association. Petrcane, Croatia, Acoustical Society of Croatia.*

Ivosevic, J., Mihalincic, I., Bucak, T., (2010). “Noise analysis of a twin piston engine aircraft”. *Congress on Sound and Vibration. Ljublana, Slovenia, European Acoustics Association.*

Jeon, J, Y., You, J., Jeong, C, I., Kim, S, Y., and Jho, M, J., (2011). “Varying the spectral envelope of air-conditioning sounds to enhance indoor acoustic comfort”. *Building and Environment, 46, 739-746.*

Johnson, S. J., Guediri, S. M., Kilkenny, C., and Clough, P. J. (2011). “Development and Validation of a Virtual Reality Simulator: Human Factors Input to Interventional Radiology Training”. *Human Factors*, *53*(6), 612-625. doi:10.1177/0018720811425042

Knobel, K. A., and Sanchez, T. G. (2006) "Loudness Discomfort Level in normal hearing individuals". *Pro Fono 18. 31-40.*

Kolcaba, K., (2001). "Evolution of the Mid Range Theory of comfort for Outcomes Research". *Nursing Outlook, 49.*

Khan, M. S., and Sundstrom, J. J. (2004). "Fr3.X1.2 – Vibration comfort in Swedish inner-city trains – A survey on passenger posture and activities". *International Congress on Acoustics, 5*V-3733-V-3736.

Kulesa, A., Krzywinski, M., Blainey, P., and Altman, N. (2015). "Sampling distributions and the bootstrap. The bootstrap can be used to assess uncertainty of sample estimates". *Nature methods. 12(6):*477-478

Larsson, P., Västfjäll, D., and Kleiner, M. (2008). "Effects of auditory information consistency and room acoustic cues on presence in virtual environments". *Acoustical Science and Technology*, *29*(2), 191-194. doi:10.1250/ast.29.191

Leatherwood, D, J., and Dempsey, T, K., (1976a). "Psychophysical relationships characterising human response to whole-body sinusoidal vertical vibration". *NASA TN D-8188. National Aeronautics and Space Administration.*

Leventhall, H. G. (2004). "Low-frequency noise and annoyance". *Noise and Health*, *6*(23), 59-72.

Leventhall, G., and Knopper, L. D. (2015). "Health-based audible noise guidelines account for infrasound and low-frequency noise produced by wind turbines". *Frontiers in Public Health*, *3*31. doi:10.3389/fpubh.2015.00031

Lewis, R., Strachan, A., and Smith, M. M. (2012). "Is high fidelity simulation the most effective method for the development of non-technical skills in nursing? A review of the current evidence". *The Open Nursing Journal*, *6*82-89. doi:10.2174/1874434601206010082

Lin, J.W., Duh, B.L., Abi-Rached, H., Parker, D.E. and Furness, T.A.III (2002). "Effects of field of view on presence, enjoyment, memory and simulator sickness in a virtual environment". In *Proceedings of IEEE Virtual Reality 2002*, pp. 164-171

Looze, d. M., Kuijt-Evers, L., and Dieen, v. J. (2003). "Sitting comfort and discomfort and the relationships with objective measures". *Ergonomics* 10.1080/0014013031000121977.

Luboz, V., Zhang, Y., Johnson, S., Song, Y., Kilkenny, C., Hunt, C., Woolnough, H. Guediri,S. Zhai,J.,Odetoyinbo,T., Littler,P., Fisher,A., Hughes,C., Chalmers,N., Kessel,D., Clough,P.J., Ward,J., Phillips,R., How,T., Bulpitt,A., John,N.W., Bello, F., Gould, D. (2013). "ImaGiNe Seldinger: First simulator for Seldinger technique and angiography training". *Computer Methods and Programs in Biomedicine, 111*419-434. doi:10.1016/j.cmpb.2013.05.014

McCrudy, D. A. (1988). "Annoyance Caused by Advanced Turboprop Aircraft Flyover Noise". *NASA Technical paper 2782.*

Mellert, V., Baumann, I., Freese, N., and Weber, R. (2008). "Impact of sound and vibration on health, travel comfort and performance of flight attendants and pilots". *Aerospace Science and Technology, 12*(Aircraft noise reduction), 18-25. doi:10.1016/j.ast.2007.10.009

Metzger, F. B. (1995). "An Assessment of Propeller Aircraft Noise Reduction Technology". NASA Contractor Report 198237.

