The Influence of Training Method on Tone Colour Discrimination

Mark Bassett

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Sydney School of Architecture, Design and Planning

University of Sydney

2018

THE UNIVERSITY OF SYDNEY

Faculty of Architecture, Design & Planning

AUTHOR'S DECLARATION

This is to certify that:

I. this thesis comprises only my original work towards the Doctor of Philosophy Degree
II. due acknowledgement has been made in the text to all other material used
III. the thesis does not exceed the word length for this degree.
IV. no part of this work has been used for the award of another degree
V. this thesis meets the *University of Sydney's Human Research Ethics Committee (HREC) requirements for the conduct of research*.

Signature(s):

Name(s): Mark Bassett

Date: 21st March 2018

The *Author's Declaration* must be inserted in the front of each thesis copy.

Abstract

The research presented in this thesis addressed the question of whether one of two training methods, *identification by continuous adjustment* (ICA) or *identification by successive approximation* (ISA), is more effective in training students using a *technical ear training program* (TETP). The work was motivated by a recognition of the schism within the audio engineering community regarding these two methods. No known empirical studies have examined the effectiveness of either training method within frequency spectrum-based student-targeted TETPs.

Preliminary work involved the exploration of the development of appropriate tests of students' tone colour discrimination ability in isolation, on tasks sufficiently different from those encountered in TETPs. Performance on these tests indicated which tasks appeared more suitable as indicators of tone colour discrimination ability. The tests were then deployed in a pilot study within a pre/post-training scenario using two groups of audio engineering students, one of which undertook an ICA and the other an ISA version of a TETP. These preliminary results indicated the suitability of a test that featured pairwise comparisons of synthetic percussive timbres to show differences in performance between the two training groups. This test was subsequently administered repeatedly in a full-scale study at regular intervals throughout a web-based TETP, in addition to before and after training.

Results of the full-scale study showed the individual differences scaling (INDSCAL)-derived stimulus spaces for both groups were similar prior to undertaking the TETP. The ISA group's post-training results were almost identical to their pre-training results, whereas the ICA groups' post-training results showed minor, but insignificant differences. Although the full-scale study found insignificant differences in

performance between training groups, the preliminary results suggest that the deployment of a pre/post-training test is an effective measure of the training method's influence on students if the test features a task that is significantly different from those trained on in the TETP.

This is the first study to investigate the transfer of tone colour discrimination ability developed within a TETP to an external task that is sufficiently different from those encountered in the TETP, and the first to investigate the influence of the ICA and ISA training methods on this ability. For researchers seeking to establish the effectiveness of a TETP, this research showed the significance of collecting data that directly relates to the goals of the TETP, and crucially, on a task that is sufficiently different from those encountered in the TETP.

The research is a first step towards establishing the influence of the ICA and ISA training methods on students' ability to discriminate tone colour. Establishing which method is most effective has important implications for audio engineering curricula and the way students learn to operate equalisers.

ACKNOWLEDGEMENTS

I wish to thank my supervisor Bill Martens for his support and patience as I relocated across the globe time after time during my candidature.

I am grateful to David Major for donating his time and expertise in programming the Tonal Ear Training program.

Professional editor, Campbell Aitken, provided copyediting and proofreading services, according to the guidelines laid out in the university-endorsed *Australian standards for editing practice*.

David Moulton was most generous in providing information on the history and development of the Golden Ears training program.

I'd like to thank the students at the SAE Institute and the University of Sydney who participated in my research.

I am fortunate to have learned how to think about sound – albeit in very different ways – from two of the best in the business, Greg Simmons and Rick O'Neil.

I would not have completed this thesis without the support and encouragement of my family and friends, thank you.

PUBLICATIONS ARISING FROM THIS THESIS

Bassett, M. (2016, September). *Max as an interactive, multi-modal teaching tool for audio engineering*. Paper presented at the Audio Engineering Society Convention 141, Los Angeles, CA.

Bassett, M., & Martens, W. L. (2016, September). *SAE Parametric Equalizer Training: Development of a technical ear training program using Max*. Paper presented at the Audio Engineering Society Convention 141, Los Angeles, CA.

McKinnon-Bassett, M., & Martens, W. L. (2013, July). *Experimental comparison of two versions of a technical ear training program: Transfer of training on tone colour identification to a dissimilarity-rating task*. Paper presented at the Audio Engineering Society Conference: 50th International Conference: Audio Education, Murfreesboro, TN.

Martens, W. L., McKinnon-Bassett, M., & Cabrera, D. (2012, April). *Perceptual evaluation of stochastic-event-based percussive timbres for use in statistical sonification*. Paper presented at the Audio Engineering Society Convention 132, Budapest, Hungary.

McKinnon-Bassett, M., & Martens, W. L. (2012). Individual differences in the relative salience of percussive timbre dimensions. *Journal of the Acoustical Society of America, 131*(4), 3434-3434.

Martens, W. L., & McKinnon-Bassett, M. (2012). Relating timbre discrimination to perceptual distances between interpolated percussive timbres. *Journal of the Acoustical Society of America, 131*(4), 3434-3434.

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

List of Equations

ACRONYMS AND ABBREVIATIONS

AES	Audio Engineering Society
AET	audio engineering technology
AIM	Australian Institute of Music
AIP	Audio information processing
CD	Compact disc
DAW	Digital audio workstation
dB	Decibels
EQ	Equalisation
GUI	Graphical user interface
Hz	Hertz
ICA	identification by continuous adjustment
INDSCAL	individual differences scaling
IQR	interquartile range
ISA	identification by successive approximation
kbps	kilobits per second
LPC	linear predictive coding
LRA	logistic regression analysis
MDS	multi-dimensional scaling
MIDI	Musical Instrument Digital Interface
NVH	noise, vibration and harshness
PCM	Pulse-code modulation
PDA	personal digital assistant
PI	performance index
Q	Quality Factor

TET	technical ear training
TETP	technical ear training program
3AFC	three-alternative forced choice
2AFC	two-alternative forced choice
UK	United Kingdom
US	United States of America
VET	Vocation Education and Training

1 INTRODUCTION

This chapter provides an overview of the history of audio engineering training programs and introduces technical ear training (TET) within the context of critical listening. Common methods used to train students to use equalisers are presented and related to the aims and significance of the research.

1.1 A HISTORY OF TRAINING FOR AUDIO ENGINEERS

On June 2nd, 1946, composer Arnold Schoenberg wrote a letter to the then Chancellor of the University of Chicago, Robert Maynard Hutchins, in reply to Hutchins' question regarding what would constitute a suitable curriculum for a university music department. One of Schoenberg's suggestions was the addition of an Academy of Music to any proposed Department of Musicology; such an academy was to consist of only 'the main fields of practical musicianship and a number of *master classes*' (Schoenburg, 1987, p. 241). Amongst the many master classes Schoenberg suggested (including voice, composition, stage and conducting), was the 'education of "soundmen"', who 'will be trained in music, acoustics, physics, mechanics and related fields to a degree enabling them to control and improve the sonority of recordings, radio broadcasting and of sound films' (ibid.). Schoenberg suggested that such training of musicians in the 'mechanical fields' should equip them to 'correct acoustic shortcomings, as, for example, missing basses, unclear harmony, shrill high notes, etc.', adding 'this can be done and it would mean a great advantage over present methods where engineers have no idea of music and musicians have no idea of the technique of mechanics' (ibid.). Schoenberg's suggestion to combine audio engineering with traditional music curricula is the first recorded communication regarding what would later constitute a unique degree that would be embraced throughout Europe and is still offered today.

The following year, in 1947, the Nordwestdeutsche Musik-Akademie (Northwest German Music Academy) in Detmold, West Germany, commenced such a degree – the *Tonmeister* (Theinhaus, 1960). This degree was the world's first formal audio engineering course, which combined audio engineering curriculum 'with profound musical education', (Theinhaus, 1960, p. 68), culminating in a three-year Diplom-Tonmeister degree. Other tonmeister programs were soon offered in Berlin, Düsseldorf, Warsaw and Stockholm (Borwick, 1973). The German term *tonmeister*, literally translated as *sound master,* refers to an undergraduate qualification that combines a traditional Bachelor of Music degree with intensive studies and professional work placement in audio engineering, commonly resulting in the BMus (Tonmeister) award. In Australia, the SAE Institute has owned the trademark since 1988 (Australian Government, 2017b). In 1998 the University of Surrey in the United Kingdom (UK), trademarked the term in the UK (Trademark Direct, 2017) and currently offer a BMus (Hons) (Tonmeister) Music and Sound Recording degree that consists of a blend of 'rigorous musical study, advanced investigation of audio engineering and mastery of sound-recording operation and practice' (University of Surrey, 2016).

In 2013 nine universities worldwide offered the tonmeister degree, with all but one (McGill University in Montreal, Canada) located in Europe. Of these nine tonmeister programs, eight are offered within prestigious music schools; the exception, in Potsdam, is located within a school of film (Jopson, 2013). It should be noted that the tonmeister program is not exclusively offered at the undergraduate level; McGill University, for example, offers a graduate tonmeister program.

A distinction is made between the tonmeister and the *toningenieur*. The toningenieur (literally translated as sound engineer) is an electrical engineer who is concerned with 'problems and instruments having to do with sound frequency techniques' (Theinhaus, 1960, p. 69). Toningenieurs build, repair and test technical instruments including microphones, amplifiers, and loudspeakers and they are not required to possess any qualifications or experience in music (Theinhaus, 1960).

Prior to the 1960s, no formal audio qualifications existed in the United States of America (US) (Schmidt-Horning, 2013). Professional and aspiring audio engineers in the US and the UK seeking training in their craft were presented with two options: obtain employment within a company such as the BBC, Decca, Columbia Records or RCA Victor and receive in-company training; or receive on-the-job training gained through the process of working up from tea-boy or runner to tape-operator and ultimately in-house engineer within a commercial recording facility (Borwick, 1973). In the US, engineers hired by a record label were required to join the label's union and only union engineers were permitted to operate the equipment in the label's recording studios (Schmidt-Horning, 2013). If an artist chose to record in an independent studio not owned by a label, a union engineer from the artist's label was required to attend all studio sessions, but in many cases, he (this role was predominately staffed by males at the time) played no part in the process (ibid.). In the US, the unions 'had always advocated training' of audio personnel and efforts were made to promote the tonmeister-style degrees (ibid., p. 129). However, these efforts ultimately failed and labels were still reluctant to provide or support musical training for audio engineers right up to the late 1950s, citing such skills as irrelevant to the position (ibid.).

Borwick (1973) highlighted the paradox that was presented to job candidates in the US looking to work as audio engineers after World War II: in the absence of formal training programs with recognised qualifications, the only criterion on which to judge applicants' suitability for work in the industry was their practical experience in that industry, the very practical experience they were seeking to obtain.

In 2017, within the education and training sectors, the international marketplace for audio courses is teeming with offerings from government and private providers in addition to traditional sandstone institutions and online training providers. Accredited qualifications in audio are widely available at vocational, undergraduate and postgraduate levels, in addition to non-accredited certificates. In 2015, in the UK alone, 384 music technology courses were listed in the *Complete University Guide* (Fisher, 2015).

A well-designed audio engineering technology (AET) training program should prepare students for employment by providing a formal learning atmosphere that adequately trains them in the skills and competencies needed to enter the workforce (Bielmeier, 2013; Tough, 2009). As such, many of today's AET training providers regularly engage with industry stakeholders who provide input into curricula and advice as to which graduate attributes, learning outcomes and competencies they think are best suited to prepare students for placement in the industry (Commonwealth of Australia, 2011; Walzer, 2015).

In a study conducted in 2006, Tough (2009) surveyed 52 AET experts within the US, investigating the order in which they ranked 160 'essential competencies [that] need to be taught in an AET program 10 years from now to prepare students effectively for a career in the audio industry of the future' (p. 234). Participants were considered experts if they met at least two of the following criteria:

- had participated in the production of a minimum of five commercially available recordings or products that each sold 500,000 copies or more;
- had been in the audio business a minimum of 10 years;
- were considered by audio peers as an industry leader; and
- were future oriented, either inventing or accessing new technologies in their work. (Tough, 2009, p. 131)

The highest-ranking music-related competency, 'Identify instruments commonly used in commercial recording', ranked 54th overall, supporting the observation that musical training is not seen as relevant for audio engineers. However, in Lightner's (1993) survey of 154 audio professionals 'engaged in recording, production, or sound reinforcement in' an eight-state region of the United States, when hiring an entry-level employee or intern, employers agreed that either an associate degree or bachelor degree was a contributing factor (p. 28). The employers ranked 'music' as the most desirable field in which to hold a formal qualification, with 'sound' ranked second. In addition, respondents ranked 'musical skills' first and 'electronic repair' second when asked 'What job qualifications do you consider most important?' (ibid., p. 47). These conflicting results suggest that the 'prejudice against musical training for recording engineers' may have subsided by the early 1990s, but it may have returned some 10 years later (Schmidt-Horning, 2013, p. 130).

The highest-ranking 'general audio' competency in Tough's study was the ability to 'demonstrate a BASIC knowledge of effects including EQ [equalisation], reverbs, delays, gates, [and] limiters' (2009, p. 168). Critical listening underpins knowledge of these effects, however, the word 'listen' does not appear on the list of Broadcast and Sound Engineering Technicians' duties on the US Department of Labor Bureau of Statistics Broadcast website (Bureau of Labor Statistics, 2016). This omission is contrary to the fact that the operation of the signal processors listed in Tough's study, and any technical decision or assessment regarding sound, is based upon the user's ability to listen critically.

1.2 Critical Listening

The term *technical listening* was first used by Letowski (1985) in his seminal paper 'Development of technical listening skills: Timbre solfeggio' to refer to 'the process of timbre evaluation in terms of both aesthetic impressions and identification of their physical correlates' (p. 241). Letowski's definition of the term limited it to timbre evaluation and combined the analytical and critical listening domains. The term *technical listening* would later be used interchangeably with *critical listening*, thus removing any reference to the analytical/aesthetic nature of the sound.

At the most basic level, an audio practitioner must be able to discriminate in the technical domain (as opposed to the musical domain) between a sound pre- and post-processing, in order to ascertain the suitability of the processor applied. Although the ability to 'assess the quality of recordings using basic critical listening skills' was ranked 13^{th} in Tough's (2009) category of general audio competencies and 42nd overall, critical listening is embedded as a fundamental skill within the four-highest-ranking audio-related competencies:

1. demonstrate a BASIC knowledge of effects including EQ, reverbs, delays, gates, [and] limiters;

2. analyse BASIC audio signal flow in the recording studio;
3. engineer recording projects as an individual; and
4. demonstrate recording session procedures for tracking, overdubbing, and mixing sessions. (Tough, 2009, pp. 317-318)

The ability to listen critically is also a key component of the Tonmeister-style curriculum. Students seeking entry to the Northwest German Music Academy's Tonmeister program in 1960 were required to meet several prerequisites, including 'good grades' in German, music and natural sciences, prior knowledge of higher mathematics, advanced piano playing, and were to possess 'a good ear' (Theinhaus, 1960, p. 69). In addition, students had to pass a series of entrance exams in mathematics, physics and music, and were required to demonstrate 'proof of a critical ear'. In the explanation of the Tonmeister course (a version of which still runs today), the Northwest German Music Academy highlights that 'an essential prerequisite' of the important function of acting 'as interpreter between the artistic and technical fields' is 'an extremely well-trained ear with the ultimate goal of lifting the listening process out of the realm of the subjective onto the level of the objective' (ibid., p. 69). The ultimate 'duty of the *Tonmeister* [is] to… be a critical judge of the over-all impression of the result of the broadcast with respect to sound' (ibid.).

1.2.1 Critical Listening versus Analytical Listening

As part of their role, audio engineers are required to listen for a variety of different reasons and to be accurate and precise when doing so, with the accuracy of such judgements and the ability to repeat them paramount to performing the role effectively and efficiently (Everest, 1982a; Letowski, 1985; Moulton, 2005; Quesnel, 1996). Indeed, Moylan suggests that 'listening and paying attention' is the most difficult job of an audio engineer, who is required to listen actively; that is, to engage 'in seeking out information' (2007, p. 89). This multidimensional analysis of sound can be divided into two categories of listening: *analytical* and *critical*. Although both methods are used in the evaluation of sound and both are examples of *active listening*, each method seeks to gather different information.

Analytical listening is concerned with evaluating the artistic elements of sound, musical communication and the interrelationship of musical events and lyrics (Moylan, 2007). When listening analytically, the audio engineer is listening musically to elements of the performance such as pitch, rhythm and intonation. Examples of analytical listening tasks include listening to ascertain whether a musical performance is in time, whether a singer is in key, or whether a guitar is in tune. Analytical listening also encompasses the emotional message delivered through the various musical elements of the performance. A recorded performance may be technically flawless from an audio engineering perspective (for example, possessing no unwanted distortion, be balanced in spectrum and level, be recorded with an appropriately low noise floor and suitable dynamic range), but it may be unsuitable due to the lack of emotion, conviction or 'soul' present in the performance and thus fail to effectively communicate with the audience. Conversely, a recording may be considered technically inferior, yet still communicate with the audience on an emotional level.

Critical listening, however, is only concerned with the technical integrity of sound quality (the sound itself) as opposed to the music and any meaning it may convey. It is the process of 'evaluating the dimensions of the artistic elements of sound as perceived parameters – out of the context of the music' (Moylan, 2007, p. 90). This type of listening involves making sound quality judgements, assessing the technical integrity of the sound based upon perception. Whilst it is possible and indeed useful to measure sound from a technical standpoint, such 'objective measures do not always give a complete picture of how equipment will sound to human ears using musical signals' (Corey, 2010, pp. 7-8), and ultimately it is the listener's perception of the sound that such technical judgements are based upon. For example, the only way to determine the reliability of perceptual encoding algorithms that 'rely on psychoacoustic models to remove components of a sound recording that are deemed inaudible' is to listen to them critically (ibid., p. 16).

Critical listening is primarily concerned with defining the perceived physical parameters of sound (Moylan, 2007). Examples of critical listening tasks include listening to ascertain whether a mix is balanced in terms of levels (it is later argued that this is an analytical task also), whether the sound spectrum is balanced appropriately, or whether the instruments have any unwanted distortion. A recording may possess all the characteristics of a world-class musical performance and thus be effective in communicating the musical ideas, message and emotion to the listener, but be unsuitable as a production due to technical flaws such as an inappropriate amount of reverberation, excessive dynamic range, or lack of stereo imaging.

Although these two listening methods clearly differ in their intent, both are used throughout the audio production process and may overlap considerably. For example, in live sound production, after ensuring input levels are correct (not too low or distorting), the first step in the process of assembling a mix for the audience is the balancing of levels. At this stage the audio engineer will set the relative levels of each instrument over which they have control (snare drum, bass guitar, vocal etc.) to sound balanced to their ear. Ensuring that the guitars are not so loud that they cover up the vocalist for example, is a technical decision, but it is also a musical decision. The conductor of an orchestra has control over the level at which each group of instruments plays at throughout the performance – a musical decision, in which the conductor ensures that the musical elements are balanced in such a way as to communicate their interpretation of the score. A balanced mix in terms of levels is thus both a technical decision and a musical one.

The application of signal processors such as reverberation, equalisation and compression is a task firmly grounded within the realm of critical listening, but it may also impact on the musical performance. The addition of an inappropriate delay to a vocal track, for example, can cause an otherwise in-time performance to be perceived as out of time and render the vocal musically unsuitable. The addition of an inappropriate amount and type of reverberation to a vocal track may render it unintelligible, sabotaging communication from the musician to the audience. Signal processing is, however, ideally used to enhance the musical message. For example, the high frequency content in the reverberation applied to a vocal performance can be removed to create a sense of sadness and melancholy. The range of frequencies (bandwidth) of a mix or instrument can be reduced to create a feeling of nostalgia or for purely creative reasons. The application of signal processing has the potential to undermine the musical message as much as it does to support it.

Tonmeisters are trained to be experts in analytical listening and critical listening, both separately and in combination. Indeed, it is the ethos of the tonmeister that critical and analytical listening are inextricable elements of the audio production process. The revised Timbre Solfeggio (Miśkiewicz, 1992) ear training program (discussed in

Chapter 2) trains students to evaluate both technical and artistic elements of a recording.

1.2.2 Critical Listening in Audio Education

Critical listening underpins numerous roles and tasks within the field of audio engineering. Placing microphones, selecting preamps, processing audio signals with equalisation, dynamic processors and effects, and balancing levels all require the audio engineer to listen critically. From recording to mixing and mastering, all stages of audio production require a trained, discerning ear to perform each role effectively. It is simply impossible to appropriately place a microphone, equalise a drum kit, compress a vocal or balance a mix without listening critically.

In his 'Reviewer's Note' in response to Letowski's 1985 paper that detailed the Timbre Solfeggio technical ear training program (TETP), Eargle commends Letowski and his colleagues on their work in the field, noting that 'there should be more such studies in American curricula' (as cited in Letowski, 1985, p. 241). Walzer (2015) suggests that including critical listening within an audio engineering curriculum 'validates the need for relevant aural skills training that legitimizes emerging commercial music production sensibilities in higher education', but highlights the fact that 'no comprehensive and audio-centred rubric exists for evaluating critical listening skills in undergraduate music technology degrees' (p. 42, p. 49). Walzer recommends a systematic approach to designing a critical listening course, which 'guides students towards autonomy in making creative, aesthetic and technical decisions through relevant lab work, historical and social context and independent discovery' (ibid., p. 46). Schaller (2015) echoes the importance placed on critical listening in audio education, suggesting that in order to develop a portfolio and obtain an internship or employment, students must develop their listening skills ‘prior to graduation and entering the audio industry’ (p. 4). Failure to do so may result in the students not fully realising the ‘technical and artistic potential of their music productions and electronic compositions’ (ibid., p. 4). Furthermore, Thompson, Mosely and Ward (2013) suggest that critical listening can ‘provide a window on issues as diverse as digital audio, data compression, acoustic time and frequency domain phenomena, loudspeaker performance/positioning and psychoacoustic phenomena’ (p. 4).

While Schaller highlights the importance of developing listening skills prior to graduation and employment, Rakowski and Trybuła (1975) suggest that when selecting candidates for entry into a Tonmeister program such as that delivered at the Warsaw Academy of Music, ‘it is particularly important to check their abilities to differentiate the timbre of sound’ (p. 1). The authors assume that the difference limen for detecting distortions in a two-alternative forced choice (2AFC) test ‘is differentiated among people and is correlated with Tonmeisters [sic] professional skill’ (ibid., p. 3).

Employers also support the importance placed on critical listening by educators; they listed ‘demonstrate a BASIC knowledge of effects including EQ, reverbs, delays, gates, [and] limiters’ as the highest-ranking audio competency in Tough’s aforementioned study – a competency reliant upon critical listening. Employers, employees, and educators alike support the notion that it is imperative for critical listening to form part of any audio education program, regardless of qualification level.

Non-tonmeister-style audio engineering programs found throughout the US and Australia traditionally focus on developing critical listening skills but vary in their engagement with analytical listening, depending on the amount of music theory, aesthetics and music production content included in the curriculum. Educational institutions take two common approaches towards teaching critical listening within their diploma, undergraduate, and postgraduate degree programs: it is either delivered as a dedicated module (course) within a program (qualification), or critical listening content is embedded in modules throughout the program.

In early 2017 in Australia, dedicated critical listening modules were not delivered by any provider in either the vocational or higher education sectors, either public or private. There were three main private providers offering audio engineering education programs; the SAE Institute, the JMC Academy and the Australian Institute of Music (AIM). These providers all offered separate undergraduate and post-graduate degree-level and competency-based Vocational Education and Training[1] (VET)-level courses in audio. As there are no dedicated critical listening modules offered in any of these providers' undergraduate courses, training in critical listening is delivered by embedding it within the various modules of the degree. For example, critical listening may be a learning outcome within a module such as 'Signal Processing'. VET-level courses are made up of a varying number of standardised units of competency. The unit of competency 'Maintain and expand music knowledge and critical listening skills' was added in 2005 and was a component in five separate Certificate III and Certificate IV level awards (Australian Government, 2017c). However, when these awards were updated in 2010, this unit of competency was removed; it no longer forms part of the Certificate IV in Music Industry qualification which superseded the five aforementioned awards (Australian Government, 2017a). There are currently no units of competency that specifically address critical listening within the Certificate IV in Music Industry

[1] In Australia, VET is a competency-based education system that focuses on workplace-specific skills and knowledge, providing awards at the Certificate I, Certificate II, Certificate III, Certificate IV, Diploma and Advanced Diploma levels.

qualification. As such, critical listening must be embedded across multiple units of competency in a similar fashion to the way undergraduate programs are delivered.

In May 2017, the SAE Institute in Australia commenced offering postgraduate qualifications in which students can specialise in audio. Two of the electives on offer in this Masters of Creative Industries program are dedicated critical listening modules: 'Critical Listening for Audio Professionals A' and 'Critical Listening for Audio Professionals B' (SAE Institute, 2017). Psychoacoustic concepts are delivered alongside interactive audio demonstrations and TETPs in an exclusively online delivery format.

At the time of writing, there were no dedicated critical listening modules within any Australian undergraduate or postgraduate programs[2]. Universities also opt for embedding critical listening within modules in both their undergraduate and postgraduate degree programs, rather than delivering dedicated critical listening modules.

In early March 2017, the Audio Engineering Society (AES) website listed 294 educational institutions offering audio engineering programs in Canada, the US and the Americas, and 195 in Europe, the Middle East and Africa (Audio Engineering Society, 2017a). While a review of the curricula of these providers is beyond the scope of this thesis, student-focused TETPs are described in detail in Chapter 2.