Meyer, G, F., Wong, L, T., Timson, E., Perfect, P., and White, M, D. (2012). "Objective Fidelity Evaluation in Multisensory Virtual Environments: Auditory Cue Fidelity in Flight Simulation". *PLOS one, Vol 7, Issue 9.* e44381. https://doi.org/10.1371/journal.pone.0044381.

Meyer, G, F., and Morse, R. (2003). "The intelligibility of consonants in noisy vowel-consonant-vowel sequences when the vowels are selectively enhanced". *Speech Communication. 41. 429-440.*

Meyer GF, Shao F, White MD, Hopkins C and Robotham AJ, "Modulation of Visually Evoked Postural Responses by Contextual Visual, Haptic and Auditory Information: A 'Virtual Reality Check'", *PLoS ONE* 8(6): e67651. doi:10.1371/journal.pone.0067651 pp. 1- 15, 2013

McNulty, J. P., and McNulty, S., (2009). "Acoustic noise in magnetic resonance imaging: An ongoing issue". *Radiography, 15*320-326. doi:10.1016/j.radi.2009.01.001

Metzger, P., (1994). "Komfortverstandis bei Kraftfahrern". *Diplomarbeit an der Freien Universitat Berlin. (PhD Thesis)*

Mwanalushi, K., (2012). "Return of the turboprops". *Low Cost and regional Airline Business. Vol 7. (http://www.lowcostandregional.com/feature/return-of-the-turboprops?session_id=hj6qf3u005q9816op1rqurkpr2 accessed 15/10/2016)*

National Institute on Deafness and Other Communication Discorders (NICD). "Noise-Induced Hearing Loss". *US Department of Health and Human Services.* NIH Publication No. 14-4233.

Nor, M. J. M., Fouladi, H. M., Nahvi, H., Ariffin, K. A., (2008). "Index for vehicle acoustical comfort inside a passenger car". *Applied Acoustics, Vol 69. Issue 4.* 343-353.

Oborne, D. J. (1978). "Techniques available for the assessment of passenger comfort". *Applied Ergonomics*, *9*(1), 45-49. doi:10.1016/0003-6870(78)90258-2

Oborne, D. (1978). "Passenger comfort--an overview". *Applied Ergonomics*, *9*(3), 131-136

.

Oborne, D., and Clarke, M. (1975). "Questionnaire surveys of passenger comfort". *Applied Ergonomics*, *6*(2), 97-103.

Oborne, D. J. (1977). "Vibration and passenger comfort". *Applied Ergonomics*, *8*(2), 97-101. doi:10.1016/0003-6870(77)90060-6

Ozcan, H. K., and Nemlioglu, S. (2006). "In-Cabin Noise Levels During Commercial Aircraft Flights". *Canadian Acoustics*, Vol 34, 31-35.

Padfield, G. D., and White, M. D. (2003). "Flight simulation in academia HELIFLIGHT in its first year of operation at the University of Liverpool". *Aeronautical Journal, 107(1075), 529-538.*

Pawlaczyk-Łuszczyńska, M., Dudarewicz, A., and Śliwińska-Kowalska, M. (2007). "Theoretical predictions and actual hearing threshold levels in workers exposed to ultrasonic noise of impulsive character—a pilot study." *International Journal of Occupational Safety and Ergonomics (JOSE). 2007; 13*(4):409–18.

Perfect, P., White, M, D., Gubbels, A, W., and Padfield, G, D. (2013). "Rotorcraft simulation fidelity: new methods for quantification and assessment". *The Aeronautical Journal. Vol 117(1189):*235-282.