1.3 Technical Ear Training

There is no question that audio engineers develop critical listening skills over time, as they are exposed to countless situations in which they are required to make

---

[2] Several providers offered critical listening modules that focused on music.

technical decisions regarding sound quality. Such sound quality assessments are made when audio engineers select and place microphones in relation to instruments and the recording space, and are developed 'casually' when they select mandatory devices such as pre-amplifiers and loudspeakers, and when they apply signal processing such as compression, delay and equalisation (Letowski, 1985). It follows that a person seeking to gain such practical experience in a shorter amount of time could effectively increase the frequency of their exposure to these critical listening scenarios through non-real-time exercises. An example of this is providing novice engineers or students with multitrack recordings for them to practise mixing, outside of the real-world commercial record production scenario. Indeed, it is common for studio interns to practise mixing after hours in commercial facilities in an effort to develop their skillset and experience. By using these offline exercises, novice audio engineers can expose themselves to critical listening scenarios at a rate much higher than normal at this stage of their career. In addition, by using non-real-time examples, they can focus on specific critical listening tasks such as equalisation in isolation, which would otherwise be impossible in commercial audio production scenarios, where an audio engineer must switch between critical/analytical listening tasks and is consistently under the pressure of time. From these offline exercises, dedicated TETPs were born.

Arguably, the first TETP was developed by David Moulton. It was a consumer-targeted program that would go on to be released as the *Golden Ears* training program in 1995 and become one of the most widely known and respected programs in the history of technical ear training (Moulton, 1993). However, the term *technical ear training* was first used by Quesnel and refers to any training program that develops a listener's technical (critical) listening skills (1990, p. ii). As such, technical ear training is an umbrella term that refers to any ear training related to audio production that does not train the listener in analytical listening or music theory. Technical ear training 'is a type of perceptual learning focused on timbral, dynamic, and spatial attributes of sound as they relate to audio recording and production' (Corey, 2010, p. 5), thus training programs such as Timbre Solfeggio (Letowski, 1985), Timbral Ear Training (Quesnel, 1990), Spectral Ear Training (Brixen, 1993), Golden Ears (Moulton, 1993), and Corey's (2010, 2017) suite of training programs are all examples of TETPs. An overview of TETPs is provided in Chapter 2, with a review of the literature provided in Chapter 3. Throughout this thesis, the term 'technical ear training' is used to refer to any critical listening training program.

The generic term *ear training* can refer to either technical ear training or musical ear training, thus a distinction must be made between the two. Solfège, a form of solmisation[3], is a music education method that assigns the seven notes of the Western musical scale a particular syllable. Although these two distinct types of ear training programs differ greatly, some TETPs use solfège techniques to identify technical parameters of sound, by assigning the sound of an English language vowel with an associated formant frequency (Letowski, 1985; Miśkiewicz, 1992; Quesnel, 2001). For example, the vowel 'ah' in the word *father* is associated with the frequency 1000 Hz (Quesnel, 2001).

[3] Solmisation is a music training method that associates each note on the musical scale with a syllable.

1.3.1 Categorising Technical Ear Training Programs

A wide variety of TETP currently exists; they range from consumer-targeted commercial programs offering only anecdotal statements from users and the occasional 'industry expert' to support their claims of effectiveness, through to programs based on empirical studies with robust experimental methodologies, designed to create expert listeners for in-house commercial testing.

1.3.1.1 Training Target

TETPs can most easily be categorised by the target user – the user that the program is designed to train. The three main categories of target user are student-targeted, employee-targeted, and consumer-targeted. There is obvious overlap between some of these categories; for example, a student enrolled in a degree program may undertake a student-targeted TETP as part of their studies, but might also use a consumer-targeted program to further refine their skills. It is also perfectly reasonable to expect a professional audio engineer to use a consumer-targeted program for the same reason. In addition, companies such as Harman (2011) have released versions of their in-house employee-targeted TETPs to the public.

Student-targeted programs are delivered as part of a structured educational program. Depending on where in the curriculum the TETP is delivered, the amount of previous audio engineering knowledge each student possesses can vary greatly. Some have almost no knowledge; others are degree-holding graduate students with several years' experience; others still are professional audio engineers with decades of experience but no formal qualifications. Most TETPs documented in the literature are delivered at undergraduate level; generally, these training programs form part of 'critical listening'-type courses that feature instruction (lectures, tutorials, etc.)

alongside the TETP. TETPs may be embedded within the curriculum as assessable items that contribute towards students' final grade, be competency-based checkpoints, or simply suggested activities that accompany lectures. The earliest documented student-targeted TETP, entitled *Timbre Solfeggio* (first reported by Rakowski and Trybuła, 1975), was delivered at the Chopin Academy of Music in Warsaw, Poland.

Employee-targeted programs are usually delivered in-house by companies looking to train their employees in several possible areas of audio engineering. Expert listeners who are not full-time employees (who work on a contractual or project basis) but are still trained by a company are included in this category for the purposes of this thesis.

There are numerous consumer-targeted TETP that vary in complexity and type, in a similar manner to employee and student-targeted programs. These products range from CD-based audio-only programs through to computer applications, mobile applications and digital audio workstation (DAW) plug-ins. They can be free, some feature buy-ins[4], and many make bold claims as to their effectiveness in order to promote the product; however, two commercial programs have proved seminal in the field.

Everest's *Manual for critical listening: An audio training course* was the first TETP to be commercially released (1982a, 1982b). The program, containing five cassettes and a 106-page booklet, would later be re-released over the course of two and a half decades in many incarnations, including the *Manual for Auditory Perception: An Audio Training Course* (1986); *Critical listening and auditory perception: The complete audio-visual*

---

[4] A buy-in or in-app purchase, is an option to purchase additional content or services within the app.

training course (1997); and the posthumous *Critical listening skills for audio professionals* (2007).

Moulton's *Golden Ears* training program was based upon a musical ear training program the author developed whilst working as a graduate teaching assistant in 1965 and was released commercially by KIQ Productions in 1993; it has been in print ever since, without modification. Numerous consumer-targeted TETPs have since been released, with many based upon the work of Everest and Moulton.

1.3.1.2 Training Type

Following the categorisation of TETPs by training target, programs can be further categorised by training type. Common types of TETPs include:

- loudness;
- pitch;
- frequency spectrum;
- spatial (reverberation, width, depth etc.);
- dynamics (volume automation, compressors, gates);
- perceptual encoding (audio data compression);
- distortion; and
- editing.

Training in loudness is generally presented using pairwise comparisons in which users are required to indicate the relative difference in loudness in decibels, or, via a simple 2AFC discrimination task in which the user is required to indicate whether they can hear a difference (alternative one), or not (alternative two), between two sounds presented separately. Pairwise comparisons such as this do not require listeners to possess any prior knowledge about the sounds, or to relate any perceived differences to physical properties – listeners are simply required to indicate whether they can perceive a difference, regardless of whether they are aware of any physical correlate. Of course, it is difficult to give feedback on this 2AFC task since the answer ‘I can’t hear a difference’ is correct when the listener can’t hear a difference. In contrast, the ABX, a variant of the three-alternative forced choice (3AFC) task described below, does have a correct answer, since X may be a match to A or B in level (and therefore, should be as loud as one and not the other presented sound).

In the 3AFC task, the listener is required to indicate whether the first presented sound is louder, quieter or identical to the second presented sound. Any preference the listener may have for one sound over the other, should they be able to discriminate between the sounds, is not recorded. Absolute identification of loudness is more difficult to implement due to the need to calibrate playback levels, and further complications such as noise floor also exist.

Training in pitch can either be presented using a 2AFC discrimination task, via pairwise comparisons of pure tone frequencies for which users are required to indicate the relative difference in pitch in Hz, or via absolute identification of the presented frequency (perceived pitch) in Hz. Absolute identification tasks such as these commonly limit the range of questions and responses to International Organization for Standardization (ISO) standard frequencies such as octave or 1/3rd octave bands.

In frequency spectrum-based TETPs, users are responding to physical changes in the frequency spectrum via perceptual changes in timbre and tone colour. In addition to 2AFC discrimination tasks, frequency spectrum-based TETPs – including Timbre Solfeggio (Letowski, 1985), Timbral Ear Training (Quesnel, 1990), Spectral Ear Training

(Brixen, 1993), components of Golden Ears (Moulton, 1993) and Corey's (2010, 2017) suite of training programs – can further be categorised into *matching*, *removing,* or *absolute identification* tasks.

In a typical matching task (Corey, 2010, 2013; Kaniwa et al., 2011; Quesnel, 2001; Quesnel & Woszczyk, 1994; Thompson et al., 2013), also termed *comparative listening* tasks (Quesnel, 1990; Quesnel & Woszczyk, 1994), the listener is presented with a reference sound (A), then a comparison sound (B). The listener must apply appropriate equalisation to (B) in order for it to be perceived as identical to (A). The listener may be permitted to switch between (A), (B) and a flat (unprocessed) sound (C) without processing. A simple matching task may require the listener to match the centre frequency of the equaliser applied to (A) – requiring the adjustment of a single parameter. The difficulty of such a task is partly dependent on the number of frequency options presented to the user. Novice users may start out with two-octave-bands, then progress to octave-bands and 1/3rd octave bands. In complex matching tasks, users may be required to identify more than one parameter of the applied equaliser, with more advanced listening tasks requiring the listener to identify the centre frequency, gain and Q-factor of the applied equaliser, for example. Combining multiple equalisers, expanding the number of options within each parameter and the addition of a time limit can all be used to add further difficulty to the task. The ability for the user to freely switch between the reference (A) and their response (B) is also an influencing factor on difficulty (Corey, 2010). Corey termed tasks that do not permit the user to hear (A) once any of the parameters of (B) have been modified as *matching memory* (2010) tasks. In this type of task, the user must commit the reference sound to memory and then match their response to this memory in a similar manner to the matching task.

Removing tasks (Kaniwa et al., 2011), also termed *bring to flat* (Quesnel & Woszczyk, 1994) and *return to flat* tasks (Corey, 2010) (also under the umbrella term comparative listening), are a modification of the matching task. As opposed to matching tasks, in which the user is required to duplicate the equalisation applied to the reference sound by processing the response sound, removing tasks require the user to remove the equalisation applied to the reference by processing the response so that it is identical to the flat sound. This is achieved by applying the inverse of the equalisation curve/s that was/were applied to the reference, thus removing (negating) the signal processing and returning the response sound 'to flat'. In this type of task, as with the matching task, the user can hear the modifications being made. As such, both matching and matching memory tasks are *active tasks*, as the user receives audible feedback of the modifications made using the equaliser (Letowski, 1985; Marui & Kamekawa, 2013; Miśkiewicz, 1992; Miśkiewicz & Letowski, 1999; Neher, Rumsey, & Brookes, 2002; Quesnel, 1990, 1996, 2001; Quesnel & Woszczyk, 1994).

Absolute identification tasks (Corey, 2010, 2013; Letowski & Letowski, 2011; Marui & Kamekawa, 2013; Quesnel, 1990, 1996, 2001; Quesnel & Woszczyk, 1994) differ from matching and matching memory tasks in that the user is not required to modify the sound. Instead, the user must identify the centre frequency of the equaliser that has been applied (or the frequency of a pure tone in the case of pitch training), for example. The listener may be able to audition a flat version of the signal. Absolute identification tasks can be increased in difficulty in the same manner as matching tasks. As there is no audible feedback provided, absolute identification tasks are passive tasks that are generally considered to be more difficult than active tasks, because the user must identify the centre frequency of the equaliser applied by comparing the presented sound to those in their long-term memory (Letowski, 1985; Marui & Kamekawa, 2013; Miśkiewicz, 1992; Miśkiewicz & Letowski, 1999; Quesnel, 1990, 1996, 2001; Quesnel & Woszczyk, 1994).

1.4 Equaliser Training Methods

Within audio engineering education, it is commonplace for equalisation to be the first signal process taught (excluding panning and gain). This is probably due to several factors, including the static nature of (non-dynamic) equalisers, the inherent simplicity of the concept and controls compared to dynamic processors, and the ubiquitous nature of equalisation throughout the field of audio production and consumer devices in general. Home hi-fi systems, personal listening devices and car stereos often feature some type of user-adjustable equalisation and it is not uncommon for commencing students in an audio engineering program with no previous audio experience to be familiar with concepts such as bass and treble. However, within audio engineering education, there are two conflicting schools of thought concerning the method by which students should be taught to learn to use equalisers. In this thesis, the two methods are termed the *Identification by Continuous Adjustment* (ICA) and *Identification by Successive Approximation* (ISA) methods. The term 'successive approximation' was first used in the context of technical ear training by Quesnel (2001, p. 88) when hypothesising about the method his students might use when performing a critical listening task.

The ICA or 'boost, search, set' method involves boosting the gain on one band of a parametric equaliser by a significant amount (approximately 10dB of gain or more) using a fairly low Q at an arbitrary centre frequency (Case, 2007, p. 117). This boost is then 'swept' (moved) throughout the frequency spectrum with the aim of locating the desired frequency. Once the frequency has been aurally identified, the gain and Q are adjusted. Katz explains:

> The classic approach is to focus the equalizer directly: starting with a large boost and fairly wide (low value) Q, sweep through the frequencies until the resonance is most exaggerated, then narrow the Q to be surgical, and finally dip the EQ the amount desired. (Katz, 2008, p. 140)

The ISA method, on the other hand, requires the user to set the frequency parameter of the equaliser by approximating the desired frequency first, then set the Q and gain parameters in a similar fashion to the ICA method with the equaliser bypassed. The user then engages the equaliser and assesses whether the approximated frequency was correct. If not, the equaliser is immediately bypassed and the user again approximates the frequency of concern based on the sound of the previous approximation. For example, the user may conclude that the previously boosted frequency was too low on the spectrum and the subsequent approximation will need to be higher on the spectrum. This process is repeated until the desired frequency is located.

In addition to the schism regarding the most effective method of training, the literature is lacking the area. While several frequency spectrum-based TETPs have been developed, there are currently no published studies examining the effectiveness of either training method.

1.5 Previous Research

To establish the effectiveness of a TETP, one needs to establish whether the goals of the training program have been met. As such, the goals of a TETP must be clearly stated at the outset and valid and reliable tests that can measure whether the goals have been met are required. The predominant method used by researchers to establish whether the goals of their TETPs have been met is to measure students' performance on the training program itself at the commencement of the training period, then measure students' performance over time, and/or, measure the students' performance at the end of the training period. In this sense, a TETP is deemed effective in meeting its prescribed goals if students' performance on the training program simply improves over time. As discussed in the literature review, in general, students' performance on a TETP is likely to improve over time; the more a student practises with a TETP, the more their performance improves, up to a point. When students' performance at the beginning of a training period is compared to their performance at the end, most TETPs show an improvement in performance; however, it should be noted that performance on any task will generally improve over time, regardless of whether additional training is provided. Many authors of TETPs conclude that a given training program is effective based on the demonstrated improvement in performance on the training program over time, rather than on some related but distinct task. For example, Kawahara et al. (2013) suggested that their TETP was effective based on observations that the average percentage of correct responses of students increased over time and the standard deviation of correct responses decreased over time, concluding that 'these results show the effectiveness of Technical Listening Training' (p. 4). These goals, which vary between training programs from the general to the specific, and the methods by which researchers conclude that they have been met, are investigated in the literature review.

In the majority of published studies presenting student-targeted TETP performance data, the training task itself, or a closely-related task utilising the same practised skills as the training task, is used as the measurement tool to ascertain whether the goals of the training program have been met (Kaniwa et al., 2011; Kawahara et al., 2013, p. 1; Liu, Wu, & Yang, 2007a; Quesnel & Woszczyk, 1994; Rościszewska, 2011; Rościszewska & Miśkiewicz, 2014). In these studies, an improvement in students' performance on the training program over time, or an improvement on a closely-related task (often delivered as a pre/post-training test), is used as evidence of the training program's effectiveness. Although these TETPs may state that they have specific goals, such as improving:

- sensitivity to timbral changes (Quesnel & Woszczyk, 1994; Rościszewska, 2011);
- 'discrimination of sound attributes' (Liu et al., 2007a, p. 1);
- 'the ability to systematically discriminate and identify sonic differences' (Kaniwa et al., 2011, pp. Introduction, Para 1);
- 'memory for timbre and develop[ing] the ability of associating the perceived characteristics of timbre with the spectral properties of sounds' (Rościszewska & Miśkiewicz, 2015, p. 1); and
- the ability to 'discriminate between different sounds' (Kawahara et al., 2013, p. 1),

performance relating to the training goals is not directly measured; rather, the goals are assumed to have been met if students simply show an improvement in performance in a TETP (or a closely related task). Quesnel concluded that his TETP was effective because students that trained using his program outperformed professional engineers on a critical listening test. Quesnel's study, which compared the performance of students on a critical listening test to that of professional engineers, is one of the most frequently cited in the literature. Quesnel (2001) argued that the TETP was effective as it had trained students on a task that 'did involve listening skills that professional audio engineers, in the absence of specific training, develop throughout their career' (pp. 69-70). Therefore, Quesnel hypothesised that the professional engineers would outperform students on the test. However, as discussed in the literature review (pp. 127-132), there are several reasons that support the prediction that students would outperform professional engineers on the test.

1.6 JUSTIFICATION

This thesis presents the argument that the most legitimate test of whether a TETP's goal has been met, such as 'improving memory for timbre', will use in that test neither the TETP employed to train students or a closely-related task. A TETP is only useful as a test if seeking to show whether students have improved their ability to undertake the TETP itself. However, a demonstrated improvement in performance on a task within a TETP, or a closely related task, does not necessarily equate to an improved memory for timbre, for example. Contrary to the prevailing paradigm in TETP research, the purpose of a TETP is not to simply improve one's performance on the training program itself, as improvement in performance on a TETP does not necessarily provide the best evidence that the training program's goals have been met.

Several researchers in the field acknowledge this argument and advocate for an external, independent test of skill transfer outside of a TETP. In the earliest published paper detailing the Timbre Solfeggio TETP, Rakowski and Trybuła (1975) suggest a positive correlation exists between performance on a critical listening entrance exam and students' subsequent 'efficiency as recording specialists', and assume a potential student's ability to detect distortions in a 2AFC test is positively correlated to their professional skill (p. 3). However, the authors caution that 'to prove such an assumption it would be necessary to estimate the validity of the test using a well-established criterion of Tonmeisters [sic] professional efficiency' (ibid., p. 3). Quesnel (1990) also suggested that 'a true measure of the validity of such a [technical ear training] system is achieved only if improvements in the exercises can also be measured in the recording studio' (p. 105). In reference to training participants for professional listening evaluations in the automobile industry, Shively and House (1998) also suggested that based on their experience, listeners achieving consistently high scores on a resonance training task 'does not mean that they are well trained for listening evaluation purposes' (p. 2). Indeed, the main objective of their study was to verify that any perceptual learning that occurred transferred outside of the training scenario to a 'listening evaluation or experiment' (ibid., p. 2).

Furthermore, the assertion that performance on the TETP is insufficient to indicate whether the goals of a TETP have been met is supported by several authors (Indelicato, Hochgraf, & Kim, 2014; Kim, 2015; Liu et al., 2007a; Quesnel & Woszczyk, 1994; Rościszewska, 2011; Rościszewska & Miśkiewicz, 2014, 2015). These authors aim to establish the effectiveness of their TETPs by looking for evidence of the transfer of developed critical listening skills outside of the TETP itself. However, as detailed in the literature review, there are no published studies that seek to establish the effectiveness of student-targeted TETPs using evidence other than an improvement on the training program itself, or a similar related task.

1.7 Aims

The aim of the research presented in this thesis was to establish the relative importance of training method on tone colour discrimination – whether 'the development of standardized audio training software benefits music technology students in measurable ways' (Walzer, 2015, p. 48). This aim required the development of a suitable pre/post-training tone colour discrimination test to ascertain whether the ICA or ISA training method is more effective in training students when using a TETP. If the research showed a difference in students' ability to discriminate tone colour based on training method, this would provide evidence of the training method's potential to develop students' critical listening skills. If the research showed a difference in students' ability to discriminate tone colour in general, regardless of training method, this would provide evidence of the transfer of developed critical listening skills outside of the TETP itself, as the pre/post-training tests for tone colour discrimination differ significantly from the tasks undertaken in either training method. The transfer of developed critical listening skills outside of the TETP, as argued in the literature review, is a requirement for establishing the effectiveness of a TETP.

1.8 Overview of Methods

This research project employed an empirical research methodology designed to develop a suitable test to establish the relative effectiveness of the ICA and ISA training methods delivered within a TETP, involving measuring students' performance on a task that is unrelated to the TETP itself. To achieve this, several tests designed to measure a student's ability to discriminate changes in tone colour were developed. The sensitivity of these tone colour discrimination tasks was then tested in isolation, outside of a TETP, via a series of preliminary studies. The tasks were deployed in a pilot study involving first-year undergraduate audio engineering students using a pre/post-training test methodology.

The TETP developed for the pilot study featured an absolute identification task in which students were required to identify the centre frequency of a parametric equaliser boost applied to pink noise. An ICA version of the TETP was administered to half the students (randomly selected) taking part in the study; these students could hear continuous changes in the spectrum resulting from adjusting the on-screen frequency selector. The other students were administered an ISA version of the TETP that muted the audio output whenever the frequency selector was adjusted, until the selector was released; as such, the ISA-trained students were unable to hear the boosted equalisation curve move throughout the spectrum. Aside from this, the training received by both ICA and ISA groups was identical.

Three tone colour discrimination tasks were developed for the pilot study. Students were asked to complete the tasks prior to the commencement of the TETP and immediately following it. The first task presented recordings of several small pellets being shaken inside two different-sized plastic containers. Students were required to estimate the number of pellets in each container via an on-screen slider. The second task required students to rate the similarity, on a scale of 0 to 10, of 72 pairs of synthetically-generated conga strike tones. The third task required students to categorise artificially generated vowel sounds on two separate formant frequency continua. The pilot study confirmed that the conga similarity-rating task had appropriate sensitivity and dynamic range to detect differences in performance between training groups when delivered in the context of a TETP, and was subsequently chosen to be deployed in the full-scale study.

The full-scale study featured a TETP that required students to identify the centre frequency of a parametric equaliser boost applied to various genres of music and pink noise. Prior to undertaking the training, students were again divided into two groups, with each group undertaking ICA and ISA versions of the training program respectively. In the full-scale study, students undertook a tone colour discrimination test before, at multiple stages during, and after the TETP.

Although multiple TETPs were developed as part of this study, and the literature review investigated how researchers established the effectiveness of their TETPs, this research project was not concerned with students' performance on TETPs. This was because it is the author's assertion that a demonstrated improvement in performance on a TETP is not suitable evidence of the effectiveness of the training program.

The next chapter presents a descriptive overview of a selection of the frequency spectrum-based TETPs developed to date. Chapter 3 presents a review of the literature, beginning with a review of the ICA and ISA training methods in the context of signal processing devices, audio engineering education and TETPs. The chapter then reviews the methods by which authors establish the effectiveness of their TETPs, using a set of criteria developed to establish the minimum requirements for TETP goals. Finally, the pre/post-training test methodology is reviewed. Chapters 4, 5 and 6 detail early work into the development of suitable tests of tone colour sensitivity, a pilot study that investigates the effectiveness of these tests in context, and the full-scale study, respectively. Chapter 7 details the TETPs developed as part of the research project. Chapter 8 summarises the research and outlines areas of future work in the field, before presenting the limitations of the research and finally, the contribution of the research to the field.

2 Technical Ear Training Programs – A Review

This chapter offers a descriptive overview of a selection of the frequency spectrum-based TETPs developed to date. For TETPs that contain multiple types of training, only the frequency spectrum-based training is detailed. The chapter is divided into two sections based on training target categories. Numerous TETPs have been omitted from this overview, as only student and consumer-targeted programs for which published papers or texts are available are included.

2.1 Student-Targeted Training Programs

2.1.1 The Teaching Machine

Greenberg and Huddleston (1971) detail the earliest documented occurrence of a student-targeted TETP in the literature. The authors describe a syllabus that was written for music students, part of which featured a mechanical 'videosonic machine' that projected slides and simultaneously played back audio. The machine, with its limited frequency response of 80-8k Hz, was designed with several goals in mind, one of which was to aid music students 'to develop an awareness of instrumental tone color as an important element of music' (ibid., p. 57). The TETP consisted of audiovisual materials that introduced students first to instrument families, then to the tone colours of individual instruments. Training featured questions in multiple-choice format. The machine automatically paused the presentation of the audiovisual material after each question in order for students to respond. If an incorrect response was recorded, a red light was triggered and the student was given a second chance to respond. The machine did not continue to the next question until the correct answer was selected. Correct responses were reinforced 'through the conversational-chaining technique of programming' (Greenberg & Huddleston, 1971, p. 59), 'a type of non-branching programme… in which the desired answer that will fill in the blank space in the frame occurs at the same point in frame where it is emphasised in the text' ('Conversational chaining', 1978). 'A pre-test, post-test and post-test-for-retention-of-knowledge' (Greenberg & Huddleston, 1971, p. 59) was proposed to evaluate the training program, however no details were published.

2.1.2 Timbre Solfeggio / Solfège

The most widely published upon student-focused TETP was first delivered at the Chopin Academy of Music in Warsaw, Poland in 1974 and variations of the course continue to be delivered today at several institutions (Letowski, 1985). Utilising the Italian word *solfeggio* – the name of a music education method used to teach memory for pitch – the Timbre Solfeggio TETP, first reported by Rakowski and Trybuła (1975), was developed by Andrzej Rakowski, Krzysztof Szlifirski, and Tomas Letowski to accompany the traditional music solfeggio course the creators also taught (Letowski, 1985). It was embedded within the academy's Tonmeister program and was designed to train and improve students' sensitivity to relative differences in timbre and their memory for absolute timbral references. Letowski (1985) later adds that a secondary goal of the Timbre Solfeggio program was to 'develop a set of qualitative descriptors to provide unequivocal codes (language) for exchanging information regarding timbre impressions' (p. 241). This early iteration of the program consisted 'of a series of [mono] listening sessions devoted to various acoustic phenomena' that were categorised into three groups:

1. linear distortions;
2. nonlinear distortions; and

3. evaluation of absolute sound pressure level, changes in level (as a result of dynamic processing), and signal to noise ratios. (Rakowski & Trybuła, 1975, p. 3)

The linear distortion category of ‘problems’ featured training in: the timbre of white noise, coloured noises, and linearly distorted (processed) speech; intelligibility of band-limited speech and the recognition of high- and low-pass filters; the effect of filters as applied to singing and musical instruments (specifically the suppression and reinforcement of formants); bandpass limiting; and the relationship between timbre and level (Rakowski & Trybuła, 1975). The nonlinear distortion category provided training in pure and complex tones, ‘coloured noises subjected to nonlinear distortions of various order’, ‘the effect of nonlinear distortions on the transmission of music’, and the threshold of detectability of ‘distortions in various musical sounds and in speech’ (ibid., p. 3).

Nine years later, Letowski (1985) reported on an updated version of the TETP, delivered at the same institution over six semesters. The updated course featured training in the memorisation of linear and nonlinear distortions, musical dynamics, sound decay and attack times and reverberation. Both ‘active’ and ‘passive’ tasks were used during the course. The active tasks required students to modify an unprocessed signal so that it matched a given processed signal, reportedly providing a ‘general orientation to the relationship between timbre changes and the physical transformations used to obtain them’ (ibid., p. 241). The passive tasks required students ‘to identify the basic timbre-related features of the sound’ (ibid.). These passive tasks were designed to develop students’ ability to detect and verbally describe the differences between two timbres, and to ’ when a sound was presented in isolation (without a reference) (ibid.). This revised version of the program initially limited the amount of signal processing and the variety of stimuli used, gradually increasing these over time.

Letowski (1985) suggested that the development of both timbre sensitivity and timbre memory relies heavily upon 'the establishment of an accepted code of terms used to describe timbre phenomena' (p. 242). Based on this, the revised program used a set of 'timbre categories' that Letowski termed 'formants', which he claimed correlated with the perceptual experience associated with the 'prevailing presence' of one of nine ISO centre frequencies (ibid., p. 242). As the training progressed, this set of nine standards (timbre categories) was expanded to include 27 1/3rd octave ISO centre frequencies.

In addition to utilising a limited set of timbre categories, the initial training featured the use of descriptive adjectives and vowel associations. Students were trained to associate standard ISO frequencies with descriptive terms such as 'sharp' and 'dark'; 'sharp', for example, was associated with a boost at 4k Hz. Letowski (1985) also suggested that the use of this vowel-likenesses 'simplifies the initial training in timbre solfeggio' (p. 242).

Initially, the main type of signal processing used with the nine timbre standards is band-pass filtering, applied to 'complex harmonic multi-tones and wideband noises' (Letowski, 1985, p. 243). Low- and high-pass filters, speech, and music are introduced later in the course, equalisers are eventually substituted for filters, and the number of timbre categories increase. Early in the training, only single modifications of the sound are made, but this later increases up to four simultaneous modifications. The components of the modification that students must identify in the two-semester basic program include the frequency of the filter/equaliser, the bandwidth and the upper and lower frequency limits of the signal. In semesters three to six, the intermediate and advanced courses require identification of the natural resonant frequencies of musical instruments (again termed ‘formants’), their relative strength and bandwidth, the audibility of distortions in lab-based and commercial recordings, and ascertaining the number of performers and their relative locations in small and large groups. The program was generally carried out in mono using a single set of loudspeakers, but other types of loudspeakers and headphones were also used at times.

A revised version of the Timbre Solfeggio TETP was implemented seven years after the 1985 Letowski paper (Hojan & Rakowski, 1999; Letowski & Miśkiewicz, 1995; Miśkiewicz, 1992). This revised program features similar training in many of the areas detailed in the earlier papers, with several new training areas:

- speech perception and sound evaluation;
- masking;
- ‘evaluation of the quality of musical recordings’; and
- ‘perceptual assessment of audio equipment’. (Miśkiewicz, 1992, p. 621)

In addition to practical training, students receive complementary theoretical instruction. This instruction covers fundamental psychoacoustic concepts including the perception of speech, music and recorded sound. Interestingly, Miśkiewicz uses the term ‘formants’ interchangeably with ‘frequency’, and uses the term ‘antiformants’ to describe the attenuation of a frequency band.

During the first year of the program, the syllabus focuses on timbre-based training using monophonic sounds. Both active and passive tasks are used, as in the Letowski’s (1985) program. Miśkiewicz (1992) suggests that the active training tasks ‘resemble real situations encountered by the sound engineer’ and cites this as the reason the active tasks are ‘well received by students’ (p. 623).

The second year of the training program focuses on ‘loudness perception and its relationship to timbre’ (Miśkiewicz, 1992, p. 623). Students are trained on loudness scaling, discrimination, memory, and summation, the effects of loudness on timbre, preferred levels of music, and speech, and preferred balance of music and speech. Students are trained to identify musical dynamic levels, that is, the manner in which the timbre of an instrument changes depending on the playing level (pianissimo versus fortissimo, for example). The pitch of complex tones was another addition to the training program, in which students were required to examine how changes in spectrum affect pitch and to assess the pitch strength of musical tones. Students are trained to identify masking, most commonly via the detection of nonlinear distortions.

Audio quality assessments of equipment and music recordings are given to students to refine their timbre description skills. These tasks include ‘semantic description and scaling of timbre’ and are introduced along with Pierre Schaeffer’s concept of the ‘sound object’ (Miśkiewicz, 1992, p. 623). The equipment that is evaluated includes several models of ‘loudspeakers, headphones, amplifiers and CD players’ (1992, p. 623). The last part of training in the second year focuses on the evaluation of the sound quality of various genres of music recordings via subjective assessments of properties including artistic elements and the overall quality of the production.

The third year of the program focuses on training students in spatial perception and awareness. Students are initially trained in horizontal localisation tasks in which they must identify the inter-channel sound pressure ratio and delay as a function of azimuth.

Active training tasks were used to train students in detecting changes in the timbre of reverberant sound via equalisation applied to reverberation that is mixed back in with the original signal. Finally, students are trained to make estimates of the 'amount of the perceptual attribute called auditory spaciousness' via judgements of width, depth and height (Miśkiewicz, 1992, p. 624). Overall, the program lasts 150 hours (Hojan & Rakowski, 1999).

The Timbre Solfège program has been adapted beyond training students of sound engineering, initially in the audio industry then the automotive industry (Miśkiewicz & Letowski, 2014). In 1994, such an industry training course was offered in the US 'with focus on perceptual capabilities of car designers and members of sound quality jury panels' (ibid., p. 1). In 'Psychoacoustics in the Automotive Industry' (2014) and '*Timbre Solfege* training in automotive industry' (1999), Miśkiewicz & Letowski (2014) detail the essential elements of an ear training course for the automotive industry, with specific focus on the adaptation of the Timbre Solfège course for this purpose. The modified program was designed to develop the following skills;

- the ability to describe the perceived characteristics of sound in an non-ambiguous manner, and communicate the description in a meaningful way to other persons involved in the sound design and development process;
- the ability to evaluate the characteristics of sounds quantitatively; and
- the ability to link individual perceived characteristics of sound to the sound's physical characteristics. (ibid., p. 2)

It was reported that these goals were achieved through theoretical lectures, demonstrations of processed sounds, and listening drills. Participants were trained in both the assessment of sound quality and sound character; the latter was achieved using ‘a clear and simple terminology, a kind of verbal code’ that the noise, vibration and harshness (NVH) engineers could use to communicate to each other and other stakeholders. In addition to qualitative descriptors, participants were trained in the quantitative evaluation of automotive sounds by rating ‘an attribute of sound on a closed scale, defined by a pair of bipolar descriptive terms’ (Miśkiewicz & Letowski, 2014, p. 3). This was achieved by training with paired comparisons, scaling tasks and absolute magnitude estimation tasks (ibid.).

As with previous iterations of the Timbre Solfège course, Miśkiewicz & Letowski (1999) refer to passive and active training with formants (timbre categories) and the use of vowels to memorise these timbre categories. The authors state that the course ‘has fulfilled its main goal of developing the listening skills of NVH engineers and prepared them for leading roles in the sound evaluation panels’, based on ‘several years of teaching the course’, ‘the results of practice tests conducted during the courses, formal and informal course evaluations by its participants, and the comments received from the participants years after the end of the training’ (Miśkiewicz & Letowski, 2014, pp. 4-5). No empirical data was provided to support these anecdotal claims of effectiveness.

2.1.3 Timbral Ear Trainer I & II

Quesnel (1990) first detailed his TETP and resulting performance data in his master’s thesis titled *A computer-assisted program in timbral ear training – a preliminary study*. The program, a partial implementation of the Timbre Solfeggio program, was designed ‘for the improvement and maintenance of the aural skills that are essential to audio engineers’ (ibid., p. 74). In a later paper, Quesnel and Woszczyk (1994) highlight that Timbral Ear Trainer I has three main goals:

1. developing long-term memory for a basic set of timbral categories;
2. progressively developing the listener's sensitivity to subtle timbre differences; and
3. developing analytic listening skills for the evaluation of complex spectral changes in terms of individual resonances, specifying for each the magnitude, Q, and center frequency. (p. 2)

In the TETP, signals are played from CD and modified using a MIDI[5]-controlled hardware digital parametric equaliser, with all listening done through headphones. Various types of noise (white, pink), speech, solo instruments and orchestral recordings are used as stimuli. Presentation of exercises and evaluation of answers is achieved using software written in C and Assembly (Figures 1–5). Feedback to users consists of indications of correctness ('right' or 'wrong').

The training program includes two types of exercises: comparative listening exercises consisting of paired comparisons, and absolute identification exercises consisting of single judgements. In the comparative listening exercises, the users are presented with two sounds, A and B, that the user could switch between. Sound A is spectrally modified with the nature of the modification unknown to the user, and Sound B is presented unmodified. The goal of the matching task is to modify the spectrum of Sound B so that it perceptually matches Sound A. The parametric equaliser parameters available for modification are centre frequency, Q and gain.

[5] Music Instrument Digital Interface

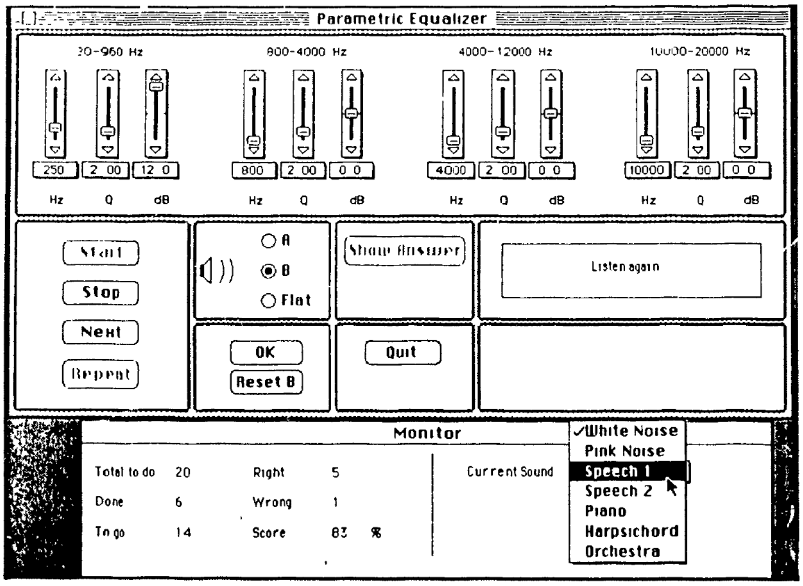

Figure 1. The first version of the Timbral Ear Trainer user interface for comparative listening exercises (Quesnel, 1990, p. 94)

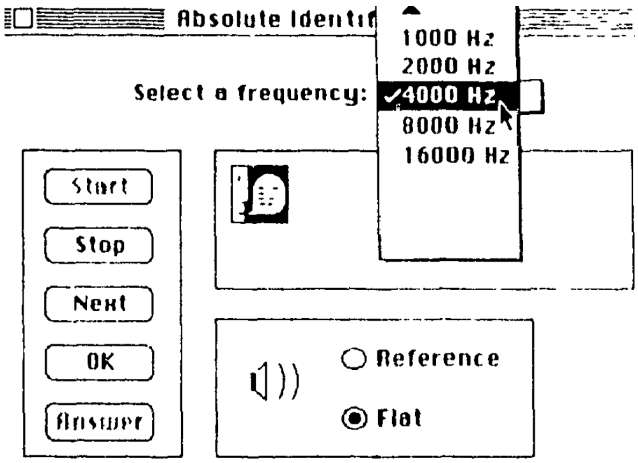

Figure 2. The first version of the Timbral Ear Trainer – user interface for absolute identification exercises (Quesnel, 1990, p. 95)

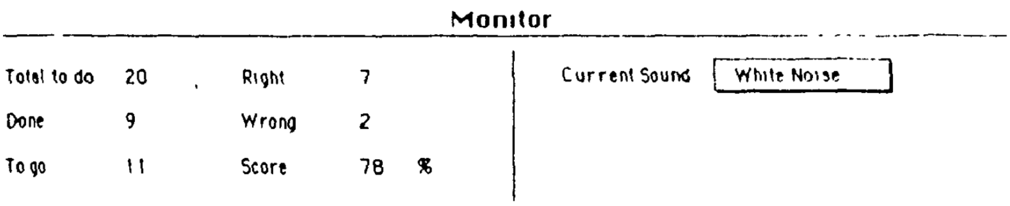

Figure 3. The first version of the Timbral Ear Trainer – monitor window (Quesnel, 1990, p. 95)

In a modified version of the comparative listening exercise, the user is presented with three sounds. Sounds A and B are spectrally modified and identical, while Sound C is presented unmodified. The user is required to modify the spectrum of Sound B so that it matches Sound C; essentially the user is required to perform a removing task by applying the inverse of the original processing applied to Sound B. Quesnel and Woszczyk (1994) later referred to this type of task as 'bring to flat' (p. 3).

The frequency spectrum is divided into four overlapping bands with one equaliser available for each band, allowing the user to modify the centre frequency, Q, and gain of each given band. Feedback is provided via the software interface showing the number of completed, correct, and incorrect exercises.

In the absolute identification exercises, the user is presented with a single sound that has been spectrally modified. The user is required to identify the specifics of the parameters that had been modified. The choice of auditioning an unprocessed version of the sound is included when the desire is to make the task easier. Following from the Timbre Solfeggio course, Quesnel promotes the use of vowel sounds as 'timbre categories'; associations between the centre frequency of resonant boosts and vowel sounds were used in the training program. A variant of the absolute identification task 'requires[d] the students to answer using vowel categories instead of using center frequency values' (Quesnel & Woszczyk, 1994, p. 4).

The Timbral Ear Trainer program is organised into '24 units of increasing complexity' initially requiring the user to identify the parameters of a single resonance, then moving in complexity to three and four resonances (Quesnel & Woszczyk, 1994, p. 4). The training is supplemented by weekly tutorial sessions with an instructor 'to provide additional help to suggest listening strategies for answering questions' (ibid.).

Two modes of interaction are available for both the comparative listening and absolute identification exercises. ‘Free mode’ consists of a random presentation of all available questions with no limit on the number of questions. In the ‘fixed mode’, a set number of problems must be completed before moving on to the next group of exercises.

Nine Sound Recording graduate students at McGill University took part in a study of the effectiveness of the training program over six months. Four students were in their second year of study and the remaining five were in their first year. The recommended training regime was ‘two periods of 45–60 minutes per week… for a total of 24 weeks’, although session durations ranged between ‘a few minutes to 2 hours’, and most users’ session durations ranged between 30 and 60 minutes (Quesnel, 1990, pp. 96–97). Performance on the training task was compared at the beginning and the end of the six-month training period for three listening exercises: comparative listening, ‘bring to flat’ and absolute identification.

For the comparative listening test, frequency was the only parameter that was varied. Frequency modifications could be applied at any of the ISO 1/3rd octave centre frequencies between 125 Hz and 16k Hz (Quesnel & Woszczyk, 1994). Ten questions were asked. For the pre-test, white noise was used as a stimulus, for the post-test, a snare drum sound was used. For the ‘bring to flat’ listening test, filter parameters were the same as the comparative listening tests, but frequency modifications could be applied at any of the ISO 1/3-octave centre frequencies between 125 Hz and 8k Hz. Two filters were applied (boosts or cuts or both) to a cello sound. For the absolute identification tasks, filter parameters were the same as the comparative listening tests, with only one filter applied to a looped snare drum sound.

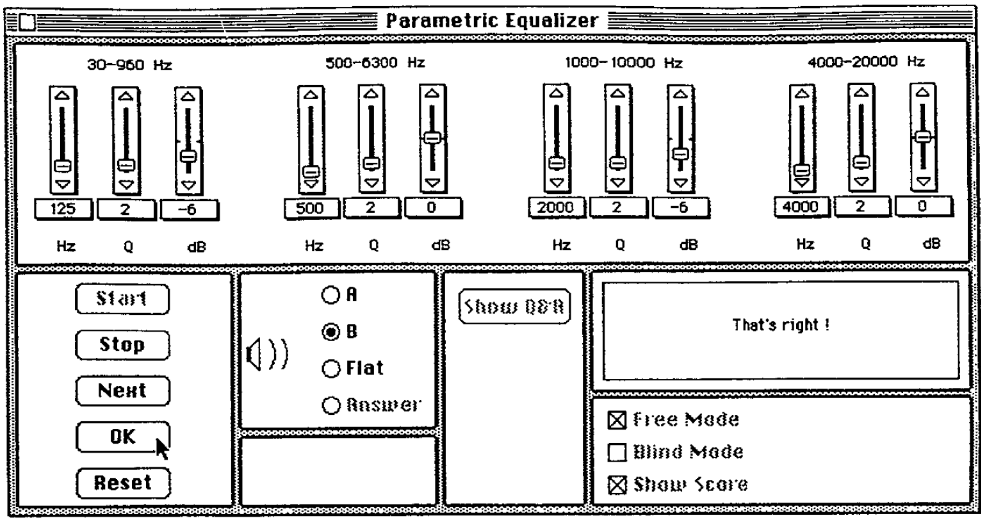

Figure 4. A revised version of the Timbral Ear Trainer user interface for comparative listening exercises (Quesnel & Woszczyk, 1994, p. 10)

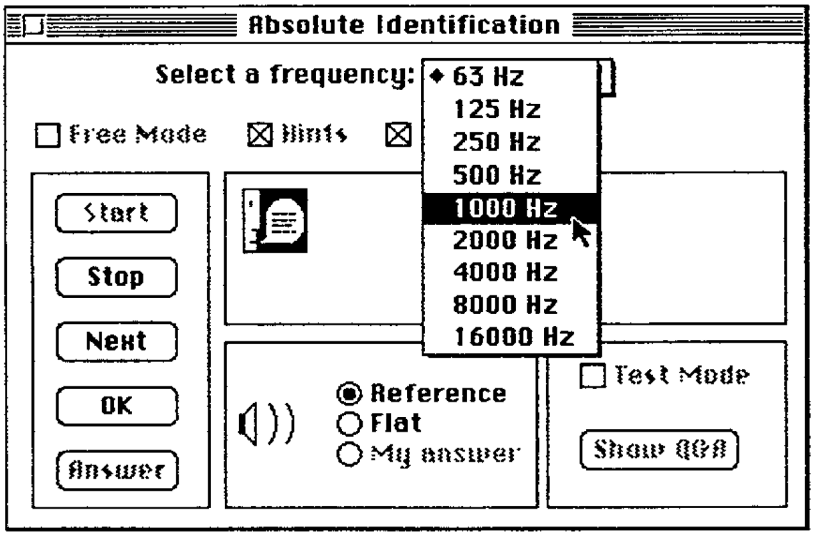

Figure 5. A revised version of the Timbral Ear Trainer user interface for absolute identification exercises (Quesnel & Woszczyk, 1994, p. 10)

Quesnel (1996) presents an updated version of the Timbral Ear Trainer program (Figures 6–10) featuring updated listening exercises and an 'adaptive' software feature. The comparative listening tasks (matching and 'bring to flat') and absolute identification tasks are again included, however the updated version contains several new listening tasks:

- ‘bandwidth evaluation’ tasks require users to duplicate filter settings (in a similar manner to matching tasks) in order to match the bandwidth of the presented sound;
- ‘similarity judgment’ tasks require users to determine if sounds presented pairwise are the same; and
- ‘vowel identification’ tasks require the user to select ‘a vowel that best correspond[s] to the resonance presented by the computer’. (Quesnel, 1996)

The revised version of Timbral Ear Trainer is adaptive, changing the ‘difficulty level of the tasks, the content of the exercises, and the feedback the user receives [based on] individual levels of performance’ (Quesnel, 1996, 3. Adaptive Training, para. 1).

The software develops listener profiles and monitors users’ progress, storing timing data and ‘information about each user's problem solving methods’ (ibid., 3.2 User Monitoring, para. 1).

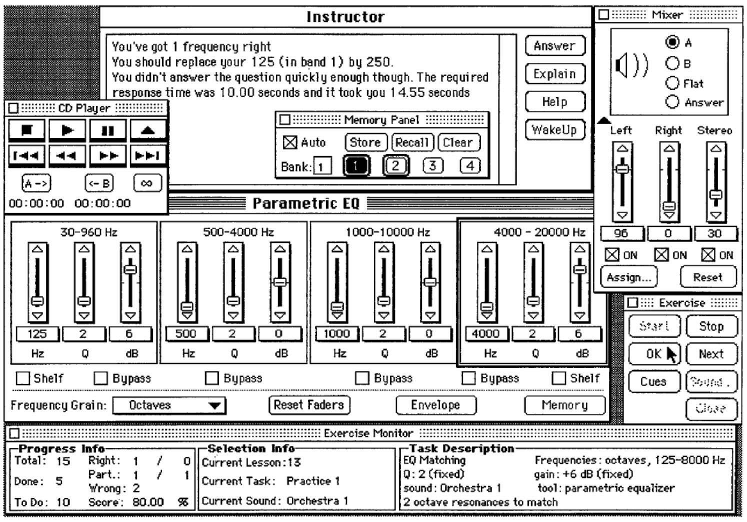

Figure 6. The 1996 version of the Timbral Ear Trainer user interface. (Quesnel, 1996)

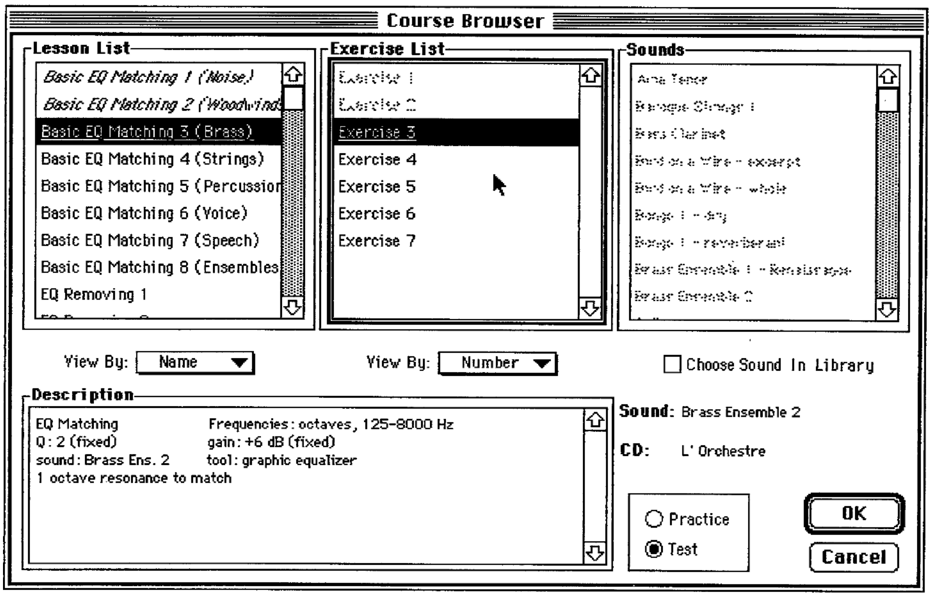

Figure 7. The 1996 Timbral Ear Trainer course browser (Quesnel, 1996)

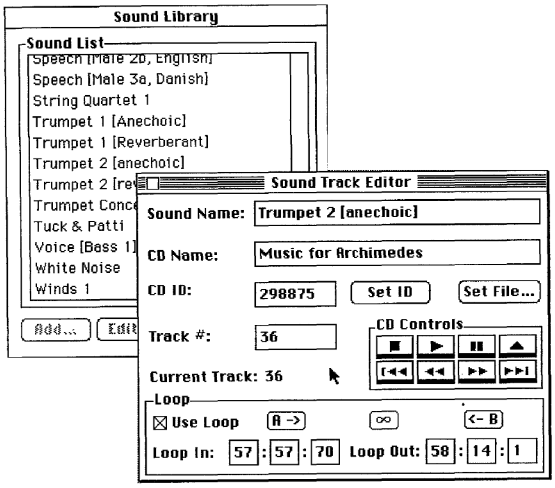

Figure 8. The 1996 Timbral Ear Trainer sound library editor. (Quesnel, 1996)

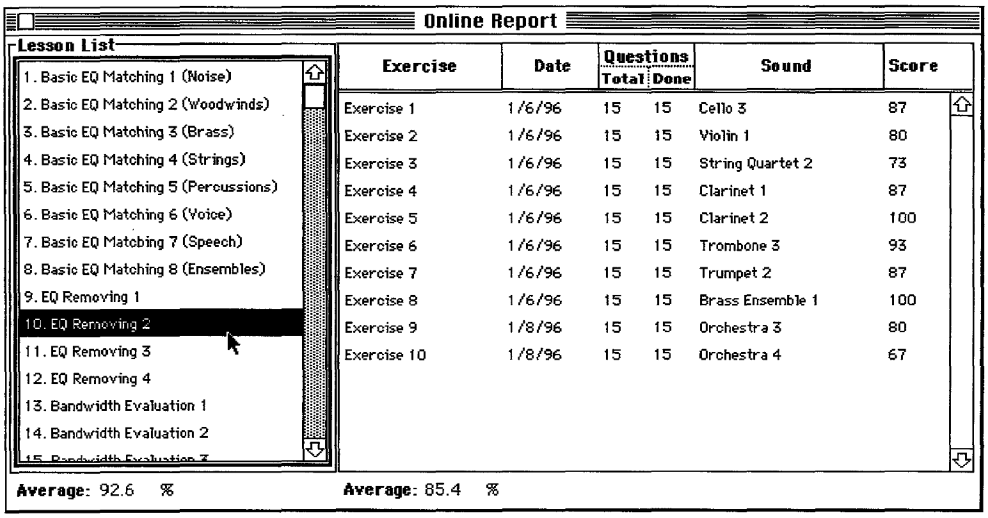

Figure 9. The 1996 Timbral Ear Trainer online report. (Quesnel, 1996)

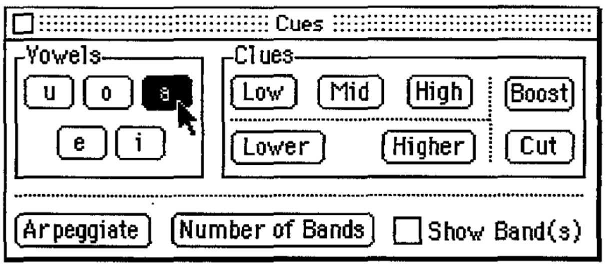

Figure 10. The 1996 Timbral Ear Trainer cues window. (Quesnel, 1996)

In Quesnel's (2001) PhD thesis, he details the latest version of the Timbral Ear Trainer program. The revised version again features one-on-one tutoring sessions delivered in combination with individual practice time using the Timbral Ear Trainer program. The program utilises 1/3rd octave centre frequencies ranging from 63 Hz to 16k Hz, 'complemented with the use of vowel labels between 250 Hz and 4 kHz' (ibid., p. 45). In his thesis, Quesnel (2001) presents the results of a listening experiment designed to answer the following question:

> How would the performance of student audio engineers with little real-world experience but formal timbral ear training (the experimental group) compare, on

an advanced equalization task, with the performance of experienced audio engineers with no such formal training (the control group)? (p. 69)

Two additional questions were then investigated:

1. Did the tasks on which the subjects in the experimental group were trained involve listening skills that are used by professional audio engineers in the course of their work?
2. Which parameters would reveal the greatest number of differences and similarities between the two groups? (p. 69).