Phillips, W.A. (1974). "On the distinction between sensory storage and short-term visual memory". *Perception and Psychophysics, 16,* 283-290.

Quehl, J., (2001). "Comfort studies on aircraft interior sound and vibration". *University of Oldenburg. (PhD Thesis).*

Reinhold, K., Kalle, S., Paju, J. (2014). “Exposure to high or low frequency noise at workplaces: Differences between assessment, health complaints, and implementation of adequate personal protective equipment”. *Argonomy Research 12(3),* 895-906.

Richards, L.G., Jacobson I.D., (1974). “Ride Quality Evaluation 1: Questionnaire Studies of Airline Passenger Comfort”. *Ergonomics, 18*(2), 129-150.

Robinson, H. J., Fernholz, M. C. (1998). “The Influence of Lamination Angles on the Interior Noise Levels of an Aircraft”. *NASA Technical Memorandum 110250.*

Rolls, E, T., Treves, A. (1998). “Neural Networks and Brain Function”. Oxford, Oxford University Press

Rolls, E, T. (1990). “A Theory of Emotion, and its Application to Understanding the Neural Basis of Emotion”. *Cognition and Emotion. 4*, 161-190.

Ruotolo, F., Maffei, L., Di Gabriele, M., Iachini, T., Masulo, M., Ruggiero, Gennaro., and Senese, V, P. (2012). “Immersive virtual reality and environmental noise assessment: An innovative audio-visual approach”. *Environmental Impact Assessment Review. 41* 10-20.

Rusconi, E., Kwan, B., Giordano, L, B., Umilta, C., and Butterworth, B., (2005). “Spatial representation of pitch height: the SMARC effect”. *Cognition. 99.* 113-29.

Salvendy, G. 2012. *Handbook of Human Factors and Ergonomics. John Wiley and Sons, New Jersey, 1752 pp.*

Sanchez, T, G., Moraes, F., Casseb, J., Cota, J., Freire, K., and Roberts, L, E., (2016). “Tinnitus is associated with reduced sound level tolerance in adolescents with normal audiograms and optoacoustic emissions”. *Scientific Reports. 6;6:27109. doi:10.1038/srep27109.*

Schretlen, D., Bobholz, J. H., and Brandt, J. (1996). “Development and psychometric properties of the brief test of attention”. *The Clinical Neuropsychologist, 10(1),* 80-89.

Sinha, A., Johnson, S., Hunt, C., Woolnough, H., Vidal, F., John, P., Nigel, W., Villard, P-F., Holbray, R., Bello, F., Gould, D. (2009). “Importance of virtual reality simulators in interventional radiology”. *Cardiovascular and Interventional Radiological Society of Europe - CIRSE 2009, Sept 2009, Lisbon, Portugal. hal-00510329*

Slater, M., (2009). “Place Illusion and plausibility can lead to realistic behaviour in immersive virtual environments”. *Philosophical Transactions of the Royal Society. 364,* 3549-3557.

Sengupta, G. (1977). "Methods of reducing low-frequency cabin noise and sonically induced stresses, based on the intrinsic structural tuning concept". *Conference on Structures, Structural Dynamics and Materials, 18th; March 21-23, 1977. San Diego, CA; US*

Vastfjall, D., Larsson, P., and Kleiner, M., (2002). "Emotion and Auditory Virtual Environments: Affect-Based Judgements of Music Reproduced with Virtual Reverberation Times". *CyberPsychology and Behaviour. 5(*1):19-32 DOI: 10.1089/109493102753685854

Ward, G. and Allport, A. (1997). "Planning and problem-solving using the five-disc Tower of London task". *Quarterly Journal of Experimental Psychology, 50A, 1,* 49-78.

Waye, K, P., Rylander, R., Benton, S., and Leventhall, H. G. (1997). "Effects on Performance and Work Quality Due To Low-Frequency Ventilation Noise". *Journal of Sound and Vibration*, *205*(4), 467.