The results of this experiment are discussed in detail in Chapter 3.

2.1.4 Spectral Ear Training

Spectral Ear Training (Figures 11–13) is a technical ear training program developed at the Danish Acoustical Institute, designed to rapidly create 'human spectrum analyzer[s]' (Brixen, 1993, p. 1). The training program has a duration of two days, separated by three to four weeks. In the period between the two training days, users are required to undertake 'intensive self-training' (ibid., p. 3). The Spectral Ear Training program requires users to identify the centre frequency of a parametric equalizer, set at one or two of 10, ISO octave-band frequencies ranging from 125 Hz to 8k Hz. Boost or cuts can be introduced at either one or two centre frequencies simultaneously. Regardless of the boost/cut presentation for a pair of sounds, the absolute magnitude of gain changes is identical (+12dB/-12dB, +6dB/+6dB etc.). The stimuli (noise, music and speech) contains typically unwanted noise such as 'ventilation systems, hum, "room noise", wind, turntable rumble, tape noise, hiss, s'es in speech etc.' (ibid.).

During the training, users are encouraged to identify the centre frequency of the parametric boost based on one of three methods:

1. identifying the frequency of the boost based on the perceived change in spectrum;
2. associate the perceived change in spectrum with a subjective term such as 'foggy', and then associate this terms with a specific frequency. Users are free to choose their own words or not use any; or
3. a modified 'sweep mode' version of (1) where the user focuses their attention 'on the frequencies present from the low frequencies towards upper part of the frequency range and vice versa'. (Brixen, 1993, p. 4).

Following the presentation of the modified spectrum, a 'flat' version is presented. Each test features 10 questions with spectral modifications randomly chosen from the 10 octave-band frequencies. Testing lasts no more than half an hour, followed by a half-hour of instruction covering the 'basics in psychoacoustics, basics in filter techniques, [and an] introduction to the self-training' (Brixen, 1993., p. 5). At the end of each training day, 'a test similar to the first one is carried out', designed to measure the 'enhancement of the listening skills which are built up during the day' (ibid., p. 6).

The first day of training introduces the training program using pink noise and solo musical instruments. Theoretical instruction in psychoacoustics is also presented and different filter types and response areas are introduced. The second day introduces 1/3rd octave centre frequencies and concentrated on the spectra of both male and female speech, pink noise and music. The theoretical instruction on the second day covers the acoustics and psychoacoustics of speech ('formants, vowels, fundamentals, energy dispersion, directional characteristics etc.' (Brixen, 1993, p. 6). Self-training, to be carried out regularly for three to four weeks, involves the use of cassettes and an answer sheet that is distributed to users during the training days.

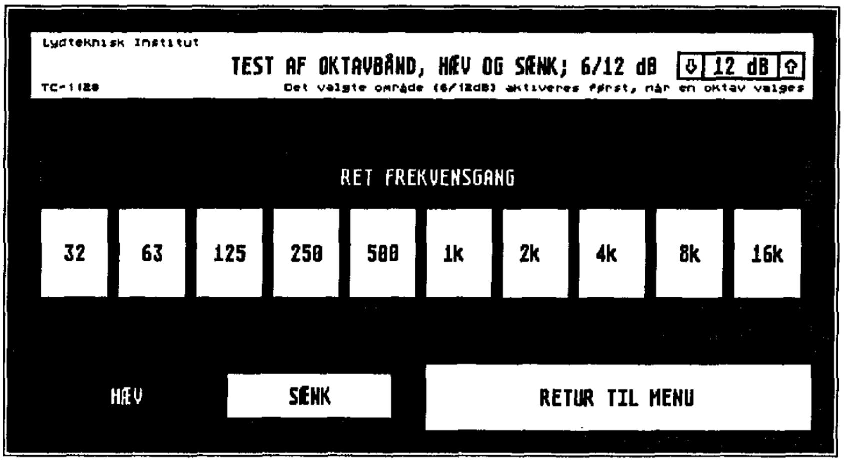

Figure 11. The Spektlyt version 1.0. TETP (Brixen, 1993, p. 16)

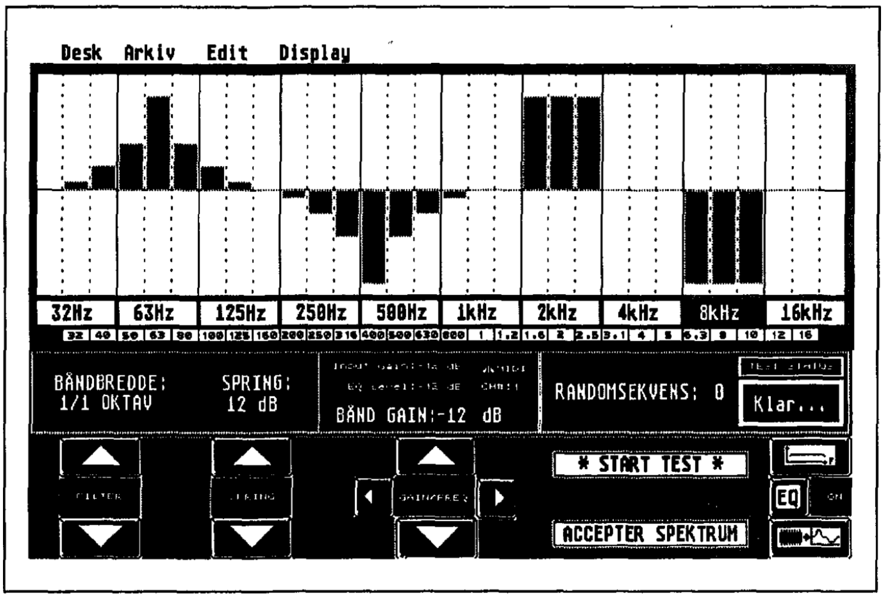

Figure 12. The Spektlyt version 2.0. TETP (Brixen, 1993, p. 17)

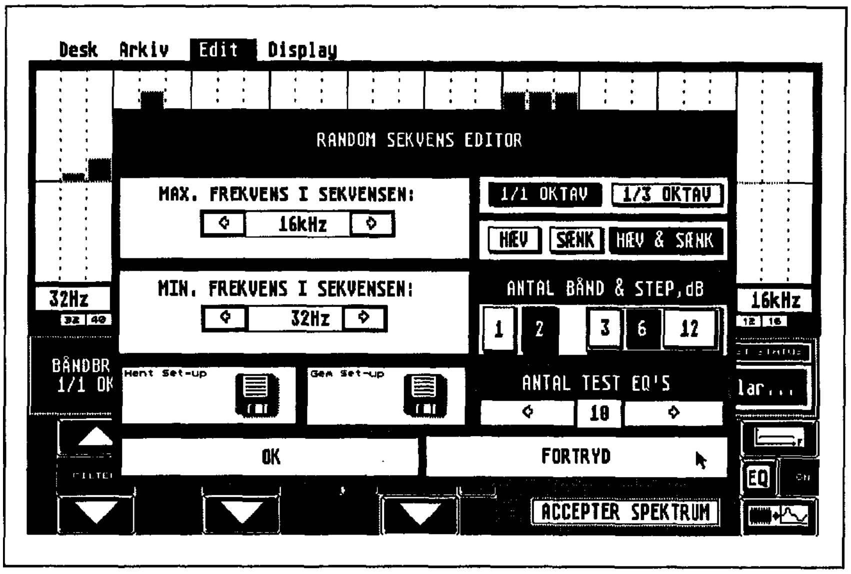

Figure 13. Editing Menu Spektlyt version 2.0. (Brixen, 1993, p. 18)

2.1.5 Technical Listening Training

Researchers at Kyushu University (Kyushu Institute of Design) developed a Technical Listening Training program designed to 'increase sensitivity to auditory differences' (Arai, Satoh, Nishimura, Ueno, & Yoshihisa, 2006) and develop an 'understanding of the relationship between acoustic properties and auditory impression' (Kawahara, Takada, Iwamiya, Nakajima, & Ueda, 2004b, pp. III -1966). The training program is delivered to both first and second-year undergraduate students in classes titled Technical Listening Training I and II, respectively (Arai et al., 2006; Iwamiya, Nakajima, Ueda, Kawahara, & Takada, 2003; Kawahara et al., 2013; Kawahara et al., 2004b). Technical Listening Training I introduces students to fundamental acoustic concepts such as frequency (pitch), level, spectrum and decibels and 'beg[an] with discrimination tasks for pitch, loudness and timbre' (Iwamiya et al., 2003). Training in pitch discrimination features pure tones presented pairwise. The reference tone is the standard A440 Hz and the second tone is either slightly higher or lower in frequency. Students are required to indicate if the second tone is higher or lower in pitch than the first. Differences between tones range from 5 Hz to 1 Hz.

Training tasks in loudness and discrimination of timbre are similar to those of pitch. For loudness, sounds are again presented pairwise, with students required to indicate whether the second sound is louder or quieter than the reference sound. Timbre discrimination tasks again feature pairwise presentation of sounds, but students are required to indicate whether the sounds are the same or different.

Following these discrimination tasks, students are introduced to identification tasks, including identifying the 'difference in sound pressure level, frequency of pure tones, center frequency of band noise, number of harmonics, amplitudes slope of harmonic components, and frequency of enhanced bandwidth of colored music' (Iwamiya et al., 2003., p. 28). Before the identification training begins, students undertake a 'rehearsal session' in each of the different areas. For sound pressure level for example, 'students listen[ed] to a series of sound pairs and correlate[d] loudness differences to differences in the sound pressure level' (ibid., p. 29).

Loudness difference identification tasks require students to identify the difference in amplitude (in dB) between two music excerpts presented pairwise, at either basic, intermediate or advanced levels. Basic levels feature 10dB increases from 0db to 30dB, intermediate levels feature 5dB increases from 0dB to 20dB or 0dB to 30dB, advanced levels feature 2dB increases from 0dB to 10dB. Frequency bandwidth training features the pairwise presentation of flat and modified sounds. The spectrum is modified using a 10dB parametric boost at one of seven octave-band frequencies ranging from 125 Hz to 8k Hz. The task requires users to identify the centre frequency of the boost. The stimuli include various genres of music. Pure tone identification training requires students to identify the frequency of a pure tone, chosen from one of seven octave-band frequencies, again ranging from 125 Hz to 8k Hz. A noise identification task is also presented, in which students must identify the centre frequency of band-limited noise. Training is also delivered, involving 'harmonic number and amplitude slope of harmonic components' designed to help 'students correlate acoustic spectrum to perceived timbre' (Iwamiya et al., 2003, p. 29).

Technical Listening Training II follows on from the first course (which is a prerequisite). The advanced version of the frequency bandwidth training is identical to the introductory version, but the magnitude of the parametric boost is reduced to 6dB. Training in identifying the frequency of low- and high-pass filters is included in the advanced program. The advanced version of the loudness difference identification task is also similar to the introductory version, but students are presented with two parts of a musical ensemble and are required to identify the difference in amplitude (in dB) between the parts. This task is designed to simulate control of a recording engineer's mixing console (Iwamiya et al., 2003). Training in identifying the spectral slope of music in addition to training in 'identify[ing] the time difference between two parts of an ensemble' (ibid., p. 30) is also provided. Training in identifying the reverberation time of music and synthesised sounds is provided, in addition to 'training involving the identification of just and mean temperament and total harmonic distortion' (ibid., p. 30). The training program is delivered through the use of personal digital assistant (PDA) terminals connected to a Windows NT-based system in an audio studio control room. The PDA, used with a pen-like device, provides visual feedback throughout the training including the percentage of correct responses.

Nishimura's (2015) follow up paper details the results of an investigation into the effect of the Technical Listening Training program on the understanding of lecture content, as opposed to listening ability. In the study, students were divided into four groups; all four received lectures on audio processing, but only two groups undertook structured technical ear training (TET). The non-TET groups listened to the same sound files as those in the TET program, but without the structured training element; they simply listened to static sound files. Playback was via headphones. The ratio (percentage) of correct answers on various tests administered throughout the course indicated the students' level of understanding of the lecture material. For the lectures on the digitization of sound, 'the same degree of understanding was demonstrated by students who participated in ear training and those who did not' (ibid., p. 4). For the sound quality and perceptual coding lectures, Nishimura suggests that the 'results indicate that students who participated in ear training demonstrated a better understanding of perceptual codecs' (ibid.).

2.1.6 Web-based Technical Listening Training

Nishimura (2013) subsequently developed a web-based multilingual version of the Technical Listening Training program that features several advantages over the previous version, including multilingual support and the ability for users to set training parameters. The training program (Figure 14) was delivered at Tokyo University of Information Sciences alongside a lecture entitled Audio Information Processing (AIP). The purpose of the ear training program within the AIP course '[was] to help students carefully listen to sounds and learn about factors that determine audio quality through audio experiences. Students are not generally required to improve their ability to discriminate sounds' (ibid., p. 3). Ear training was delivered for the first 10 weeks of the 12-week semester in alignment with the theme of the lecture.

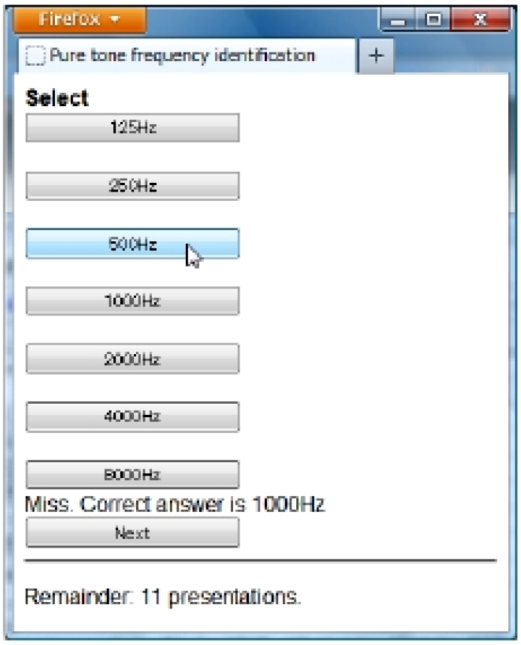

Figure 14. web-based multilingual version of the Technical Listening Training program. (Nishimura, 2006b)

2.1.7 Beijing Union University Technical Ear Training Program

Liu, Wu and Yang (2007a; Liu, Wu, & Yang, 2007b) presented the results of a technical ear training program delivered at the Beijing Union University. The goal of the training program was to speed the 'develop[ment] [of] the subjects' discrimination of sound frequency, sound level and sound spectrum' (Liu, Wu, & Yang, 2007a, p. 2). The program trained students in six areas of critical listening:

1. Identification of pure tones;
2. Identification of the difference in Hz between two presented pure tones;
3. Identification of the difference in dB between two presented sounds;
4. Aural identification of musical instruments; and
5. Identification of frequency response irregularities. (pp. 2-3)

A second paper, published simultaneously, details the results of the pure tone identification training only (Liu, Wu, & Yang, 2007b).

2.1.8 Personalised Timbral Ear Trainer

Kaniwa et al. (2011) describe an adaptive personalised technical ear training program entitled Personal Timbral Ear Trainer that 'controls the training task based on the trainees' previous performance' (p. 1). The program is based on Quesnel's Timbral Ear Trainer program, but has been modified to be adaptive in nature. The system modifies the presentation of frequencies based on each individual student's performance on previous questions. Frequencies that a student performs poorly on are flagged as 'weaknesses' and are subsequently presented more frequently in the program. The authors suggest that the adaptive nature of the program allows students to 'effectively study technical listening without an instructor' (ibid).

Kim, Kaniwa, Terasawa, Yamada, and Makino (2013) published a follow up paper further detailing the Personal Timbral Ear Trainer program. The authors set out to establish whether 'there [was] a listener-dependent idiosyncrasy in identifying spectral differences' and whether 'adaptively regenerated training contents (for a specific trainee) assist in acquiring timbre identification ability more effectively than non-adaptive ones' (ibid., p. 425). The TETP requires users to individually identify the centre frequency of 25 parametric filter boosts or cuts at one of seven octave-band frequencies ranging from 125 Hz to 8k Hz in an absolute identification task (Q was not stated). The user can switch between an unprocessed and processed version of the sound at any time. The stimuli are pink noise, orchestral music, solo piano and drums. Users can also import their own sound files.

2.1.9 University of Lethbridge TETP

Schaller and Burleigh (2015) detail an online TETP embedded within the Bachelor of Music program at the University of Lethbridge. The authors describe several TET applets (small computer programs) that are used within the Digital Audio Arts curriculum, the first of which is the Noise Spiral.

The Noise Spiral applet was 'developed to facilitate the learning process of frequency band recognition' by requiring students to identify the centre frequency of various parametric boost or cuts, first applied to pink noise, then music and other stimuli including their own audio files (Schaller & Burleigh, 2015, p. 5). The gain was initially set to +12dB and subsequently reduced gradually to +3dB, likewise the number of possible frequencies starts at octave bands before increasing to 1/2 and 1/3rd octave bands. German vowels can be overlaid onto the Noise Spiral graphic, which the authors suggest 'are a great help… [as] even people who do not have previous experience working in music or music technology disciplines … are familiar with the timbre of spoken vowels' (ibid.). In order to assess students' progress, online listening quizzes are scheduled throughout the course.

2.1.10 McGill University

Martin and Massenburg (2015) describe a TETP that features a set of matching exercises delivered using the Pro Tools® DAW. First-year students at McGill University's graduate program in Sound Recording undertake 'timbral ear training', whilst second-year students are trained 'on several more advanced skills' (ibid., p. 1). The authors state that the goal of both TETPs 'is to enable the students to work more quickly with a greater degree of precision' (ibid.). The authors suggest that their TETP differs from others in the way in which the program's difficulty increases; traditional TETPs ‘increase the difficultly by making changes increasingly subtle’, whereas the McGill program increases in complexity in addition to subtlety (ibid., p. 2). The authors also differentiate their TETP from others as the McGill program requires users to match individual tracks within a mix, as opposed to the more traditional global manipulation of a finished stereo mix. Finally, the TETP required users to perform dynamic signal processing – matching the automation of levels, panning and effects. These more advanced tasks are introduced ‘in a two-semester Advanced Technical Ear Training course’ (ibid.).

The frequency-spectrum tasks contained in the Advanced Technical Ear Training course are ‘Advanced Equalization Matching’ tasks. These removing tasks feature two separate audio tracks in Pro Tools® with identical musical content. One of the tracks is unprocessed (the ‘question track’), while the other (the ‘answer track’) has one or more resonances introduced using a high-Q parametric equaliser. Students are required to insert an equaliser on the ‘answer track’ and apply the inverse of the equalisation that has already been applied, in order to make it sound the same as the ‘question track’. The authors state that to their knowledge ‘this is the only exercise of its kind that adds this complexity to EQ training’; however, Corey’s (2010, 2017) Technical Ear Training program (detailed in Section 2c) allows users to train with matching and removing tasks, with multiple (up to three) high-Q filters applied. Corey’s TETP also increases difficulty by increasing complexity (the addition of more bands of equalisation). Additional advanced equalisation training is featured in the ‘Mix-alike’ assignments. In these assignments, students are given a professionally mixed stereo mix-down and a multitrack session of the mix. Students are required to, amongst other tasks, match the equalisation applied to the lead vocal track (post-compression) within the mix.

2.2 Consumer-Targeted Training Programs

This section details the three consumer-targeted frequency spectrum-based programs for which published information is available. Moulton's Golden Ears and Everest's Critical Listening and Auditory Perception programs are regarded as seminal works in the field that also achieved commercial success. Corey's Audio Production and Critical Listening program follows on from the work of Moulton and Everest, providing free public access to a suite of interactive technical ear training programs.

2.2.1 Golden Ears

In 1966, as part of his duties as a graduate instructor at the Juilliard School of Music, David Moulton was required to create and teach musical ear training for first and second year undergraduate students (Moulton, 2014). In 1969 he started a small recording studio, and taught one of his first clients to use a sound reinforcement system (ibid.). As the client did not understand 'what terms such as LF or 120 Hz might mean or sound like', using pink noise and a 10-band graphic equaliser, Moulton created a series of drills based on the traditional musical ear training that he had received and then later taught at Juilliard (ibid., p. 2). Moulton (2014) subsequently incorporated these drills into the university classes that he taught, using them:

> in almost all of [his] teaching from 1969, as a private teacher and then Chair of the Sound Recording Technology program at the State University of New York, through 1993 [as part of their Tonmeister Studies program], [and] as academic chairman of Berklee College of Music's Music Production and Engineering Program'. (pp. 3-4)

Moulton's technical ear training program was 'also incorporated into the National Public Radio Music Recording Workshop curriculum from 1984 through 1996' and has been used at 'the University of Massachusetts at Lowell, Emerson College, the School of the Museum of Fine Arts in Boston and the University of Maine at Augusta' (Moulton, 2014, p. 4). Other institutions to deliver the program include the University of California, New York University, the Danish Acoustical Institute and the National Broadcasting Corporation (Moulton, 2014).

In 1993, KIQ productions approached Moulton to release his ear training program on CD, a product that was later released as the famed *Golden Ears* training program which has been 'in print ever since, without modification' (Moulton, 2014). The primary goal of the program was to develop 'the ability to accurately describe the physical nature of what we are hearing' (KIQ Productions Inc., 2005, p. 1). The *Golden Ears* training program is divided into four volumes;

1. Frequencies;
2. effects and processing;
3. time domain – delay and reverb drills; and
4. master frequencies – 1/3 octave and dual-octave drills. (Moulton, 1995)

Volume one: 'Frequencies - Disc 1' features sets of drills each containing 10 samples, with each sample containing either pink noise or music. Each sample is first presented flat for a period of three seconds, then a modification is made to the spectrum by applying a boost or cut to one of the standard octave band centre frequencies using a graphic equaliser. The modified signal is then presented for four seconds before the flat signal is again presented. The user must identify the centre frequency at which the modification took place. The drills are separated by overlapping frequency ranges into '"low" (octaves one through five, 31-500Hz), "mid" (octaves four through eight, 250-4000 Hertz), and "high" (octaves six through ten, 1KHz-16KHZ)' and by gain (Moulton, 1995, p. 9). Drill 1 for example, has modifications (and possible answers) only at the lowest five octave band frequencies and only contains 12dB boosts. Drill 9 contains modifications (and possible answers) only at the highest five octave band frequencies and only contains 12dB cuts. However, in drill sets 13 and 14, all 10 octave band frequencies are included. Each drill is preceded by a warm-up drill and answer sheets are provided at the end of the chapter. Volume one: 'Frequencies - Disc 2', features similar drills to Disc 1, but the last 10 of the 14 drills comprise either boosts or cuts to the spectrum. In these drills, users are required to identify not only the centre frequency of the boost, but whether the spectrum has been boosted or cut.

Volume two: 'Effects and Signal Processing' features the presentation of paired comparisons in sets of five, in which the signal is first presented flat, then with 'some sort of signal processing or audio anomaly added' (Moulton, 1995, p. 27). The 31 possible signal processing changes are grouped into six 'families' of modifications; 'amplitude change, [total harmonic] distortion, compression, equalization, stereophony, and time-delay / reverberation' (ibid.). In addition, the paired comparisons feature an unmodified version of the signal as a reference. Aside from simple differences in level between the sounds, the amplitude difference drills feature a gradual increase in level over the duration of the sample (ibid.).

Volume three: 'Time Domain – Delay and reverb drills' focuses on the development of sensitivity to time intervals. Disc 1 of this volume focuses on the perception of short delay times, whilst Disc 2 is centred on pre-delay and reverberation. The drills are presented in sets of 10. In the delay drills, users must identify both the delay time from a limited set of options, and the channel in which the delay was applied. In the decay/reverb drills, the user must identify both the pre-delay and reverb time for each sample, again from a limited set of values.

Volume four: 'Master frequencies – 1/3 octave and dual-octave drills' focuses on developing advanced timbre discrimination ability. The first disc requires users to identify the 1/3rd octave band centre frequency and the type of gain modification (boost or cut), whilst the second disc requires the user to identify two simultaneous octave band modifications. The frequencies in this volume are again grouped into 'low, 'mid' and 'high' sets, with pink noise and music as stimuli.

Moulton gauged the success of the program based upon observing students' 'improved abilities to hear equalization changes and accurately describe them technically, as well as by indirect reports from them about the confidence they felt in their work', in addition to the students' success as graduates (totalling approximately 3000) (Moulton, 2014, pp. 5-6).

2.2.2 Critical Listening / Auditory Perception

In 1982, F. Alton Everest presented a paper titled 'Instruction in critical listening' at the 72nd AES convention in Anaheim, California. The paper was based on his *Manual for critical listening: An audio training course* (Everest, 1982b), which describes the first commercially available technical ear training course, now one of the most widely cited and respected critical listening courses in use today (Figure 15). Everest's critical listening course, released on cassettes, was designed to accelerate the process of learning to 'detect minute deviations from accepted norms of audio quality', and to demonstrate the 'close relationship between audio faults and audio technology' alongside the training program (1982a, p. 1).

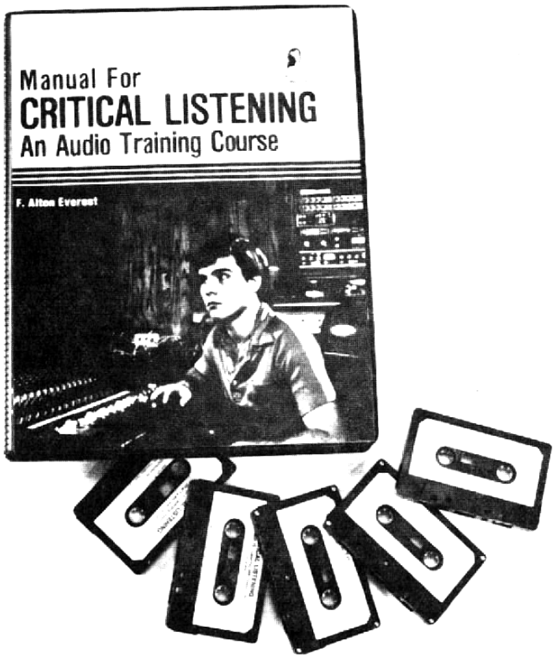

Figure 15. Everest's Manual for critical listening: An audio training course. (Everest, 1982b)

In his seminal paper, Everest (1982a) suggests that 'the discriminating aspect of human hearing surely can be improved through training and experience' (p. 2) and that this improvement can be accelerated through a structured training program, although no evidence of the effectiveness of the training program has ever been presented. The course itself consists of 10 lessons presented on five double-sided cassettes and a 106-page manual with the following sections:

1. Estimating the frequency of sound;
2. Estimation of sound level changes;
3. Estimating frequency band limitations;
4. Frequency response irregularities;
5. Judgment of sound quality;
6. Detecting distortion;
7. Reverberation effects;
8. Signal versus noise;

9. Voice colorations; and
10. Listening with discrimination. (1982a, Fig. 2)

Each lesson concludes with a listening test. The manual contains sketches, the narrator's script, definitions, and frequency response graphs. The training program is entirely passive and features proficiency tests throughout requiring users to (for example) detect 'wow' and 'flutter' in music and speech, with associated evaluation (answer) sheets provided. The 10th lesson is a review, in which musical selections are presented. Each musical section opens and closes with unmodified audio, with modified audio contained between. Everest released multiple iterations of the training program with minimal modifications beyond presentation format.

In 1986 Everest released the *Manual for auditory perception: An audio training course* in which he presented 'eight units designed for instruction in the psychoacoustical aspects of human hearing' (Everest, 1986, p. xx):

1. Loudness, pitch, and timbre;
2. How one sound masks another;
3. How the ear analyzes sound;
4. Non-linearities in the auditory system;
5. The perception of delayed sounds;
6. Why some sounds are more pleasant than others;
7. How we locate sounds; and
8. True binaural listening. (ibid., p. vii)

The training course again featured a manual and cassettes, but unlike the original course that was devoted to technical ear training, the *Manual for Auditory Perception* course focused primarily on psychoacoustics and auditory perception.

In 1997 the two programs – *Manual for critical listening: An audio training course* and *Manual for auditory perception: An audio training course* – were combined and released as *Critical listening and auditory perception: The complete audio-visual training course* (Everest, 1997). Although the content of the training programs is identical to the two separate earlier versions, the format was upgraded, with the audio material released on five CDs.

Two years after Everest's death in 2005, *Critical listening skills for audio professionals* was released (Everest, 2007). In a similar manner to the 1997 release, this book was divided into two sections: the first focused on critical listening, and the second on auditory perception. The audio material was again released on CD, but in MP3[6] format (320kbps). It is unclear why a revered course on critical listening was ultimately released on a lossy format (one that discards some of the audio data), because, as Corey (2010) suggests, encoded audio files (such as MP3) 'should never be used for EQ exercises, even if they have been converted back to PCM[7]' (pp. 48-49). The use of MP3-encoded audio is especially confusing, considering the book's preface concludes by stating that '[w]ith compact disc sound, [the course] promises to be a solid contribution for the future' (Everest, 2007, p. viii).

---

6 An audio coding format for digital audio that provides reduced file size.
7 Pulse-code modulation.

2.2.3 Audio Production and Critical Listening: Technical Ear Training

In 2010, Corey published *Audio production and critical listening: Technical ear training,* a book focused on technical ear training that included a CD-ROM containing software modules that align with each section of the text. The text begins by presenting a detailed introduction to active and critical listening, technical ear training, and sound reproduction system configurations. In addition to spectral balance and equalisation, the book contains chapters and associated training modules on:

- 'spatial attributes – delay and reverberation;
- dynamic range control – compression/limiting and expansion;
- sounds or qualities of sound that can detract from recordings – distortion and noise; and
- audio excerpt cut-off points – source-destination editing'. (Corey, 2010, p. xii)

The aim of the book and associated software modules was 'to explore critical listening as it relates to typical types of audio signal processing' and to 'present[] some ideas for developing critical listening skills and potentially reduce[] the time it takes to develop them' (Corey, 2010, pp. x-xi). Corey outlined three main goals of the text and associated software training modules:

- to facilitate isomorphic mapping of technical parameters and perceived qualities of sound;
- to heighten awareness of subtle features and attributes of sound, and to promote a greater ability to differentiate among minute changes in sound quality or signal processing; and
- to increase the speed with which one can identify features of sound, translate between auditory perceptions and signal processing control parameters, and

decide on what physical parameters need to be changed in a given situation. (ibid., p. xii)

Corey (2010) hypothesised 'that increased sensitivity in one area of critical listening (such as equalization) will facilitate increased awareness and sensitivity in other areas (such as compression and reverberation) as a result of overall improved listening skills' (p. 13).

Chapter two focused on spectral balance and equalisation areas of technical ear training. After introducing the many ways in which the audio engineer's decisions shape the spectral balance of a signal (equalisation, microphone choice and placement, monitors, room acoustics, playback levels etc.), Corey provides a brief overview of equaliser types, before introducing the associated software training module for the chapter, the 'Technical Ear Trainer – Parametric Equalization Practice Module'. The equalisation practice module (Figure 16), designed so the user can 'practice hearing the sonic effect of various equalization parameters', allows the user to select pink noise or import a mono or stereo audio file as the stimulus (Corey, 2010, p. 37). Playback can be set to mono or stereo and if an audio file has been imported the user can select a region to loop over if desired. The module offers four different types of frequency spectrum-based training as described in Chapter 1: matching, matching memory, return to flat, and absolute identification.

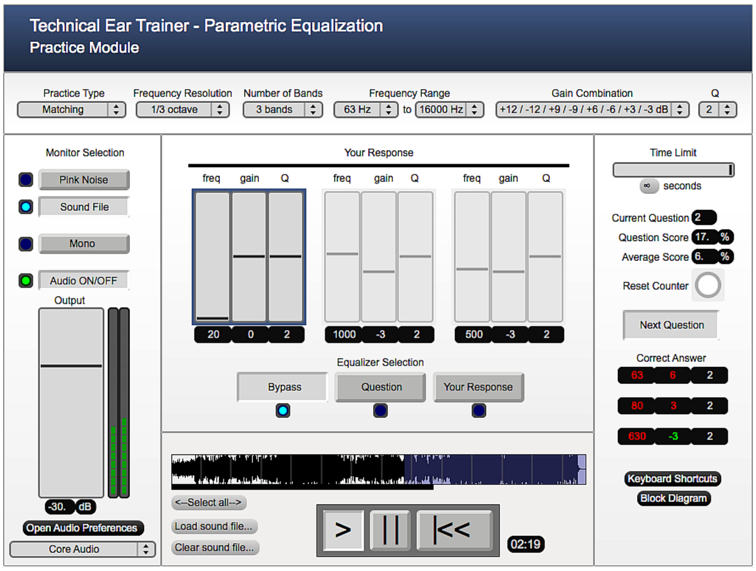

Figure 16. Technical Ear Trainer – Parametric Equalization Practice Module. (Corey, 2010)

Users can select one, two or three simultaneous bands to practice with; the frequency range is selectable between 63 Hz and 16k Hz, and the frequencies within the bands can be set to octave or 1/3rd octave. For each training session, the gain and Q are selectable, but the Q is fixed for each question. The time limit for each question can be set at between one and 30 seconds. Questions are randomly generated by the software module within the limitations chosen by the user. The user's responses (frequency and gain) are coloured green if correct or red if incorrect, and a running 'question score' and 'average score' are displayed. The program features keyboard shortcuts designed to speed up response times and reduce the amount of mouse-related work on screen. Indeed, almost all the parameters of the training program can be controlled using shortcuts, allowing the user to abandon use of the mouse almost entirely.

Corey (2010) advises readers to initially limit their practice to 'short but regular practice times on a daily or several times a week basis', suggesting that in order to avoid fatigue in the early stages, '10- to 15-minute practice sessions are probably best' (p. 37). Corey suggests that over time users should gradually increase their practice, but cautions '45 to 60 minutes is going to be the upper useful limit' per practice session (ibid., p. 38). Corey concludes the chapter by stating that the training software 'can serve as an effective tool for progress in technical ear training and critical listening when used for regular and consistent practice' (ibid., p. 49), but to date no empirical data relating to this training program has been presented.

In a later paper, Corey (2013) provides an overview of the implementation of his *Audio production and critical listening: Technical ear training* program within a classroom setting at the University of Michigan. He cites one of the main learning objectives of the program as 'making clear, unambiguous distinctions between memories for adjacent octave or third-octave frequencies' (ibid., pp. 3-4). In contrast to the full version of the training program, in this implementation, students were only permitted to perform absolute identification tasks. Students are encouraged to begin the training using pink noise as opposed to music, as pink noise has equal power per octave on average and thus 'typically represents frequency content more evenly than most music recordings' (ibid., p. 3). Thus, when a boost is applied at a given frequency using pink noise as the stimulus, 'we can be assured that a boost at another frequency will produce an effect of comparable strength' (ibid.).

Corey (2013) presents anecdotal evidence of increased performance on the technical ear training tasks, suggesting that '[a]s the class progresses, students become more accurate, more confident, and faster in their response times' (p. 4). He draws attention to a test he provides students in which they are required to identify the centre frequency of a +12dB parametric boost applied to pink noise or familiar recording. On this test, he claims that students are 'typically […] able to correctly identify alterations by ear within about 5 seconds' and 'consistently make correct judgments most of the time' (ibid.). Corey concludes by suggesting that '[t]echnical ear training can help increase sensitivity to changes in timbre, develop memory for timbre and its associated physical properties, and increase the speed in which one can make correct identifications of timbral alterations' (ibid., p. 5). The metrics by which performance on technical ear training is measured are discussed in detail in Chapter 3.

In 2017, Corey released the second edition of *Audio production and critical listening: Technical ear training* as part of the 'Audio Engineering Society Presents' series (Corey, 2017). Unlike the previous version that included a CD-ROM disc containing the technical ear training applications, the new version of the text was released along with new, publicly available web browser-based versions of the applications (Figure 17). The updated text 'includes information on objective measurements of sound, technical descriptions of signal processing, and their relationships to subjective impressions of sound' (Audio Engineering Society, 2017b, p. 1).

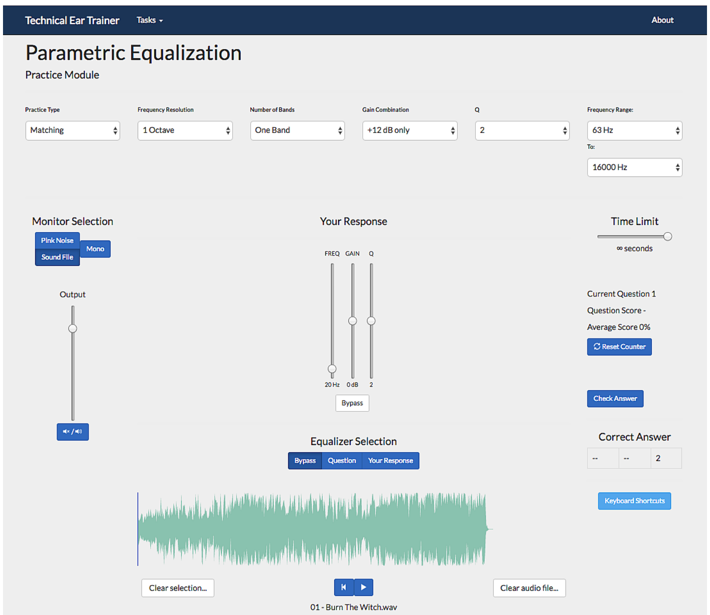

Figure 17. Technical Ear Trainer – Parametric Equalization Practice Module. (Corey & Benson, 2017)

2.3 Summary

This chapter provides a descriptive summary of frequency spectrum-based TETPs targeted at students and consumers for which published papers or texts are available (Table 1). Since the deployment of the Timbre Solfeggio program in 1974, TETPs have continued to grow in complexity and developers have become more interested in establishing the effectiveness of training programs.

Table 1 - Overview of Student and Consumer-targeted TETP aims.

Student-targeted Training	
Training Program	**Aims**
The Teaching Machine	'To develop an awareness of instrumental tone color as an important element of music' (Greenberg & Huddleston, 1971, p. 59)
Timbre Solfeggio / Solfège	Train and improve students' sensitivity to relative differences in timbre and their memory for absolute timbral references.
	'Develop a set of qualitative descriptors to provide unequivocal codes (language) for exchanging information regarding timbre impressions' (Letowski, 1985, p. 241)
	To provide a 'general orientation to the relationship between timbre changes and the physical transformations used to obtain them' (Letowski, 1985, p. 241)
	Identify and describe 'the basic properties of sound quality' (Letowski, 1985, p. 241)
	Develop 'the listening skills of NVH engineers and prepared them for leading roles in the sound evaluation panels' (Miśkiewicz & Letowski, 2014, pp. 4-5)
Timbral Ear Trainer I and II	'The improvement and maintenance of the aural skills that are essential to audio engineers' (Quesnel, 1990, p. 74)
	Develop long-term memory for a basic set of timbral categories;
	Progressively develop the listener's sensitivity to subtle timbre differences
	Develop analytic listening skills for the evaluation of complex spectral changes in terms of individual resonances, specifying for each the magnitude, Q, and center frequency. (Quesnel and Woszczyk, 1994, p. 2)
Spectral Ear Training	Rapidly create 'human spectrum analyzer[s]' (Brixen, 1993, p. 1).
Technical Listening Training	'Increase sensitivity to auditory differences' (Arai, Satoh, Nishimura, Ueno, & Yoshihisa, 2006)
	Develop an 'understanding of the relationship between acoustic properties and auditory impression' (Kawahara, Takada, Iwamiya, Nakajima, & Ueda, 2004b, pp. III -1966)
Web-based Technical Listening Training	'To help students carefully listen to sounds and learn about factors that determine audio quality through audio experiences' (Nishimura, 2013, p. 3)

Beijing Union University Technical Ear Training Program	'Speed the 'develop[ment] [of] the subjects' discrimination of sound frequency, sound level and sound spectrum' (Liu, Wu, & Yang, 2007a, p. 2)
Personalised Timbral Ear Trainer	Based on Timbral Ear Trainer I and II
University of Lethbridge TETP	'Facilitate the learning process of frequency band recognition' (Schaller & Burleigh, 2015, p. 5).
McGill University	'Enable the students to work more quickly with a greater degree of precision' (Martin and Massenburg, 2015, p. 1)
Consumer-targeted Training	
Training Program	**Aims**
Golden Ears	'Develop 'the ability to accurately describe the physical nature of what we are hearing' (KIQ Productions Inc., 2005, p. 1).
Critical Listening / Auditory Perception	Accelerate the process of learning to 'detect minute deviations from accepted norms of audio quality', and to demonstrate the 'close relationship between audio faults and audio technology' alongside the training program. (Everest, 1982a, p. 1)
Audio Production and Critical Listening	'Facilitate isomorphic mapping of technical parameters and perceived qualities of sound'
	'Heighten awareness of subtle features and attributes of sound, and to promote a greater ability to differentiate among minute changes in sound quality or signal processing'
	'Increase the speed with which one can identify features of sound, translate between auditory perceptions and signal processing control parameters, and decide on what physical parameters need to be changed in a given situation. (Corey, 2010 ,p. xii)
	'Make clear, unambiguous distinctions between memories for adjacent octave or third-octave frequencies' (Corey, 2013, pp. 3-4)

The next chapter presents a review of the relevant literature in the field and examines the methods that TETP researchers have used to gauge the effectiveness of their programs using three criteria, developed to establish the minimum requirements for TETP goals; professional relevancy, measured performance relating to goals, and performance measured separately from the TETP. The chapter begins with an examination of the division of opinion within the audio engineering community regarding the two main methods used to train students to operate equalisers.

3 LITERATURE REVIEW

The focus of this research project was the influence of the training method used in technical ear training programs on audio engineering students' ability to discriminate tone colour. In reviewing the literature on this topic, the schism within the audio engineering community regarding the two main methods used to train students to operate equalisers, and the deployment of these training methods within TETPs, is first examined. Despite numerous audio engineering texts promoting the ICA method, and several authors cautioning against its use, this review found that there are no published studies on the merits of either the ICA or ISA methods. As such, there is no published data to support claims about the effectiveness or ineffectiveness of either training method.

The author developed three criteria to establish the minimum requirements for TETP goals if a TETP is to be assessed for effectiveness; they must be *professionally relevant*; researchers must *measure performance relating to the goals;* and this performance must be *measured separately from the TETP, on a task that students have not previously trained with.* The approaches that researchers have taken to establish the effectiveness of their TETPs were then investigated using these criteria. The scope of this review was limited to student-targeted TETPs (as defined in Chapter 1) for which published information was available. In addition, only frequency spectrum-based TETPs were included in the review; this included any TETP that trains users by presenting modifications of the frequency spectrum. Performance data were available for many of the TETPs that met these inclusion criteria, but the availability of such data was not a prerequisite for inclusion in this review. The review did not include TETPs aimed at consumer or employee audiences, those focused on training students using nonspectral-based tasks, or those for which no published information exists. The following eight TETPs were included in the review:

1. Timbre Solfeggio / Solfège (Hojan & Rakowski, 1999; Letowski, 1985; Letowski & Miśkiewicz, 1995; Miśkiewicz, 1992; Rakowski & Trybuła, 1975; Rościszewska, 2011; Rościszewska & Miśkiewicz, 2014, 2015);
2. Timbral Ear Trainer I & II (Quesnel, 1990, 1994, 1996, 2001; Quesnel & Woszczyk, 1994);
3. Spectral Ear Trainer (Brixen, 1993);
4. Technical Listening Training (Arai et al., 2006; Iwamiya et al., 2003; Kawahara et al., 2013; Kawahara et al., 2004b);
5. Web-based Technical Listening Training (Arai et al., 2006; Nishimura, 2006a, 2006b, 2013, 2015, n.d.);
6. Beijing Union University Training (Liu et al., 2007a, 2007b);
7. Personalised Timbral Ear Trainer (Indelicato et al., 2014; Kaniwa et al., 2011; Kim, 2015; Kim et al., 2013; Webster University, 2014); and
8. University of Lethbridge Training (Schaller & Burleigh, 2015).

The review demonstrated that none of the above TETPs met all three developed criteria, as performance relating to the goals of these TETPs was either measured on the TETP itself or on a similar or identical task to that which students had previously trained on as part of the TETP. The utility of pre/post-training tests in establishing the effect of the TETP on trainees' critical listening skills was also reviewed.

3.1 LEARNING TO OPERATE EQUALISERS

This section details the ICA and ISA methods of training students to use equalisers and the commonly associated criticisms and approbations. The difficulty student may face when using the two methods is examined, as is the use of the methods throughout the audio education and professional audio industries. The deployment of the methods within digital audio signal processors is assessed, and finally the use of the methods within TETPs is investigated.

3.1.1 ICA AND ISA METHODS

The contention surrounding the most effective method of training students to learn to identify the 'sound' of certain frequencies when operating equalisers centres on the differences between the ICA and ISA methods. Criticisms of the ICA method focus largely on the exposure of the listener to what one author termed a 'barrage of sweeping frequencies' that allegedly 'dull[s] the senses' and lacks 'accuracy and cohesiveness' (Stavrou, 2003, pp. 166-167). These criticisms contrast with the ISA method, which some authors claim cultivates listening skills 'more quickly' than the ICA method, and helps to 'develop a deep feeling for the qualities of different spectral regions' (Case, 2007, p. 117), serving to 'strengthen your ear while maintaining your perspective and sensitivity' (Stavrou, 2003, p. 167). Katz suggests a third, lesser-known method that relies on traditional music solfège:

> For engineers who have a musical background – keep a keyboard handy to determine the key of the song and use your sense of relative pitch to determine the problem note. Then translate that note to a frequency with a converter and dip that frequency. (Katz, 2008, p. 140)

3.1.2 Difficulty of Use

Despite the criticisms of the ICA method, it may be an easier initial learning technique for novice engineers. Case suggests that 'until one has had several years behind a console… there is no harm in taking advantage' of the method, noting that the ICA method:

> is fast, relatively easy, and has the extra benefit that the producer and any band members or label executives who might also be in the control room will be able to follow along sonically and be supportive of the EQ decisions. (Case, 2007, p. 117)

While suggesting that the 'savvy, experienced engineer' 'eventually skips the boost and search steps entirely and instead reach[es] immediately for the frequency range they wish to manipulate', Case (2007) acknowledges that the ISA method is more difficult for novice audio engineers to use in the first instance, as the method requires them to 'imagine – in their mind's ear – the frequency ranges in need of alteration (p. 117). Although 'the goal is to skip the boost and search parts', in discussing the obstacles that novice audio engineers encounter when learning how to professionally operate an equaliser, Case notes that 'critical listening skills are developed over a lifetime and require careful concentration, good equipment, and a good monitoring environment. No one learned the difference between 1 kHz and 1.2 kHz overnight' (ibid., p. 118). Novice audio engineers have a relatively limited amount of critical listening and equalisation experience and are therefore yet to build an exhaustive memory for the sounds of the various frequency spectrum regions. This makes the ISA method more difficult to use in the early stages of an audio engineer's career. However, despite the challenges associated with this method, Case encourages novice audio engineers to 'try [it] from time to time, during low-pressure recording sessions' (ibid., p. 117).

The ICA method, however, is not a technique suited solely to those unfamiliar with the sound of various frequency bands, nor is it a technique exclusively used by the novice audio engineer, as 'even the famous, expensive engineers resort to the boost, search, and set approach on occasion' (Case, 2007, p. 117). Brixen (1993) also suggests that the ICA method has a place in professional audio situations. As audio engineers are often required 'to determine what is not present in the spectrum', Brixen suggests that the 'listening procedure […] has to include a kind of "sweep mode" examining the spectrum' (p. 2).

3.1.3 AUDIO EDUCATION AND INDUSTRY TEXTS

Most audio engineering schools in Australia prescribe set texts that suggest students use the ICA method when learning to identify the sound of different frequencies. *Modern recording techniques* (Huber & Runstein, 2014), the set text for the three largest audio engineering schools in Australia (the SAE Institute, the Australian Institute of Music, and the JMC Academy) and at the SAE Institute campuses in United States, promotes the ICA method. Numerous prominent audio engineering texts used throughout audio engineering education institutions globally (Bartlett & Bartlett, 2008; Berman, 1999; Case, 2007; Crich, 2005; Gibson, 1996; Gottlieb, 2007; Hirsch & Heithecker, 2006; Izhaki, 2008; Katz, 2002; Keane, 2007; McGuire & Pritts, 2008; Rose, 2008; Shea, 2005; Strong, 2011, 2012), and in industry magazines (White & Houghton, 2008) also promote the ICA method.

3.1.4 SIGNAL PROCESSORS

In addition to audio engineering texts, several DAW and third-party plug-in manufacturers promote the ICA method by the inclusion of frequency-band isolation features in their equaliser plug-ins. The EQ3 7-Band and EQ3 1-Band equalisers included in Avid®'s Pro Tools® software feature a 'band-pass mode' that temporarily applies a band-pass filter at the selected frequency (Figure 18). The feature is designed to allow the user to 'solo and fine-tune each individual filter before reverting the control to notch filter or peaking filter type operations' (Avid Technology Inc., 2015, p. 14).

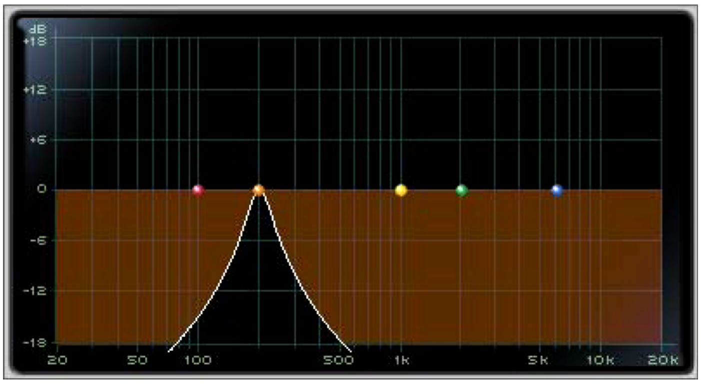

Figure 18. 'Band-pass' mode enabled on the EQ3 7-band equaliser supplied with Pro Tools® (Avid Technology Inc., 2015, p. 14).