White, M, D., Perfect, P., Gareth, P, D., Gubbels, A, W., and Berryman, A, C. (2012). "Acceptance testing and commissioning of a flight simulator for rotorcraft simulation fidelity research". *Journal of Aerospace Engineering. Vol 227*(4),663-686 DOI: 10.1177/0954410012439816

Whitlow, J.B. Jr., Sievers, G.K. (1984). "Fuel savings potential of the NASA Advanced Turboprop program". *Conference: Aviation fuel conservation symposium, Washington, DC, USA, 10 Sep 1984*

Wills, A. J. (2003). "Cognitive Psychology Experimental Package". *Retrieved from* http://willslab.co.uk. (Accessed 10/10/13)

Witmer, B. G., and Singer, M. J. (1994). "Measuring immersion in virtual environments". (ARI Technical Report 1014). Alexandria*, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.*

Woszczyk, W., Bech, S., Hansen V. (1995). "Interactions Between Audio - Visual Factors in a Home Theatre System: Definition of Subjective Attributes", *99th Audio Eng. Soc. Convention., New York, Preprint No. 4133*

Yang, W., Kang, J., (2004). "Acoustic comfort evaluation in urban open public spaces". *Applied Acoustics, Vol 66, Issue 2,* 211-229.

Zeyada, Y., and Hess, R. A. (2000). "Modeling Human Pilot Cue Utilization with Applications to Simulator Fidelity Assessment". *Journal of Aircraft*, *37*(4), 588-597.

Zhang, D., Zhang, M., and Lui, D., Kang, J. (2016). “Soundscape evaluation in Han Chinese Buddhist temples”. *Applied Acoustics, 111, 188-197.*

Ziaran, S. (2014). “The assessment and evaluation of low-frequency noise near the region of infrasound”. *Noise Health, 16(68):10-7. doi: 10.4103/1463-1741.127848.*

Appendices

Appendix A - Descriptive Statistics for Pilot Study (Chapter 3)

A.1 Tower of London Descriptive Statistics

This table shows the descriptive statistics of the Tower of London task. This includes mean (M) and standard deviation (SD) of the planning time and execution times of participants at each dB level in the pilot study.

Tower of London Planning & Exection Time (ms) N=21							
Decidels 70		Decibels 76		Decibels 70		Decibels 76	
Planning Time		Planning Time		Execution Time		Execution Time	
M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
7.97	6.93	6.13	3.25	11.62	8.99	10.97	9.01

A.2 Tower of London Statistics

Chi squared test to assess impact of decibel level on number of errors on Tower of London Task

χ^2 (1, N=21) = 1.546, p = .107

A.3 Brief Test of Attention Descriptive Statistics

This table shows the mean (M) and Standard deviation (SD) of the number of errors participants made at each noise level 70 dB(A) and 76 dB(A) during the Brief test of attention task in the pilot study.

Number of Errors N=21			
Decibels 70		Decibels 76	
BTA		BTA	
M	*SD*	*M*	*SD*
1.86	2.71	1.43	1.78

A.4 Brief Test of Attention Statistics

Paired t-test to determine impact of decibel level on errors in the Brief Test of Attention Task

t (20) = .629, p = .2685

Appendix B

B.1 Matlab code for SIT

```
% new contest u
global sig n dBatt
stimlist = ''; %change paths
datapath = ''; %change paths
stimNo=200; firstStim=1;  % no of stimuli to be used and offset in list
dBatt =0;

   % set up stimulus sequence
noStim=1; noRead =0;
fid1 = fopen(stimlist,'r');

for c = 1:firstStim + stimNo -1;

      clear name;
      nc=1;
      c = fscanf(fid1,'%c',1);
    while ((c ~= ' ')&(~feof(fid1)));name(nc)=c; c = fscanf(fid1,'%c',1); nc=nc+1;end;
    m1 = fscanf(fid1,'%f',1);
    m2 = fscanf(fid1,'%f',1);
    m3 = fscanf(fid1,'%f',1);
      m4 = fscanf(fid1,'%f\n',1);

      noRead=noRead+1;
      if(noRead >= firstStim)
         sig(noRead).name =name;
         sig(noRead).start = round(m1);
         sig(noRead).conStart = round(m2);
         sig(noRead).conStop = round(m3);
         sig(noRead).stop = round(m4);
         sig(noRead).R = 0; % placeholder for response


         switch (name(4))
         case 'B'
            sig(noRead).id=1;
         case 'C'
            sig(noRead).id=2;
         case 'D'
            sig(noRead).id=3;
            if (name(5)) == 'H'; sig(noRead).id=16; end; %th
         case 'F'
            sig(noRead).id=4;
         case 'G'
            sig(noRead).id=5;
         case 'J'
            sig(noRead).id=6;
         case 'K'
            sig(noRead).id=7;
         case 'L'
```

```
        sig(noRead).id=8;
      case 'M'
        sig(noRead).id=9;
      case 'N'
        sig(noRead).id=10;
      case 'P'
        sig(noRead).id=11;
      case 'R'
        sig(noRead).id=12;
      case 'S'
        sig(noRead).id=13;
        if (name(5)) == 'H'; sig(noRead).id=14; end;
      case 'T'
        sig(noRead).id=15;
      case 'V'
        sig(noRead).id=17;
      case 'W'
        sig(noRead).id=18;
      case 'Y'
        sig(noRead).id=19;
      case 'Z'
        sig(noRead).id=20;
       otherwise
         fprintf('Error assigning id, name is %s\n',name);
      end

    end;

end;




n=0;

  fprintf(' n = %d\n',n);
  %for (i=1:stimNo)
  %fprintf(' Stim %d = %s   con in %d-%d\n',i,sig(i).name,sig(i).conStart,sig(i).conStop)
  %end;




% build user interface

bStart=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.2,0.95,0.6,0.05],...
   'String','- - -  START  - - -','Callback','if(n==0)n=1; contest_a(1); end;');

bR1=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.02,0.7,0.15,0.15],...
   'String','aBa','Callback','sig(n).R=1; n=n+1;  contest_a(1);');
bR2=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.22,0.7,0.15,0.15],...
   'String','aCHa','Callback','sig(n).R=2; n=n+1;  contest_a(1);');
bR3=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.42,0.7,0.15,0.15],...
   'String','aDa','Callback','sig(n).R=3; n=n+1;  contest_a(1);');
bR4=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.62,0.7,0.15,0.15],...
    'String','aFa','Callback','sig(n).R=4; n=n+1;  contest_a(1);');
bR5=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.82,0.7,0.15,0.15],...
   'String','aGa','Callback','sig(n).R=5; n=n+1;  contest_a(1);');
```

```
bR6=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.02,0.5,0.15,0.15],...
   'String','aJa','Callback','sig(n).R=6; n=n+1;  contest_a(1);');
bR7=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.22,0.5,0.15,0.15],...
   'String','aKa','Callback','sig(n).R=7; n=n+1;  contest_a(1);');
bR8=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.42,0.5,0.15,0.15],...
   'String','aLa','Callback','sig(n).R=8; n=n+1;  contest_a(1);');
bR9=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.62,0.5,0.15,0.15],...
   'String','aMa','Callback','sig(n).R=9; n=n+1;  contest_a(1);');
bR10=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.82,0.5,0.15,0.15],...
   'String','aNa','Callback','sig(n).R=10; n=n+1;  contest_a(1);');

bR11=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.02,0.3,0.15,0.15],...
   'String','aPa','Callback','sig(n).R=11; n=n+1;  contest_a(1);');
bR12=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.22,0.3,0.15,0.15],...
    'String','aRa','Callback','sig(n).R=12; n=n+1;  contest_a(1);');
bR13=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.42,0.3,0.15,0.15],...
   'String','aSa','Callback','sig(n).R=13; n=n+1;  contest_a(1);');
bR14=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.62,0.3,0.15,0.15],...
   'String','aSHa','Callback','sig(n).R=14; n=n+1;  contest_a(1);');
bR15=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.82,0.3,0.15,0.15],...
   'String','aTa','Callback','sig(n).R=15; n=n+1;  contest_a(1);');

bR16=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.02,0.1,0.15,0.15],...
   'String','aTHa','Callback','sig(n).R=16; n=n+1; contest_a(1);');
bR17=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.22,0.1,0.15,0.15],...
   'String','aVa','Callback','sig(n).R=17; n=n+1;  contest_a(1);');
bR18=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.42,0.1,0.15,0.15],...
   'String','aWa','Callback','sig(n).R=18; n=n+1;  contest_a(1);');
bR19=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.62,0.1,0.15,0.15],...
   'String','aYa','Callback','sig(n).R=19; n=n+1;  contest_a(1);');
bR20=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.82,0.1,0.15,0.15],...
    'String','aZa','Callback','sig(n).R=20; n=n+1;  contest_a(1);');

%bStop=uicontrol(gcf,'Style','push','Units','normalized','Position',[0.2,0.005,0.6,0.05],...
 %  'String','- - - Stop - - -','Callback','n=200;  contest_a(1);');

%bStop=Exit_Callback(hObject, handles)
% hObject    handle to Exit (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
%close(handles.figure1)
```