The Massenburg Design Works® MDW Hi-Res Parametric EQ (Figure 19) features patented IsoPeak® (Massenburg DesignWorks LLC, 2009) technology, promoted as '[a] simple and fast tool to isolate that frequency you're looking for' (Avid Technology Inc., 2010, p. 1). Once IsoPeak® mode has been activated in the plug-in, the user guide suggests that the listener '"[s]weep" the frequency listening to the effect of the peak boost to find the artifact or area of interest' then 'adjust the Q and the Boost/Cut' as desired (Avid Technology Inc., 2010, p. 13). In demonstrating the five-band version of this equaliser, George Massenburg promotes the Isopeak® function's ability 'to quickly switch back and forth between a dip and a very narrow peak', allowing the user to find 'inharmonic artefacts in an acoustic guitar' (GML LLC, 2014).

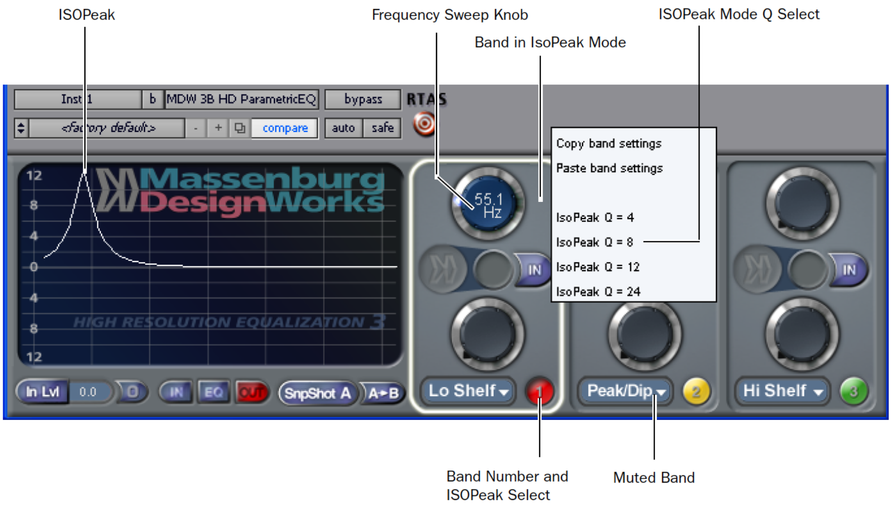

Figure 19. MDW® Hi-Res Parametric Equaliser in IsoPeak® mode. (Avid Technology Inc., 2010, p. 13)

The *Mastering with Ozone* guide details a similar feature included in the Ozone® 7 plug-in (Figure 20), which provides the user with 'an "audio magnifying glass"', designed to aid in 'pinpointing the location of a frequency in the mix without messing up your actual EQ bands' (iZotope Inc., 2015, p. 22).

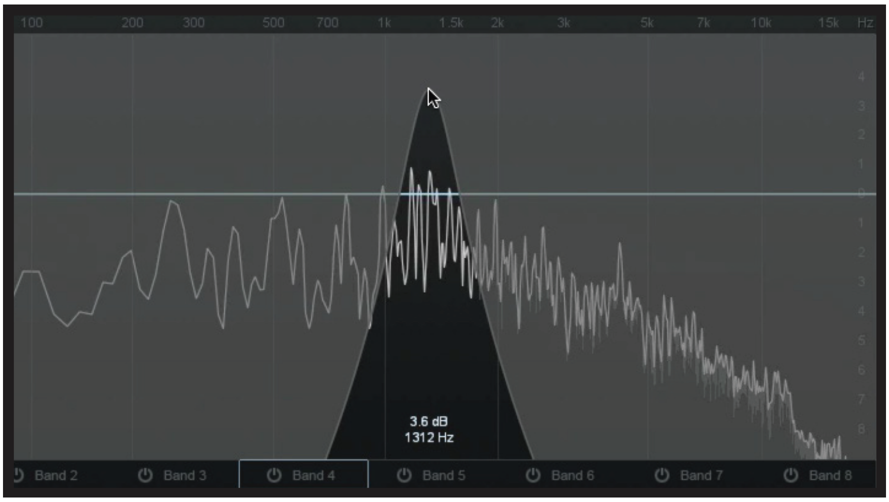

Figure 20. The 'audio magnifying glass' feature in Ozone® 7 (iZotope Inc., 2015, p. 22)

The AirEQ® equaliser features two modes that facilitate the ICA method, both of which are designed to allow the user to 'precisely hear the resonances [they] search for, while keeping the original gain of the bands' (Eiosis LLC, p. 41). The first mode, entitled Frequency Finder, temporarily sets the currently selected frequency's gain to maximum. The second mode, entitled Band Solo, 'solo[s] the output of the selected band while keeping the Gain and/ or Q at their maximum value' (ibid.).

By including a frequency region isolation feature, it follows that users will want to sweep the isolated frequency region, thus including such a feature promotes the ICA method by default. Plug-in designers encourage the ICA method most likely because of its ease of use compared to the ISA method particularly for novice users. The ability to isolate a frequency region is certainly not necessary to operate an equaliser, but it is an additional feature that can be sold to users.

3.1.5 Technical Ear Training Programs

Students enrolled in the Timbre Solfège TETP at the Fryderyk Chopin University of Music are instructed to use the ISA method, that is, 'to make informed judgments of timbre characteristics before manipulating the settings of the equalizer' (Rościszewska & Miśkiewicz, 2015, p. 2). This instruction is provided to students as the authors believe that 'such an approach develops the skill of timbre identification more effectively than sweeping throughout the frequency spectrum and matching the sound timbre by hit-and-miss' (ibid.).

Students enrolled in the Timbral Ear Trainer I program are 'trained to "track down" dips by actively "looking for" the spectral region where an octave or 1/3 octave is missing' (Quesnel, 2001, p. 47). This ICA-style method requires users to 'sweep' throughout the spectrum, requiring a large number of frequency parameter adjustments compared to the ISA method. Quesnel hypothesised that professional audio engineers would perform fewer adjustments than his students, as the professionals 'would be better at analyzing the spectral differences and would apply the right adjustments more directly' (ibid., p. 87), in contrast to the students who he thought 'would use more trial-and-error strategies and successive approximations to achieve a correct answer' (ibid., p. 88). However, when Quesnel examined the number of adjustments participants made to all the parameters of the test interface (including frequency), he found a statistically significant difference ($F(1,9) = 6.08$, $p \leq 0.05$) between the means of the two groups, even though the error bars in Figure 21 overlap. The mean number of adjustments for the professional audio engineers was 89, compared to the students' 58.

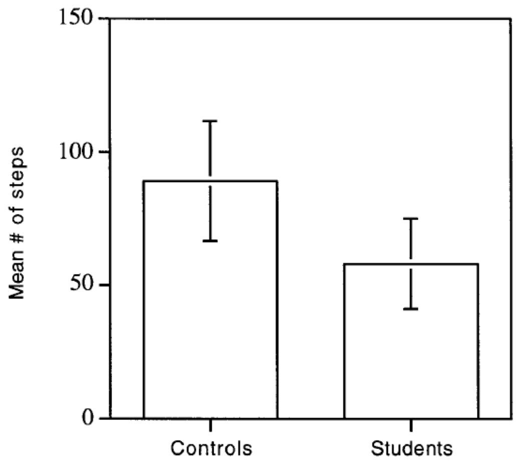

Figure 21. The mean number of adjustments for the professional and student participants (Quesnel, 2001, p. 90).

3.2 Establishing the Effectiveness of Technical Ear Training Programs

To establish whether a TETP is effective, one must look to the stated goals of the program; only if the program meets these goals can the researcher declare it effective. Conversely, in the absence of clearly stated goals, researchers are unable to know with any degree of certainty whether a TETP is effective. Fortunately, most TETPs for which published information exist do have clearly stated goals, but the mere existence alone of a training goal is insufficient. In this review, the researcher's argument was that to declare a TETP effective, the goals of the program must first meet three criteria: they must be *professionally relevant*; researchers must *measure performance relating to the goals;* and this performance must be *measured separately from the TETP, on a task that students have not previously trained with.* Only once these criteria are met should the researcher examine whether users' performance on a task relating to the goals of the TETP has changed because of the training. The following sections examine the goals of the eight student-targeted spectral-based TETPs for which published information is available, against these three criteria.

3.2.1 PROFESSIONALLY RELEVANT GOALS

Audio engineering education programs and the TETPs they commonly utilise are designed to train students to become competent audio professionals (Bielmeier, 2013; Davis & Parker, 2013; Scheirman, 2013; Tough, 2009, 2010). TETPs are specifically designed to increase students' exposure to critical listening scenarios via the repeated presentation of simulated professional audio tasks. Developers suggest that practising with their TETPs speeds up the development of certain skills that ultimately serve students in their future audio careers. It follows that if the goal of a TETP is to develop a certain skill, then for the TETP to be effective, not only must students develop the skill, but the skill must be relevant to students' future professional audio careers. Therefore, the first criterion that a TETP goal must meet is that it must be professionally relevant. A TETP is of little use to students if it does not train them to develop or refine a skill that is useful in their future audio careers. Based on the stated goals of their programs, it is evident that most developers of TETPs acknowledge this criterion. Indeed, it is uncommon to find a TETP with a stated goal that is *not* professionally relevant. The following section presents the eight student-targeted spectral-based TETPs, then an assessment of their goals against the criterion of professional relevancy.

3.2.1.1 Timbre Solfeggio / Solfège

Numerous goals have been published for the Timbre Solfeggio / Timbre Solfège (hereafter Timbre Solfège) TETP programs. These goals vary from general 'listening skills'-type goals, to specific goals relating to the perception of timbre. Rakowski and Trybuła provided the least specific goal – that the program was designed to 'train[] the aural dispositions of future Tonmeisters' (1975, p. 4). Miśkiewicz (1992, p. 621) and Hojan and Rakowski (1999, p. 625) also cited the 'develop[ment] of listening skills' as the goal of the program. Other authors, however, have been more specific when defining the goals of iterations of the Timbre Solfège programs. The goals for this TETP are divided into two categories: those that relate to timbre perception, and those that relate to language.

The first category relates to the development 'of timbre perception skills' (Letowski, 1985, p. 241), specifically the development of both sensitivity and memory for timbre (Rakowski & Trybuła, 1975). Letowski (1985) asserted that the goal of timbre solfège was to 'increase timbre sensitivity and improve (expand) timbre memory' (p. 241). Miśkiewicz (1992) stated that the program trained 'the skill of memorizing and identifying changes of timbre that arise from certain modifications of the spectrum' (p. 621). Letowski and Miśkiewicz (1995) suggested that the program developed 'listening skills that include acute sensitivity to differences among sounds and accurate auditory memory' (p. 917), while Rościszewska and Miśkiewicz suggested the program trained 'the ability to detect and memorize the characteristics of timbre produced by modifications of the sound spectrum, in particular, changes of the spectrum envelope shape' (2014, 2. Teaching methods for developing the acuity of memory for timbre, para. 1), and developed 'auditory cues for identification of the sound spectrum characteristics' (2015, p. 2).

The second goal category of the Timbre Solfège program addresses the use of language within the TETP. Letowski stated the secondary training goal of the program 'was to develop a set of qualitative descriptors to provide unequivocal codes (language) for exchanging information regarding timbre perception' (1985, p. 241), 'develop[ing] in the students the command of the vocabulary used for timbre description' (2015, p. 2).

These goals for the Timbre Solfège TETP program meet the first criterion in that they are professionally relevant. The ability to discriminate changes in timbre is paramount within a professional audio production environment as it underpins critical listening – a key skill in the effective operation of almost all audio signal processing devices. Having first discriminated changes in timbre, the ability to communicate these changes to fellow audio engineers, producers and musicians is a vital part of the audio production process. As such, developing an appropriate vocabulary to communicate timbre attributes and changes in timbre also meets the criterion of professional relevancy.

3.2.1.2 Timbral Ear Trainer I & II

In 1990, Quesnel published a thesis entitled *A computer-assisted program in timbral ear training – a preliminary study*. This document introduced the Timbral Ear Trainer I TETP, which was 'a partial implementation' of the Timbre Solfège program

(Quesnel & Woszczyk, 1994, p. 2). In the thesis, Quesnel lists four goals for the training program:

1) develop in long-term memory a set of timbral category references against which sounds to be analyzed can be compared;
2) provide the student with a knowledge of the relationships between perceived timbral changes and the physical parameters of sounds responsible for these changes and under her/his control;
3) develop perception acuity of fine timbre differences (paired comparisons); this is the ability to compare a given timbre with an absolute reference; and
4) develop the ability to describe the timbre of a given sound (absolute judgement); this is the ability to compare a given timbre with a personal internal reference. (Quesnel, 1990, p. 83)

Quesnel (1990) concluded that the training program 'could help improve timbre perception acuity and memory of sound engineers' (p. 104). In a 1994 paper detailing the training program, Quesnel and Woszczyk stated that the program had three main goals:

5) developing long-term memory for a basic set of timbral categories;
6) progressively developing the listener's sensitivity to subtle timbre differences; and
7) developing analytic listening skills for the evaluation of complex spectral changes in terms of individual resonances, specifying for each the magnitude, Q, and center frequency. (p. 2)

In a later paper, also published in 1994, Quesnel stated that the training program was aimed at developing (8) 'memory for a basic set of timbre categories', (9) 'sensitivity to small timbre differences', and (10) 'listening strategies for the evaluation of timbre' (p. 38). In 1996, Quesnel detailed a revised version of the training program, (11) 'aimed at developing the ability to evaluate quantitatively timbre differences between sounds' (1996). In his 2001 PhD thesis, Quesnel stated that 'an important goal of this [thesis] is to better understand the actions a listener is performing during timbral manipulation task [*sic*] and to subsequently use this information to improve the usefulness and efficiency of the training' (p. 7). However, in the thesis, the only stated goal of the TETP itself was (12) 'to develop memory for a set of octave resonance references centered from 63 to 16000 Hz' (p. 45).

The Timbre Solfège program section above addresses the third, sixth and ninth goals of the Timbral Ear Trainer program – the development of timbre discrimination skills (timbre sensitivity) – and shows their professional relevance. Likewise, the section addresses the fourth goal of the Timbral Ear Trainer program relating to the communication of changes in timbre, demonstrating that it too is professionally relevant.

The first, fifth, eighth and 12th goal of the Timbral Ear Trainer program, namely, storing timbre category references (frequencies) in long-term memory, is primarily useful to a professional audio engineer when operating an equaliser. When aiming to identify the frequency of a bandwidth of concern (for example, when identifying the centre frequency of an unwanted resonance captured during the recording process), possessing a long-term memory of timbral category references (frequencies) is advantageous. These references provide anchor points throughout the spectrum that aid in the identification of the frequency region of concern. Using these spectral anchor points, the audio engineer first estimates which of the octave reference frequencies (stored in long-term memory) is closest to the centre frequency of the perceived resonance. The audio engineer then assesses whether the resonance is higher, lower, or equal to the reference frequency. By using internal timbre category reference frequencies, a specific starting point from which to home in on the frequency of concern replaces what would otherwise be a general approximation. As such, these goals meet the first criterion of being professionally relevant.

Following on from the ability to discriminate changes in timbre and then compare the perceived timbre to timbre category references in long-term memory, goals two, seven and eleven of the Timbral Ear Trainer program address the ability of the student to quantify the difference between two sounds by identifying the parameters of the individual resonances. When manipulating the spectrum using a signal processor, the user ultimately achieves the desired sound by modifying one or more parameters of a signal processing device such as an equaliser. When seeking to remove an unwanted resonance, for example, the ability to associate perceived spectral irregularities with the physical attributes of sound and the associated parameters of a signal processor is integral to achieving the desired sonic outcome. To remove an unwanted resonance using an equaliser, the engineer must quantify the parameters of the resonance, including frequency, gain, and Q. Quantifying the attributes of the perceived spectral irregularities in this manner, as opposed to ‘blindly’ modifying the parameters of a signal processor until achieving the desired sonic outcome, results in a more efficient use of the processor. This is comparable to the difference between the ISA method and the ICA method when identifying a resonance.

The 10th goal, the development of 'listening strategies for the evaluation of timbre', is more general, but it remains professionally relevant if the developed listening strategies are useful.

3.2.1.3 Spectral Ear Trainer

In a 1993 paper, Brixen detailed the Spectral Ear Trainer TETP, developed at the Danish Acoustical Institute, designed to train students in 'the recognition of spectral changes' and to rapidly achieve 'the ability of being a "Human Spectrum Analyzer"' (pp. 1, 3). Recognition (identification) of spectral changes requires the use of timbre discrimination skills, namely timbre sensitivity and the storage and recollection of timbre category reference frequencies in long-term memory. In the preceding sections, both skills are shown to be professionally relevant.

3.2.1.4 Technical Listening Training

The Technical Listening Training TETP was designed to improve students' auditory sensitivity and improve their 'understanding of the relationship between acoustic properties and auditory impression' (Iwamiya et al., 2003, p. 1). The authors define 'auditory sensitivity' as 'the ability to discriminate between different sounds' and 'identify various types of perceived differences, such as differences in pitch, loudness and timbre' (Iwamiya et al., 2003, p. 28). Arai et al. (2006) state the program was 'designed to help listeners develop a sensitive ear for the general sense of sound including sound discrimination' (p. 347). The authors dissect 'professional listening' into three abilities; 'the ability to discriminate between different sounds, the ability to correlate the auditory difference with the physical properties of sounds, [and] the ability to imagine the proper sounds when given the acoustic properties of the sounds' (Kawahara et al., 2013, p. 2).

The preceding sections present the argument that timbre discrimination is a professionally relevant skill for audio engineering students to develop, along with the ability to identify (quantify) the difference between two sounds (via the correlation of perceived differences with the physical properties of sound). The ability to imagine or hear in one's mind a sound when provided with physical properties alone directly relates to the storage and associated recollection of timbre category reference frequencies stored in long-term memory. When endeavouring to identify the centre frequency of a low-frequency resonance, for example, the audio engineer compares the perceived sound to memorised timbre category reference frequencies. The 'sound' of each of these reference frequencies is stored alongside the associated physical parameters of the sound. For example, the 'sound' of a resonant boost at 500 Hz is stored alongside the frequency of the boost itself. When hearing a resonant boost at 500 Hz, the audio engineer searches long-term memory for a matching sound, and if found, recalls the frequency stored alongside the sound itself. The storage and recollection of these sound–parameter pairings is integral to an audio engineer's ability to identify the frequency of a resonant boost, for example. This skill is therefore a professionally relevant skill for audio engineering students to develop.

3.2.1.5 Web-based Technical Listening Training

The stated goals for the Web-based Technical Listening Training TETP are identical to the goals of the previously detailed Technical Listening Training program; namely, the web-based program is 'designed to improve auditory sensitivity' (Nishimura, 2006a, p. 3071).

3.2.1.6 Beijing Union University Training

The TETP delivered at Beijing Union University was designed with the goal of improving 'subjects' discrimination of sound attributes', (Liu et al., 2007a, p. 1). That the ability to discriminate changes in timbre is being professionally relevant is demonstrated in previous sections.

3.2.1.7 Personalised Timbral Ear Trainer

The Personal Timbral Ear Trainer program was designed 'to improve the listening of sound engineers so that they can skilfully modify and edit the structure of sound' (Kaniwa et al., 2011). The program employs an absolute identification task, with the goal of increasing:

> a trainee's ability to identify a modified spectrum, describe it in terms of technical parameters (center frequency, Q, and gain), and eventually build a long-term memory of the internal reference on which a trainee will rely for future identification tasks. (Kaniwa et al., 2011).

The ability to discriminate between two sounds, the subsequent identification of the parameters of the modified sound, and the storage in long-term memory of timbre category reference frequencies are shown in sections above to be professionally relevant goals for a TETP.

3.2.1.8 University of Lethbridge Training

The Noise Spiral applet developed at the University of Lethbridge was 'developed to facilitate the learning process of frequency band recognition' (Schaller & Burleigh, 2015, p. 4). The goal of 'teach[ing] frequency band recognition' (ibid., p. 1) is professionally relevant, as argued previously.

3.2.1.9 Summary

The first criterion that a TETP goal must meet is that it must be professionally relevant and as such, the goals of the eight selected TETPs were presented and reviewed against the first criterion of professional relevancy. As shown, the stated goals of all TETPs meet this criterion. However, it is subsequently argued that to consider a TETP effective, the stated goals of the training program must also meet a second criterion: performance relating to the goals of the program must be measured.

3.2.2 Measuring Performance Relating to Goals

Once a TETP goal has met the first criterion of professional relevancy, it must then meet a second criterion concerning the measurement of user performance. A TETP may have a clearly stated, professionally relevant goal such as 'develop timbre discrimination ability', but in order to state that the goal has been met, users' performance on a task relating to this goal must be measured. In this example, it is not possible to directly measure a user's skill level in discriminating timbre; rather, the user must demonstrate this skill through performance on a task. It is this performance that provides an indication of the level of skill the user possesses. Hence, to ascertain whether a user's skill level has changed as a result of undertaking the TETP, one must measure their performance on a task that demonstrates this skill. To establish a baseline level of performance, ideally the researcher gathers performance data before and after users undertake the TETP to enable detection of any improvement in a skill. Without both sets of data, the researcher lacks evidence to support the claim that a user's skill level has increased as a result of undertaking the TETP (this pre/post-training methodology is discussed further in Section 3.3). It follows that if the goal of the training program is to develop a certain skill, without performance data on a task that demonstrates improvement of this skill, one cannot claim that the training program meets this goal and subsequently that the program is effective.

To support claims that users' timbre discrimination skills have improved because they trained using a TETP, many authors present performance data on a task they claim demonstrates this skill. The following section presents a review of the eight TETPs in relation to the second criterion of measured performance, to establish whether authors present goal-related performance data to support claims of their training program's effectiveness.

3.2.2.1 Timbre Solfeggio / Solfège

In their seminal 1975 paper, Rakowski and Trybuła claim that students' sensitivity and memory for timbre 'can certainly be trained and improved' after training with the Timbre Solfège TETP, which the authors claim 'prove[d] to be an efficient form of training the aural dispositions of future Tonmeisters' (pp. 3-4). However, the authors do not indicate what these conclusions were based on and don't refer to, or present, associated performance data.

Letowski (1985) claims that both sensitivity and memory for timbre 'can be developed successfully by training on a systematic basis' using an updated version of the Timbre Solfège TETP (p. 240). Letowski went further, claiming that 'almost every student improve[d] greatly' in both their sensitivity and memory for timbre 'after several laboratory classes' (p. 243). He maintained that 'in some cases the precision in the description of the physical properties of the sound based on the timbre evaluation alone [was] astonishing' (ibid.). These claims were based on anecdotal evidence gathered during the delivery of the TETP.

Miśkiewicz (1992) details an 'expanded and updated version of the [Timbre Solfège] program', and maintains that 'technical listening skills may be improved by systematic training' (p. 624). This claim was based on the author's experience in delivering the TETP and reports from graduates in the field (p. 621).

Letowski (1995) also claims that the Timbre Solfège TETP was successful in developing 'acute sensitivity to differences among sounds and accurate auditory memory' (p. 917). This claim, and the claim that the 'listening abilities necessary for analyzing timbre may be developed and considerably improved by systematic training', are based on the experience of the various researchers who have delivered the program over the years (p. 920).

In 2011, Rościszewska presented the results of a test designed to establish the effectiveness of the Timbre Solfège TETP. This paper was the first to provide empirical, performance-based evidence in support of the effectiveness of the training program in the 30 years since the original implementation. The test was administered to successive 'groups of students over a period of 10 years' and was designed to 'show that sensitivity to changes in timbre may be considerably improved by systematic training' (2011, p. 1). The test, delivered at the beginning and end of the Timbre Solfège program, presented students with a 2AFC task in which they were required to indicate whether a series of two sounds presented pairwise were different or identical. Performance data from the test showed that the 'average percentage of correct responses obtained at the end of the academic year is higher than at the beginning of training' with one exception (p. 4).

In 2014, Rościszewska and Miśkiewicz detailed a new test that was 'introduced for the assessment of timbre memory acuity' for participants in the Timbre Solfège program (2014, 'Purpose of the test'). The test, which was delivered at the beginning and end of the Timbre Solfège program, required students to 'adjust the frequency of a formant in a noise burst to duplicate from memory the timbre of a reference noise burst' (ibid.). The authors added that 'the accuracy of formant frequency adjustment' and 'the acuity of students' short-term memory for timbre' improved as a result of the training, noting that 'in most cases the improvement was substantial in its amount' (ibid., 'Results and discussion'). The authors concluded that 'most students improve[d] both the[ir] short-term and long-term memory for timbre during the course' (ibid., 'Conclusions').

In 2015, Rościszewska and Miśkiewicz again presented performance data for the Timbre Solfège TETP in support of clams that the program 'improves memory for timbre and develops the ability of associating the perceived characteristics of timbre with the spectral properties of sounds' (p. 1). In this paper, the authors present results for three different tests and detail five tests overall. The first test is administered throughout the Timbre Solfège training program as it is designed to monitor students' progress throughout the training program. At the beginning of the test, students are required 'to identify one modification of the sound spectrum in each test trial: the bandwidth cut-off frequency or the formant frequency' (p. 4). The test increases in difficulty at later stages in the training program, with students required to identify 'two spectral modifications, including low- and high-frequency bandwidth cut-offs and formants' (p. 4). At the end of the training program, a second test is administered whereby students are first required to again identify the parameters of two spectral modifications introduced into music recordings; following this, a third test requires students to perform a matching task. The fourth test detailed in the paper, designed to test students' ability to discriminate timbre, is identical to the test detailed by

Rościszewska (2011), and the fifth test is identical to that detailed by Rościszewska and Miśkiewicz (2014). Based on the results of these tests 'as well as extensive practical experience gained from teaching' the program, Rościszewska and Miśkiewicz (2015) conclude 'that timbre evaluation skills may be effectively developed by systematic training' (p. 6).

Of the eight papers reviewed that detail the Timbre Solfège program, three present performance data relating to the training program's goals. Not only are the stated goals for the Timbre Solfège program measurable, but several authors have recently endeavoured to collect performance data relating to these goals in an attempt to establish the effectiveness of their TETP. As such, the goals for the Timbre Solfège TETP meet the second criterion in that performance relating to these goals has been measured.

3.2.2.2 Timbral Ear Trainer I & II

Quesnel collected preliminary performance data for nine students during the development of the Timbral Ear Trainer I TETP. He recorded their performance on a comparative listening exercise at the beginning and the end of the six-month training period, drawing two conclusions. Firstly, he noted that all students' performance (percentage of correct answers) improved, and quite significantly in some cases (Quesnel, 1990). Secondly, he noted that 'individual ability within the group varied substantially at the beginning whereas it was much more homogeneous at the end' (p. 97). Quesnel also examined response times, concluding that 'the time spent on each problem within the exercise' was shorter at the end of the training program than the beginning (p. 100), and decreased for most students by a factor of two. However, he did not find a 'clear relationship between practice time and performance' (p. 100).

In 1994, Quesnel and Woszczyk claimed that the Timbral Ear Trainer I program was an ‘effective training method for the development of listening skills’ (p. 1). This claim was based on the performance of seven students who undertook a pre-post-training test, delivered two years after Quesnel’s initial study. Speed and accuracy data was collected through three tests administered before and after training with the Timbral Ear Trainer I program: a comparative listening exercise, a bring to flat exercise, and an absolute identification exercise. For the comparative listening exercise, ‘most students obtained a perfect score at the beginning of the training period’, with the two students who did not achieving slight improvement at the end (p. 5). For the bring to flat exercise, the authors noted that ‘individual ability within the group varied substantially at the beginning of the training whereas at the end, all students obtained a perfect score’ (ibid.). For the absolute identification test, the authors noted that ‘significant improvement occurred for all students’ (ibid.). Quesnel and Woszczyk also introduced the Timbral Ear Trainer II TETP, which was under development at the time.

Quesnel’s 1994 paper provided an overview of the Timbral Ear Trainer II TETP but did not present performance data. In 1996, Quesnel published a paper detailing the Timbral Ear Trainer II program, stating that ‘experiments [were] in progress in order to evaluate the effectiveness of the training’ (Current and Future Work, para. 1).

In 2001, Quesnel published the results of an experiment conducted as part of his PhD research. The experiment involved two groups of participants: students who had undertaken the Timbral Ear Trainer II TETP, and professional audio engineers who had not. Both groups of participants ‘had to modify the comparison stimulus (each trial equalization) [using a parametric equaliser] until it matched the "flat", unprocessed reference, which was the same for each trial’ (2001, p. 75). Quesnel concludes that overall, the students 'performed significantly better' than the professional audio engineers on 11 measures (p. 96). Quesnel argues that the training program was effective because the students outperformed the professional engineers.

In each of the various iterations of the Timbral Ear Trainer program, Quesnel collected performance data to support claims of the TETP's effectiveness. As such, the goals for Timbral Ear Trainer I and II meet the second criterion in that performance relating to the goals of the training programs was measured.

3.2.2.3 Spectral Ear Trainer

Brixen (1993) suggests that 'to some extent it is possible to establish an ability of being a "human spectrum analyzer"' after undertaking the author's Spectral Ear Training TETP (p. 9). This claim was based on the author's experience in delivering the training program, and to date no performance data has been presented. As performance data relating to the goals of the Spectral Ear Training program were not presented or measured, the program's goals do not meet the second criterion. It is insufficient to base claims of the effectiveness of a TETP solely on anecdotal evidence or observations of improvements in students' performance during the training program.

3.2.2.4 Technical Listening Training

The authors of the Technical Listening Training program claim that 'students [who undertake the training program] improve their sound sensitivity and understanding of the relationship between acoustic properties and auditory impression' (Iwamiya et al., 2003, p. 27). To accompany their 2004 paper, Iwamiya et al. presented a demonstration of their TETP, claiming that participants 'will experience that auditory sensitivity may be improved through this program' (Kawahara, Takada, Iwamiya, Nakajima, & Ueda, 2004a). However, in their 2003 and 2004 papers that detail the training program, no performance data was provided to support these claims (Iwamiya et al., 2003; Kawahara et al., 2004b). In a 2013 paper, Kawahara et al. presented the average correct answer ratios of participants as evidence of the effectiveness of the Technical Listening Training program offered at Kyushu University; however, only performance data for the 'identification of sound pressure level[s]' was presented (2013, p. 4). Performance data relating to the frequency-spectrum-related components of the TETP, that is, the goal of training students to identify perceived differences in timbre, was not presented. As there is no performance data published on this goal, the TETP goal does not meet the second criterion of measured performance.

3.2.2.5 Web-based Technical Listening Training

Authors of the Web-based Technical Listening Training program affirm that participants demonstrated 'significant improvements' in their frequency discrimination abilities (Nishimura, 2006a, pp. 3072–3073). However, to date performance data relating to this training program goal has not been published (Arai et al., 2006; Nishimura, 2006a, 2006b, 2013, 2015). As such, the goals of the TETP do not meet the second criterion.

3.2.2.6 Beijing Union University Training

The authors of the TETP offered at Beijing Union University claim that after undertaking the training program 'most subjects' improved their ability to discriminate attributes of sound (Liu et al., 2007a, p. 1). This claim was based on the performance of participants in a series of tests, one of which required participants to identify the parameters (centre frequency and gain polarity) of a parametric filter applied to pink noise and music. The tests were administered at the beginning and end of the TETP, with the authors noting that 'subjects made great progress after training' with the average mark on the composite exam increasing after undertaking the training program (p. 3). In terms of specific results regarding the parametric filter identification task, only the post-training scores are presented, making it impossible to compare pre/post-training scores.

It is evident that performance relating to the participants' ability to discriminate frequency response irregularities in the Beijing Union University TETP was measured and partial performance data were presented. As such, the goals of the TETP offered at Beijing Union University meet the second criterion.

3.2.2.7 Personalised Timbral Ear Trainer

The creators of the Personalised Timbral Ear Trainer program Kaniwa et al., (2011) present results from an initial study of the effectiveness of an adaptive TETP. In order to establish the effectiveness of the program, the authors designed an experiment that compared the performance of a control group on a TETP, to that of an experimental group. The control group trained using a TETP that generated random questions, whereas the experimental group trained using a TETP that generated questions 'using the weighted random function [that was] dynamically updated according to the trainee's previous training scores' (Kaniwa et al., 2011, 4. Experiment to Evaluate Proposed System, para. 1). The authors find that 'the average correct answer rate of the Proposal group was higher than that of the Conventional group' (ibid., 5.1 Analysis by Mean and Standard Deviation, para. 1). The authors conclude that their adaptive TETP 'could assist trainees in improving their ability to identify differences more effectively', as 'adaptive feedback would assist a trainee in acquisition of timbre-identification ability' (Kim et al., 2013, p. 424). Kim suggested that based on the 'importance of individualized training for acquisition of spectrum-identification and spectrum-matching skills', adaptive training tasks 'should be implemented by default into technical ear training programs used in audio production industry and education' (2015). As the authors measured performance on a task relating to the goal of the TETP, the goals of the Personalised Timbral Ear Trainer TETP meet the second criterion.

3.2.2.8 University of Lethbridge Training

The authors of the TETP delivered at the University of Lethbridge made no claims as to the effectiveness of their programs other than stating that 'the development of this skill [frequency band recognition] can be accelerated by using tools for technical ear training' (Schaller & Burleigh, 2015, p. 4). This TETP does not meet the second criterion as performance data was not collected.

3.2.2.9 Summary

To ascertain whether a user's skill level has changed as a result of undertaking the TETP, one must measure their performance on a task that demonstrates this skill. As such, the goals of the eight selected TETPs have been compared against the second criterion, which a TETP's goals must meet in order for the TETP to be declared effective. For the majority of programs reviewed, authors collected performance data relating to the goals of their TETP in order to support claims of their program's effectiveness. In the next section, the specifics of how the authors established the effectiveness of their TETP are reviewed.

3.2.3 Measuring Performance Externally to a Technical Ear Training Program

This researcher asserts that for a TETP to be considered effective, the goals of the TETP must first meet two criteria: the stated goals must be professionally relevant and performance relating to the goals must be measured. In the next section, a final criterion is proposed, about which there is some controversy in the field. It is proposed that for a TETP to be considered effective, after meeting the first two criteria, performance relating to the goals of the program must be measured separately from the TETP on a task that is either completely novel to students or significantly different from a previously experienced task. This section presents a review of how authors of TETPs for which performance data is available ascertained the effectiveness of their TETP, in the context of this third criterion. The TETPs reviewed in this section are limited only to those that met the second criterion, as one cannot review data that was not collected. As such, the section addresses only four of the eight TETPs reviewed previously. The tests used by authors to establish the effectiveness of their TETPs were reviewed first. The task undertaken by students in these tests were then compared against the tasks used in the TETPs.

3.2.3.1 Timbre Solfeggio / Solfège

Rościszewska was the first author to publish performance data relating to the Timbre Solfège TETP. The paper, published in Polish, 'present[s] results of a test used to assess the progress of sound engineering students in timbre discrimination tasks, at various stages of the *Timbre Solfège* course' (Rościszewska, 2011, p. 1). Students completed the test before the commencement of the Timbre Solfège TETP, then again at the completion of the program after two semesters of training. The test consisted of 44 trials, each of which contained sounds presented pairwise. Students performed a 2AFC task in which they indicated whether the two presented sounds were different or identical. For the pairs of sounds that differed, the spectrum was modified via the introduction of a parametric boost applied at 125, 500, 2k or 8k Hz. Gain values for the boosts were 2, 3.5, 5 or 6.5 dB. The Q value of the filter was not detailed. Thirty-two pairs containing differing sounds were presented, along with 12 pairs containing identical sounds. The stimulus used was classical music, specifically ‘chords performed by a string orchestra in the dynamics forte’ (p. 2). The test was administered over loudspeakers at a level of 70 phons (Rościszewska & Miśkiewicz, 2015, p. 4). The mode of playback (mono or stereo) was not indicated. The test was administered to successive groups of students over a period of 10 years. The paper presents pre/post-training data for 97 students.

The test recorded the percentage of correct answers and the percentage of ‘false alarms’ (false positives – trials in which the student indicated the pairs were different, but they were not). These percentages were used to calculate a final result (*P*) by ‘normalizing the number of correct answers according to the formula’:

$$P[\%] = \frac{x - a}{100 - a} . 100$$

where x = the percentage of correct answers

α = percentage of false alarms

Equation 1. The normalising formula (Rościszewska, 2011, p. 3)

Students’ test results at the commencement of the training program in October and at the end of the training in May the following year are presented in Figure 22.

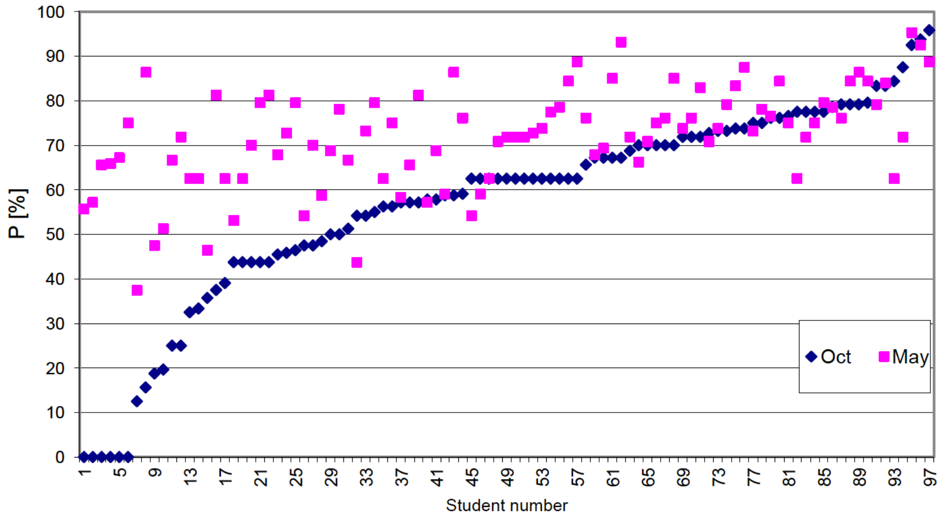

Figure 22. P[%] results for 97 participants before (Oct) and after (May) undertaking the Timbre Solfège TETP (sorted by October. score). (Rościszewska, 2011, p. 3)

Rościszewska (2011) noted that although 18 of the 97 participants (19%) 'received a worse result on the second test than the first', the magnitude of performance decline 'is small (an average of 7 percentage points)' (p. 3). Figure 23 shows the correlation between the P[%] in October and May.

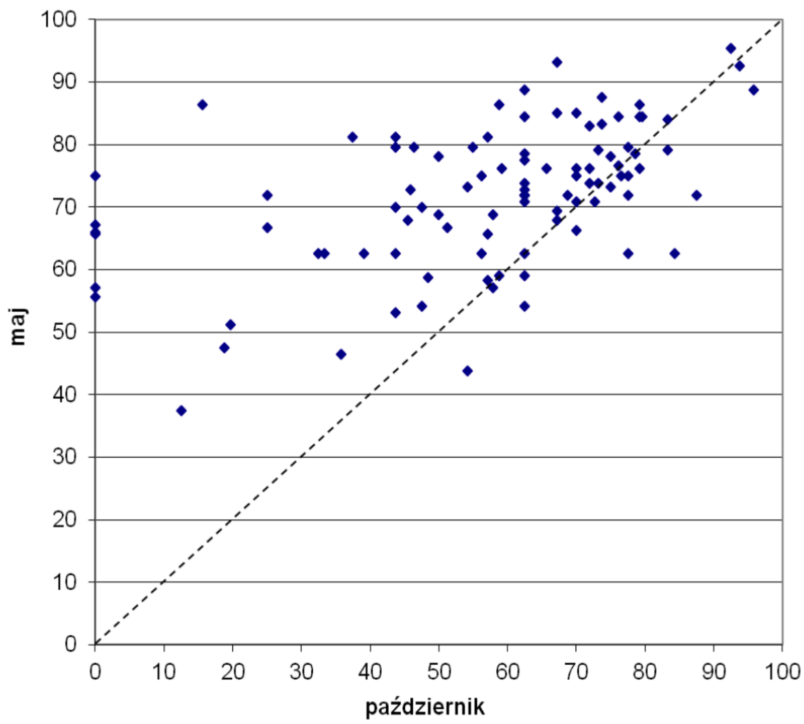

Figure 23. Correlation between the October (październik) and May (maj) results. (Rościszewska, 2011, p. 4)

The average percentage of correct responses, P (ignoring 'false alarms'), for all students versus the various filter gain levels for each frequency is shown in Figure 24. Rościszewska (2011) 'noted that in all cases, the largest increase in the number of correct answers is between the lowest values of gain control: 2 dB and 3.5 dB' (p. 5). Note the decrease in performance at 125 Hz. Rościszewska explains the decrease as being due to frequency masking caused by a strong fundamental present in the stimulus at 100 Hz.

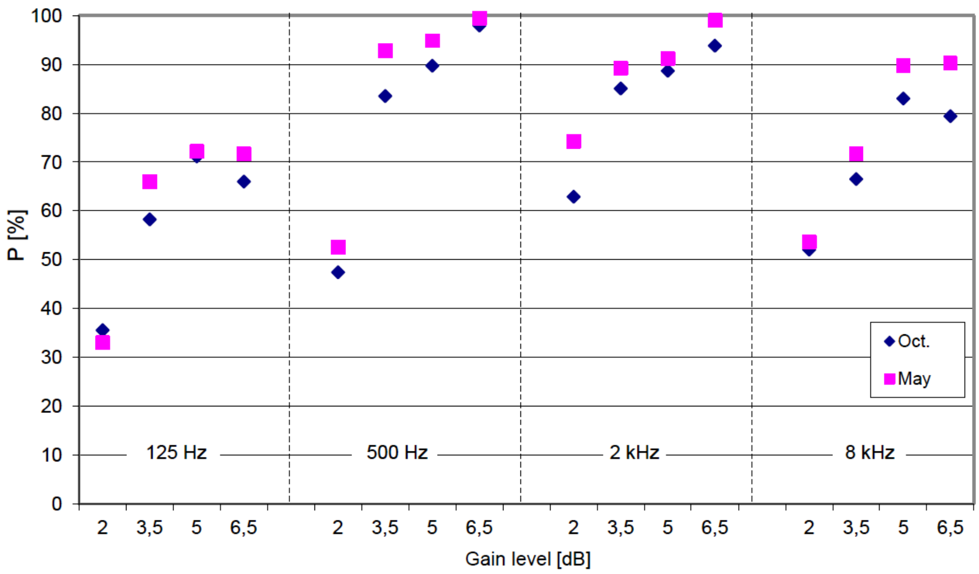

Figure 24. P[%] for each filter gain level and frequency. (Rościszewska, 2011, p. 5)

Rościszewska also examined the effectiveness of the TETP based on students' initial skill level (performance on the first test). After taking the initial test, students were divided into two groups. Group A (n=30) consisted of students whose performance on the initial test did not exceed P = 50% and Group B (n=67) consisted of the remainder of the students. The means and standard deviations for the two groups are shown in Figure 25.

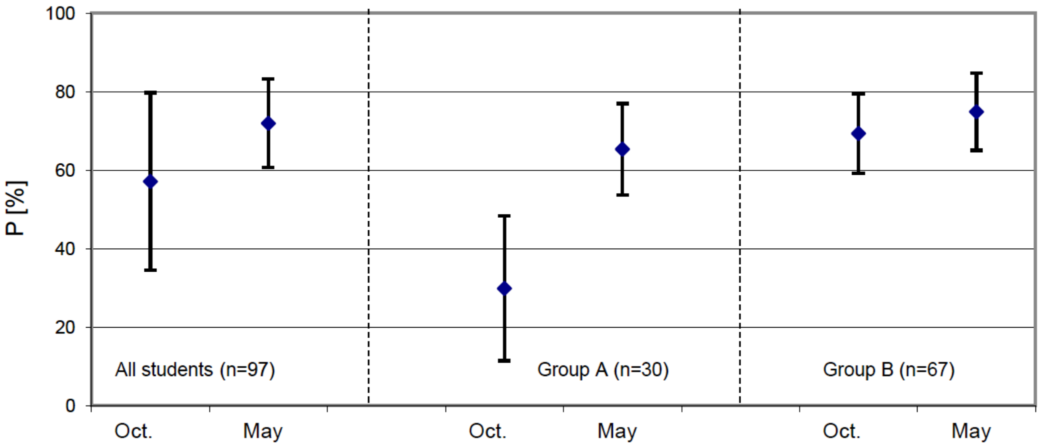

Figure 25. Mean and standard deviation P[%] results. (Rościszewska, 2011, p. 5)

These results show that the mean score for Group A improved from 30% to 65%, whereas Group B's performance increased from 70% to 75%. Rościeszewska (2011) also noted that the 'dispersion of test results conducted in May for both groups are much lower than in October, reflecting the greater stability of the responses of students at the end of the year' (p. 5). Finally, Rościszewska presented the average percentage of false positives for both groups (Figure 26). The mean percentage of false positives for all students decreased, with Group A's percentage decreasing from 47% to 21% (p. 6).

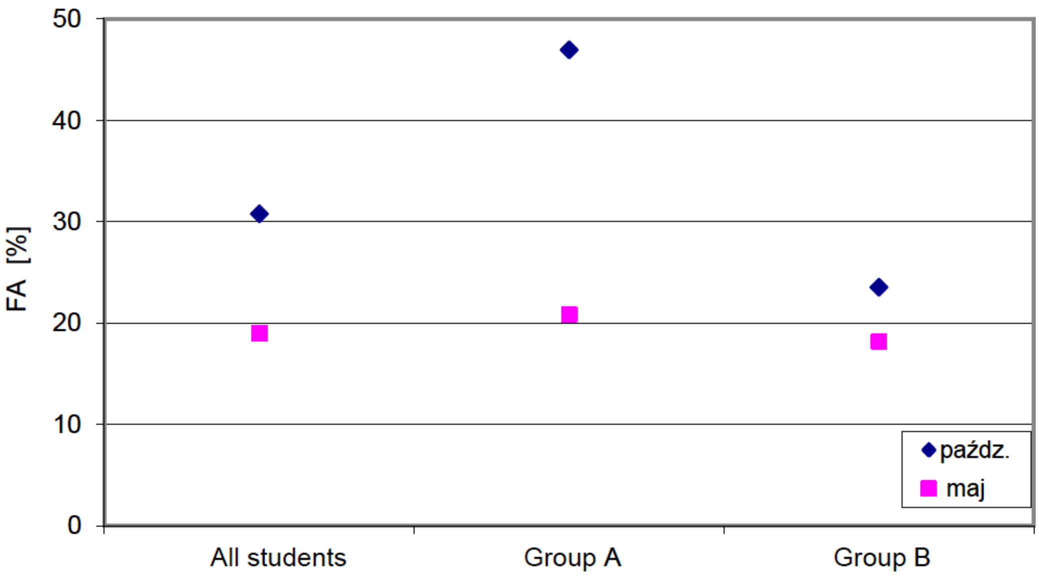

Figure 26. Percentage of false alarms for the October (październik) and May (maj) results. (Rościszewska, 2011, p. 6)

The test ‘for the assessment of timbre memory acuity’ detailed by Rościszewska and Miśkiewicz (2014, ‘3.1 Purpose of the test’) required students to perform a matching task using a parametric equaliser applied to a one-second noise burst. This test was administered twice, at the commencement and conclusion of the first year of the Timbre Solfège TETP. For each trial, students were presented with a burst of noise (the ‘standard’) containing a spectral irregularity introduced using a parametric filter at ‘125, 250, 500, 1000, 2000, 4000, [or] 8000 Hz’ (2014, ‘3.2. Procedure of testing’) with a gain of 12 dB and Q of 5.0. After a pause, students were then presented with an unmodified noise burst (the ‘variable’), and were required to modify the spectrum of this flat signal using a parametric equaliser to duplicate the sound of ‘the standard’. The test had two differing conditions defined by the duration of the pause between signals. The first condition featured a pause of five seconds, the second a pause of three minutes. The tests were administered in stereo using headphones at a ‘level of about 65 phons’ (2014, ‘3.2. Procedure of testing’).

Performance data from 39 students for conditions one and two are presented in Figures 27 and 28 respectively. The shaded square in the figures indicates ‘deviation[s] of 600 and less cents (half of an octave)’ (Rościszewska & Miśkiewicz, 2014, ‘3.3. Results and discussion’). The authors note that ‘in most trials the formant frequencies were adjusted more accurately after training’ for the first condition, with 63% of the points lying below the diagonal line (ibid.). The authors conclude ‘that the acuity of students’ short-term memory for timbre has improved after training’ (ibid.). For the second condition, the authors again positively conclude ‘that the acuity of long-term memory for timbre has improved after training’ based on the 64% of the points lying below the diagonal line (ibid.). Following these comments, the authors note that the raw data in

Figures 27 and 28 ‘suggest that the improvement of the acuity of timbre memory is moderate at the end of training as it is manifested in only about 2/3 of the trials in the two conditions of the test’ (ibid.).

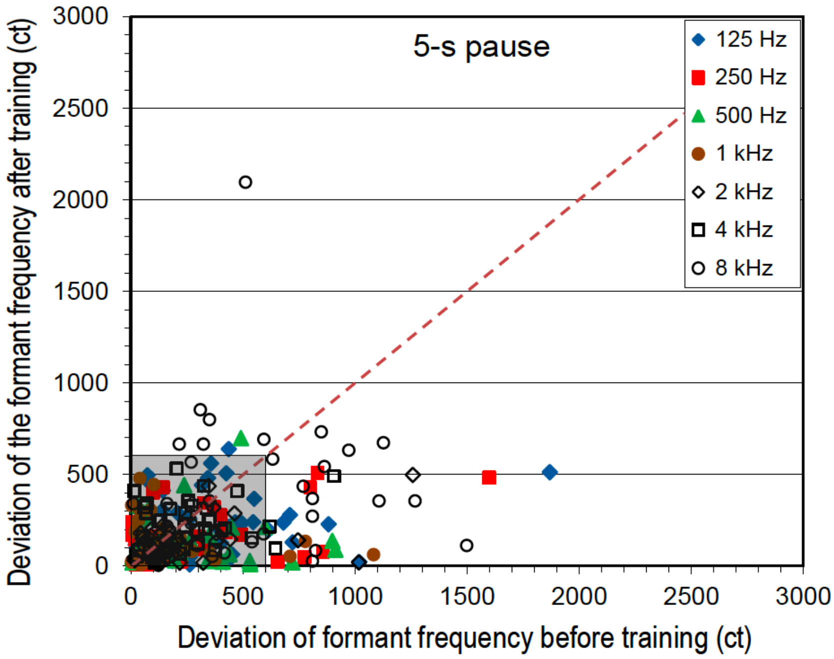

Figure 27. Deviation from the centre frequency in cents, for the first condition (five second pause), before (abscissa) and after (ordinate) training. (Rościszewska & Miśkiewicz, 2014, ‘3.3. Results and discussion’)

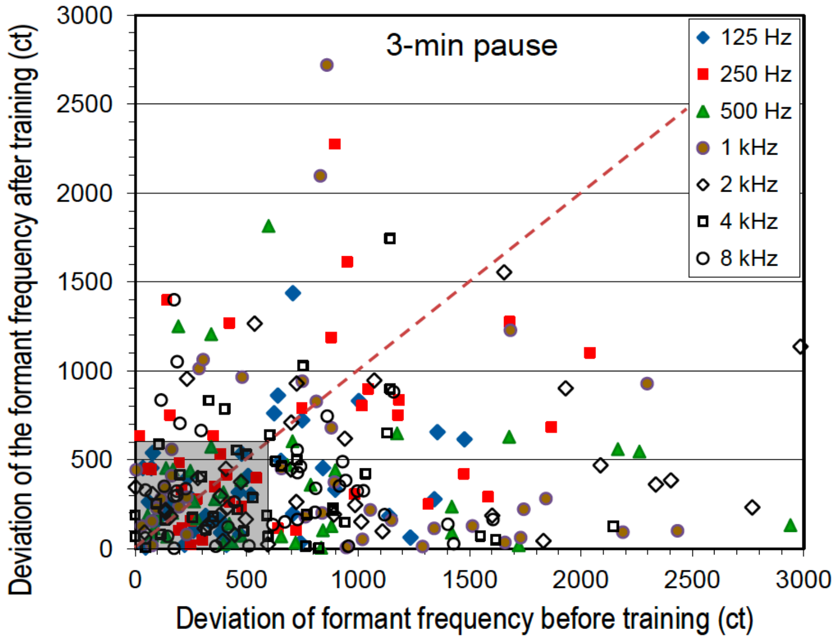

Figure 28. Pre-training deviation from the centre frequency in cents, for the second condition (three minute pause), before (abscissa) and after (ordinate) training. (Rościszewska & Miśkiewicz, 2014, ‘3.3. Results and discussion’)

The authors then present average performance data 'for each student across all formant frequencies' as shown in Figures 29 and 30 (Rościszewska & Miśkiewicz, 2014, '3.3. Results and discussion'). The authors note that 74% of students in condition one 'adjusted the formant frequencies more accurately after training', compared to 90% of students in condition two ('3.3. Results and discussion').