Appendix C – Original Presence Questionnaire

(Witmer & Singer, Vs. 3.0, Nov. 1994)*
Revised by the UQO Cyberpsychology Lab (2004)

Participant number:________________ Age:_____ **SEX:** male / female

Characterize your experience in the environment, by marking an "X" in the appropriate box of the 7-point scale, in accordance with the question content and descriptive labels. Please consider the entire scale when making your responses, as the intermediate levels may apply. Answer the questions independently in the order that they appear. Do not skip questions or return to a previous question to change your answer.

WITH REGARD TO THE EXPERIENCED ENVIRONMENT

1. How much were you able to control events?

|________|________|________|________|________|________|________|
NOT AT ALL SOMEWHAT COMPLETELY

2. How responsive was the environment to actions that you initiated (or performed)?

|________|________|________|________|________|________|________|
NOT RESPONSIVE MODERATELY RESPONSIVE COMPLETELY RESPONSIVE

3. How natural did your interactions with the environment seem?

|________|________|________|________|________|________|________|
EXTREMELY ARTIFICIAL BORDERLINE COMPLETELY NATURAL

4. How much did the visual aspects of the environment involve you?

|________|________|________|________|________|________|________|
NOT AT ALL SOMEWHAT COMPLETELY

5. How natural was the mechanism which controlled movement through the environment?

|________|________|________|________|________|________|________|
EXTREMELY ARTIFICIAL BORDERLINE COMPLETELY NATURAL

6. How compelling was your sense of objects moving through space?

|________|________|________|________|________|________|________|

NOT AT ALL MODERATELY COMPELLING VERY COMPELLING

7. How much did your experiences in the virtual environment seem consistent with your real world experiences?

|________|________|________|________|________|________|________|

NOT CONSISTENT MODERATELY CONSISTENT VERY CONSISTENT

8. Were you able to anticipate what would happen next in response to the actions that you performed?

|________|________|________|________|________|________|________|

NOT AT ALL SOMEWHAT COMPLETELY

9. How completely were you able to actively survey or search the environment using vision?

|________|________|________|________|________|________|________|

NOT AT ALL SOMEWHAT COMPLETELY

10. How compelling was your sense of moving around inside the virtual environment?

|________|________|________|________|________|________|________|

NOT COMPELLING MODERATELY COMPELLING VERY COMPELLING

11. How closely were you able to examine objects?

|________|________|________|________|________|________|________|

NOT AT ALL PRETTY CLOSELY VERY CLOSELY

12. How well could you examine objects from multiple viewpoints?

|________|________|________|________|________|________|________|

NOT AT ALL SOMEWHAT EXTENSIVELY

13. How involved were you in the virtual environment experience?

|________|________|________|________|________|________|________|
NOT INVOLVED — MILDLY INVOLVED — COMPLETELY ENGROSSED

14. How much delay did you experience between your actions and expected outcomes?

|________|________|________|________|________|________|________|
NO DELAYS — MODERATE DELAYS — LONG DELAYS

15. How quickly did you adjust to the virtual environment experience?

|________|________|________|________|________|________|________|
NOT AT ALL — SLOWLY — LESS THAN ONE MINUTE

16. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

|________|________|________|________|________|________|________|
NOT PROFICIENT — REASONABLY PROFICIENT — VERY PROFICIENT

17. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?

|________|________|________|________|________|________|________|
NOT AT ALL — INTERFERED SOMEWHAT — PREVENTED TASK PERFORMANCE

18. How much did the control devices interfere with the performance of assigned tasks or with other activities?

|________|________|________|________|________|________|________|
NOT AT ALL — INTERFERED SOMEWHAT — INTERFERED GREATLY

19. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

|________|________|________|________|________|________|________|
NOT AT ALL — SOMEWHAT — COMPLETELY

*Original version : Witmer, B.G. & Singer. M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence : Teleoperators and Virtual Environments*, *7*(3), 225-240. The factor structure of the Presence Questionnaire. *Presence, 14*(3) 298-312. Revised factor structure: Witmer, B.J., Jerome, C.J., & Singer, M.J. (2005). The factor structure of the Presence Questionnaire. *Presence*

Appendix D – Consent Form

UNIVERSITY OF LIVERPOOL

Title of Project: Analysis of effects of background noise on cognitive task performance and comfort.

Name of lead Researcher: Mr. M. Aldridge

Email: ps0u9264@liverpool.ac.uk

Telephone:

Address: School of Psychology
Eleanor Rathbone Building,
Bedford Street South,
Liverpool,
L69 7ZA

Please initial box

1. I confirm that I have read and understood the information sheet for the above study. ☐

2. I confirm that I do not have any hearing problems and that I am aged between 18 and 50 years. ☐

3. I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason. ☐

4. I understand that none of my personal details will be recorded and that my responses are anonymous. ☐

5. I confirm that I have been given the opportunity to ask questions and have them answered. ☐

6. I understand that I can have access to the data, and ask for it to be destroyed if I so wish ☐

7. I agree to take part in the above study. ☐

______________________________ ________________ ___________________________

Name of Participant Date Signature

_______________________________ ________________ ___________________________

Researcher Date Signature

<u>1 copy for participant, 1 copy for researcher</u>

Appendix E – Debriefing Sheet

UNIVERSITY OF LIVERPOOL

Debriefing Information

Thank you for taking part in this research.

The aim of this study was to investigate the effects of background noise on the participants' ability to complete a variety of tasks, and comfort. The tasks included; Tests of short term memory, auditory identification, a vigilance task, and finally a questionnaire. These were used to test the effect of varied levels of background noise and fidelity on the participants perceived comfort and cognitive performance. We expect both measures to be correlated.

We expect comfort rating to be affected by the situational context: we expect participants to find airplane noise in aircraft environments more acceptable than in soundproof rooms.

The results of this study will provide invaluable data in the measurement of the effects of background noise on cognitive performance. This will subsequently aid in the development of standardized measures for further study.

If you have any questions or comments regarding the experiment, please discuss them with the experimenter now. Alternatively, you can contact Georg Meyer:

School of Psychology

Eleanor Rathbone Building,

University of Liverpool

email georg@liv.ac.uk

telephone 0151 7942579.

This study has been approved by The University of Liverpool's Ethics Committee. The University has formal procedures to deal with complaints and for the reporting of adverse effects. If you or a representative wishes to raise a concern about the study, in particular about the conduct of the study or the individual involved, which would be inappropriate to raise with the principal investigator, please use the complaints procedure. Complaints should be addressed to the Research Governance Officer in RBS (ethics@liv.ac.uk; 0151 794 8727). Please provide the study name or a description, the principal investigator (Georg Meyer) and the nature of the complaint.