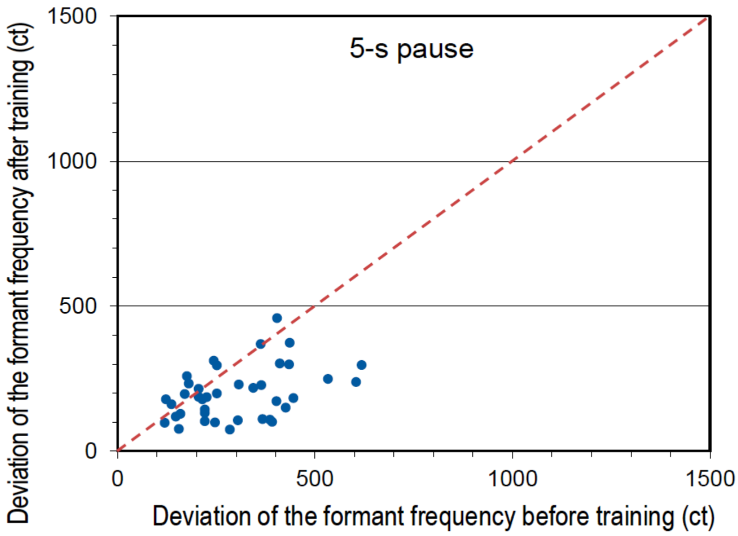

Figure 29. Average performance data for all formant frequencies for the first condition, before (abscissa) and after (ordinate) training. (Rościszewska & Miśkiewicz, 2014, '3.3. Results and discussion')

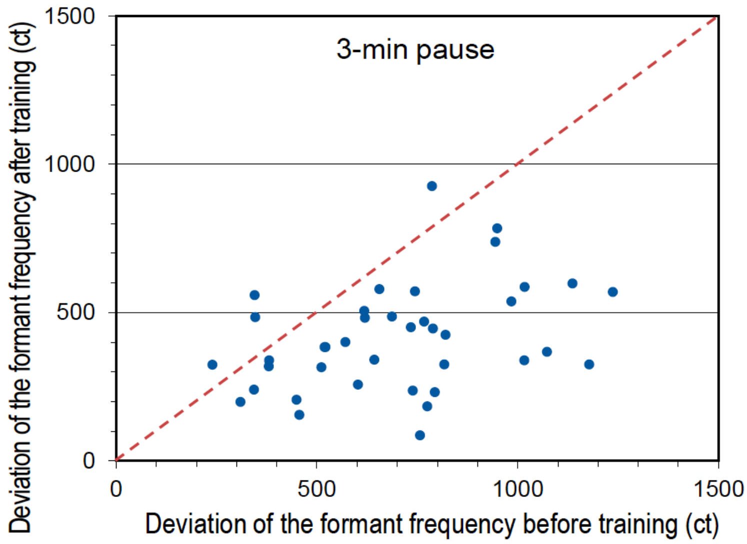

Figure 30. Average performance data for all formant frequencies for the second condition, before (abscissa) and after (ordinate) training. (Rościszewska & Miśkiewicz, 2014, '3.3. Results and discussion')

The mean and standard deviation for all centre frequencies is then presented for both conditions, as shown in Figures 31 and 32 respectively. The authors note that for both conditions the mean deviation from the centre frequency was lower in the post-training test, with a greater improvement in post-test performance evident in both conditions. The authors highlight that the deviation and the improvement were greatest in the second condition. (The authors also discuss the apparent frequency-dependent deviation particularly in the second condition, but this is beyond the scope of this review.)

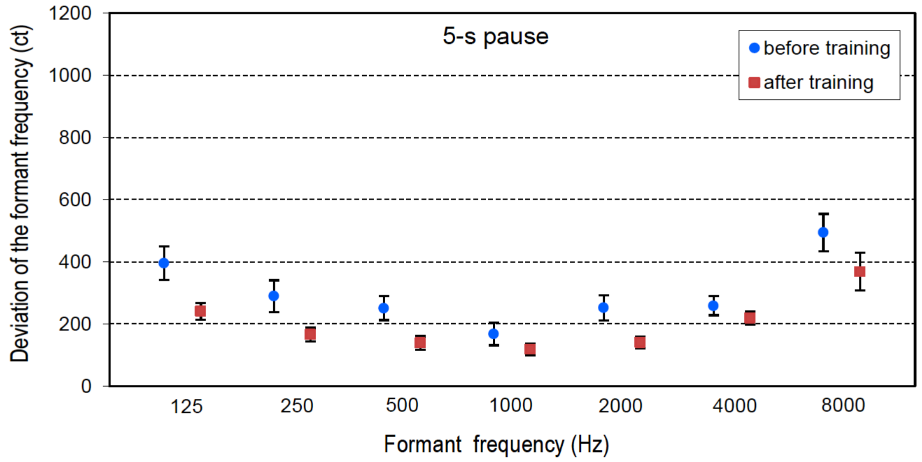

Figure 31. Deviation from the centre frequency in cents for all students for the first condition, before and after training. (Rościszewska & Miśkiewicz, 2014, '3.3. Results and discussion')

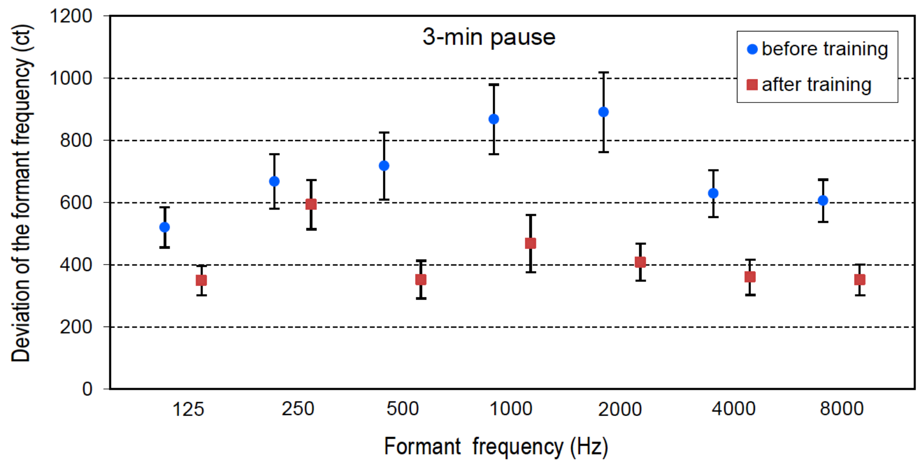

Figure 32. Deviation from the centre frequency in cents for all students for the second condition, before and after training. (Rościszewska & Miśkiewicz, 2014, '3.3. Results and discussion')

As previously discussed, the Timbre Solfège TETP requires students to perform matching tasks using a parametric equaliser applied to white noise, initially at one of nine octave band ISO-standard frequencies. Rościszewska and Miśkiewicz's test required students to perform a matching task using a parametric equaliser applied to white noise under two conditions. Condition one featured a five-second pause between the 'standard' and the 'variable', while condition two featured a three-minute pause.

In their most recent paper, Rościszewska and Miśkiewicz (2015) detail five separate tests used to assess the effectiveness of the Timbre Solfège TETP. These were a progress test given throughout the training program, an absolute identification task and a matching task (both administered at the end of the training program), and the previously detailed 2AFC timbre discrimination task and matching task.

The progress test consisted of 15–20 samples of music containing spectral irregularities introduced using one of several different equaliser types. Initially, students identified a single equaliser, but as the course progressed, the test increased in difficulty to 'include various combinations of two spectral modifications, including low- and high-frequency bandwidth cut-offs' and parametric boosts (Rościszewska & Miśkiewicz, 2015, p. 4). The parametric boosts were at one of nine octave band ISO-standard frequencies as previously detailed. The authors note that the 'tests [were] given to the students frequently throughout the course and [gave] the instructor useful feedback about the progress of training' (ibid.). Performance data was not provided for this test. The second test, administered at the end of the Timbre Solfège training program, required students to perform an absolute identification task. Students were presented with a musical excerpt 'in which segments of unprocessed sound (flat frequency response) [were] interleaved with segments spectrally shaped through the use of an equalizer' (ibid.). Each processed segment of music was modified using a combination of two separate equalisers, selected from low-pass filters, high-pass filters, and parametric boosts, set at octave and 1/3rd octave ISO frequencies. Students identified spectral irregularities introduced to 20 symphonic recordings, and 20 segments in rock recordings. The only presented information regarding students' performance on this test is the authors' comment that 'the scores of the best students amount to about 85%' (ibid.). The third test, also administered at the end of the Timbre Solfège TETP, required students to perform a matching task using a graphic equaliser. Students identified three simultaneously boosted frequency bands at 1/3rd octave ISO standard frequencies. Performance data was not supplied for this test. The authors detail the fourth and fifth tests, but these tests and the resultant performance data were presented previously (Rościszewska, 2011; Rościszewska & Miśkiewicz, 2014).

The Timbre Solfège TETP commences with the presentation of spectral irregularities introduced using a parametric equaliser, initially using 'white and pink noise, then music and speech' as stimuli (Rościszewska, 2011). At the beginning of the program, the spectral irregularities are introduced at one of nine octave band ISO standard frequencies and 'as the training proceeds, the nine basic formant categories are expanded to 27 one-third-octave bands extending from 40 through 16000 Hz' (Letowski & Miśkiewicz, 1995, p. 919). Students are required to perform passive and active tasks involving these equalised signals. Passive tasks require the student to 'detect, identify, and verbalize timbral differences between two sounds' (ibid.). Initially this task is performed using a single parametric boost, but 'the tasks become more complex and include two, three, or more modifications introduced at the same time (ibid.). The active task involved students completing a matching task, as previously described.

The test Rościszewska (2011) administered before and immediately following the Timbre Solfège TETP required students to detect, in a 2AFC test, spectral irregularities introduced by a parametric equaliser at '125 Hz, 500 Hz and 2 kHz and 8 kHz' (p. 2). Although students did not perform 2AFC timbre discrimination tasks in the Timbre Solfège TETP, the training program did require them to: 'detect … changes of timbre related to variations of the spectrum envelope' (Miśkiewicz, 1992, p. 622), 'detect… timbral differences between two sounds' (Letowski & Miśkiewicz, 1995, p. 919), and 'to detect differences between the timbres of two similar sounds' (Letowski, 1985, p. 241). Specifically, students in the Timbre Solfège TETP trained to detect spectral irregularities introduced by a parametric equaliser, initially at one of nine octave band ISO-standard frequencies: '63 Hz, 125 Hz, 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, 8 kHz and 16 kHz' (Rościszewska, 2011, p. 1). The gain values of the parametric boosts were not indicated. Although performance relating to the goals of the program was measured separately from the TETP, the test was identical to a task in the TETP. As such, this does not meet the current study's third criterion for effectiveness assessment, as performance relating to the goals of the program was measured on a task with which students were familiar.

The Timbre Solfège TETP trains students to perform a matching task (as detailed by Letowski, 1985; Letowski & Miśkiewicz, 1995; Miśkiewicz, 1992). Students are presented with a reference signal that has its spectrum modified in some way using one or more equalisers. Students then switch to a flat version of the signal and modify this signal to 'make both output signals subjectively identical' (Letowski, 1985, p. 241).

Several different stimuli are used in the Timbre Solfège TETP, including white and pink noise (Letowski, 1985; Letowski & Miśkiewicz, 1995; Miśkiewicz, 1992; Rościszewska, 2011). The centre frequency of the parametric boosts introduced are chosen from nine ISO-standard octave band frequencies (63 Hz, 125 Hz, 250 Hz, 500 Hz, 1k Hz, 2k Hz, 4k Hz, 8k Hz and 16k Hz) (Rościszewska, 2011). The gain values of the parametric boosts used in the training program are not indicated.

The test Rościszewska and Miśkiewicz (2014) administered prior to, and immediately after the training program required students to perform a matching task in which spectral irregularities were introduced using a parametric equaliser applied to noise (type not specified). The centre frequencies of the parametric boosts were chosen from seven ISO-standard octave band frequencies (125 Hz, 250 Hz, 500 Hz, 1k Hz, 2k Hz, 4k Hz, and 8k Hz). As previously detailed, the test featured two conditions: a five-second pause and a three-minute pause between the presentation of the modified signal and the flat signal. Although performance relating to the goals of the program was again measured separately from the TETP, the test was similar to a task that students trained on in the TETP; the first condition (five-second pause) was identical to the Timbre Solfège TETP task. The authors argued that the second condition (three-minute pause) tested 'the acuity of long-term memory' – something that students do not specifically train on in the TETP (2014, 3.3. Results and discussion, para. 4). However, the stimulus, modifications and overall task is identical to that which students perform in the Timbre Solfège TETP. As performance relating to the goals of the TETP was measured using a task with which students were familiar, the test does not meet the third criterion.

Rościszewska and Miśkiewicz's (2015) first, second and third tests required students to identify the parameters of one or more equalisers (including parametric), perform an absolute identification task, and perform a matching task – all identical to training tasks used in the Timbre Solfège TETP. The fourth and fifth tests (Rościszewska, 2011; Rościszewska & Miśkiewicz, 2014) were reviewed (see above) and were also found to be identical or almost identical to the tasks students trained on in the TETP. As such, the third criterion is not met.

In response to the increase in P[%] for Group A after undertaking the Timbre Solfège TETP, Rościszewska notes that students have the ability to develop tone colour discrimination skills 'that are not yet developed at the beginning of their studies' (2011, p. 6). Based on this, Rościszewska concludes that a tone colour discrimination test was not a reliable test of the candidate's suitability when used as an entrance exam. This contradicts Rakowski and Trybuła (1975), who suggest that 'in the selection of candidates for future Tonmeisters it is particularly important to check their abilities to differentiate the timbre of sound' (p. 1).

Rościszewska (2011) claims that 'sensitivity to changes in timbre may be considerably improved by systematic training', however, she cautions that although the majority of students demonstrate an increase in performance (P[%]) between the pre- and post-training tests, these results are not due to training with the Timbre Solfège TETP alone (p. 7). Indeed, Rościszewska concludes the paper with the observation that timbre discrimination skills are not solely developed during the Timbre Solfège TETP, but also during 'other exercises conducted… in the studios' (p. 6).

3.2.3.2 Timbral Ear Trainer I & II

Quesnel's (1990) preliminary study collected performance data for nine students on the Timbral Ear Trainer I TETP on two tasks: a 'comparative listening exercise' (matching/removing task) and an absolute identification task. Quesnel used the difference in students' performance on the TETP task at the beginning of the training sessions and six months later as evidence of the effectiveness of the training program. His comparative listening task required students to adjust the frequency of a parametric equaliser (gain and Q fixed at +12 dB and 2.0 respectively) applied to unprocessed white noise, so that the perceived timbre matched that of a processed reference signal. The parametric boost was chosen from one of 9 ISO-standard octave bands from 63 Hz to 16k Hz. In the absolute identification task, students identified the centre frequency of the parametric boost applied to the signal. The parameters of the equaliser were identical to those used in the comparative listening task. Twenty questions were presented for both the comparative listening and absolute identification exercises, with two questions repeated within each exercise.

Quesnel (1990) makes three observations about the comparative listening exercise results (Figure 33): first, the performance (percentage of correct answers) 'improved for all students'; second, for some students 'the difference [was] quite significant'; and third, individual variance between subjects was considerable at the beginning of training, whereas at the end 'it was much more homogenous', with all 'but one [student] obtaining perfect scores' (p. 97).

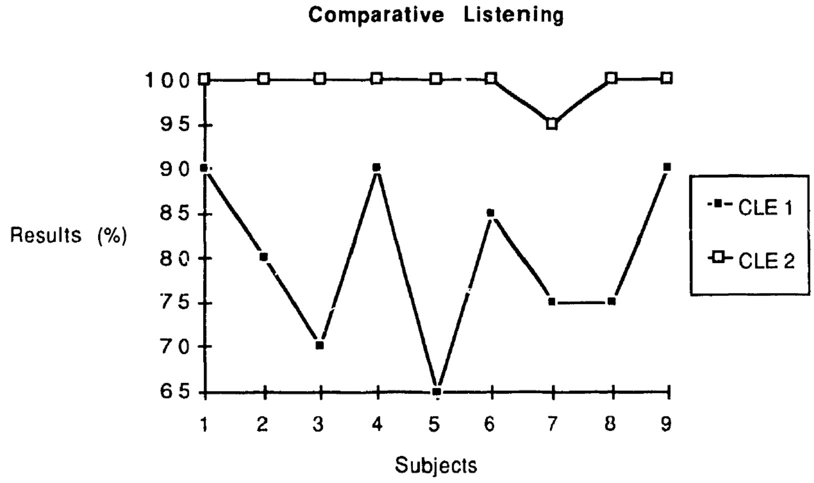

Figure 33. Percentage of correct responses for the nine participants for the comparative listening exercise at the beginning (CLE 1) and the end (CLE 2) of the Timbral Ear Trainer I TETP. (Quesnel, 1990, p. 98)

Quesnel (1990) makes three observations regarding the response time for the comparative listening exercises (Figure 34). First, he notes that 'all students responded more quickly at the end' of the training period (p. 100). Secondly, he notes that the difference between the average response times (Time 1 compared to Time 2) was significantly different, with the response time for most students 'decreas[ing] by a factor greater than 2' (100). Finally, he observes that 'the homogenization effect of practice that was observed in the score data was not apparent in the time data' (p. 100).

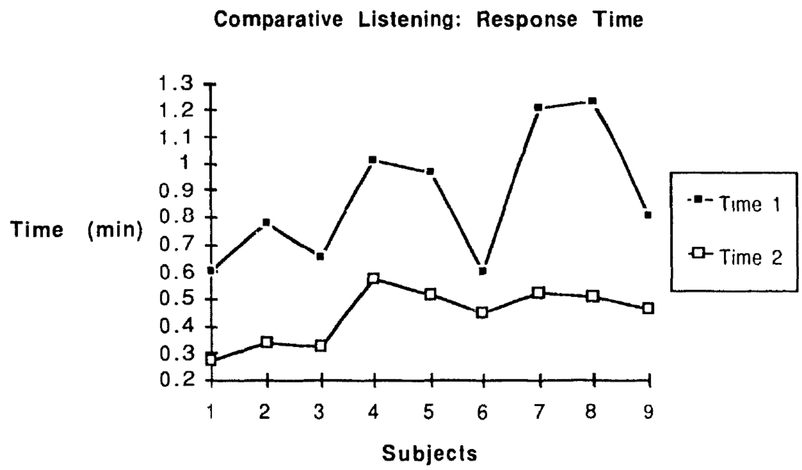

Figure 34. Average response times for the nine participants for the comparative listening exercise at the beginning (Time 1) and the end (Time 2) of the Timbral Ear Trainer I TETP (Quesnel, 1990, p. 99)

In reference to the performance data for the absolute identification task (Figure 35), Quesnel (1990) notes that the absolute identification exercise was more difficult than the comparative listening exercise, because, in contrast to the comparative listening exercise, the 'students' abilities … remained unequal after the training period' (p. 101).

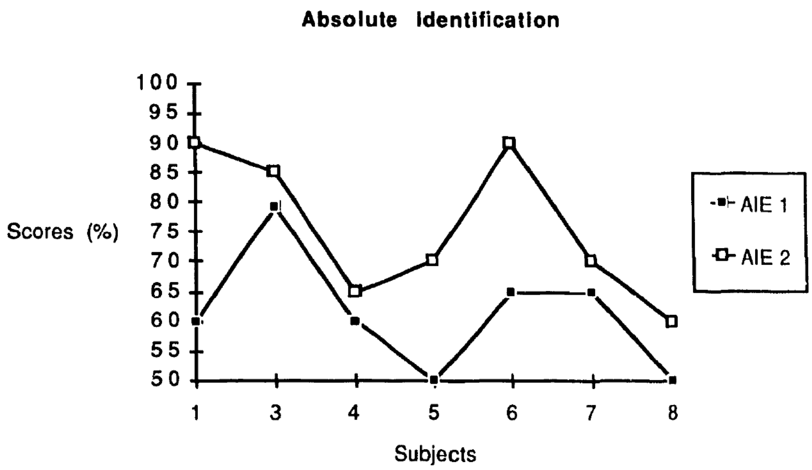

Figure 35. Percentage of correct responses for the nine participants for the absolute identification listening exercise at the beginning (AIE 1) and the end (AIE 2) of the Timbral Ear Trainer I TETP. Only seven students of the original nine participated in the absolute identification exercise at the beginning and end of the training period. The number '2' is missing from the horizontal axis. (Quesnel, 1990, p. 103).

Quesnel (1990) highlights that students who performed better at the beginning of the training program also performed better at the end. He suggests that a lack of practice time was a possible explanation for the absence of 'asymptotic performance' (p. 101). Despite this, Quesnel observes that all students' performances improved, although the magnitude of the improvement was small for some students. He offered several explanations for poor improvement, including: too much practice time (listening fatigue), not enough practice time, and 'errors in the use of the program' (p. 102). Quesnel also suggests that 'some students might have lower perceptual ability limits than others', suggesting that more research was needed to 'determine the [practice] time necessary to obtain asymptotic performance' (p. 102). Quesnel concludes by cautioning that ‘the data presented here do not allow any final conclusions to be drawn about the validity of the system in its current state as a training/teaching tool for timbre perception’, although he noted ‘that such a system could be a useful tool for sound engineers to improve the perceptual abilities they need in their profession’ (p. 102). Quesnel did not elaborate on the type of perceptual abilities involved.

The performance data for the second deployment of the Timbral Ear Trainer I program (Quesnel & Woszczyk, 1994) again resulted from students undertaking comparative listening and absolute identification exercises. The comparative listening exercises were slightly different to Quesnel’s 1990 study in that two types of exercises were implemented. The first, a matching task, was identical to that delivered by Quesnel (1990), as detailed above. The second comparative listening exercise was a ‘bring to flat’ task, as explained in Chapter 1. The absolute identification task was also identical to Quesnel’s (1990), as detailed above. Seven students participated in the second implementation of the Timbral Ear Trainer program.

For the matching task (Figure 36), most (5/7) students achieved a perfect score at the beginning of the training program. The two students who did not showed a slight improvement after the training program. The authors note that such ‘results [were] not surprising considering the low difficulty level of the exercise’ (Quesnel & Woszczyk, 1994, p. 5). In reference to the bring-to-flat task (Figure 37), the authors note that ‘individual ability within the group varied substantially at the beginning of the training whereas at the end, all students obtained a perfect score’ (ibid.). Finally, in reference to the absolute identification task (Figure 38), the authors make two relevant observations: that ‘significant improvement occurred for all students’ and that ‘performance levels [were] again more homogeneous at the end of the training than at the beginning’ (ibid.).

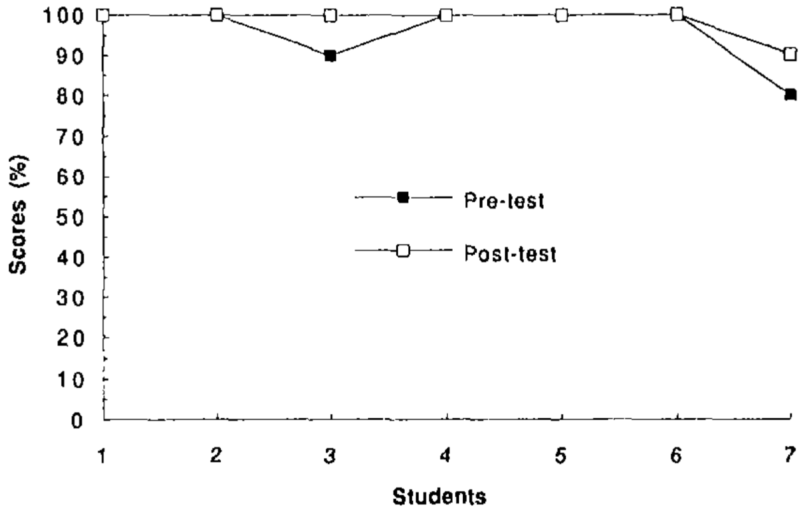

Figure 36. Matching performance scores for the second iteration of the Timbral Ear Trainer I program (Quesnel & Woszczyk, 1994)

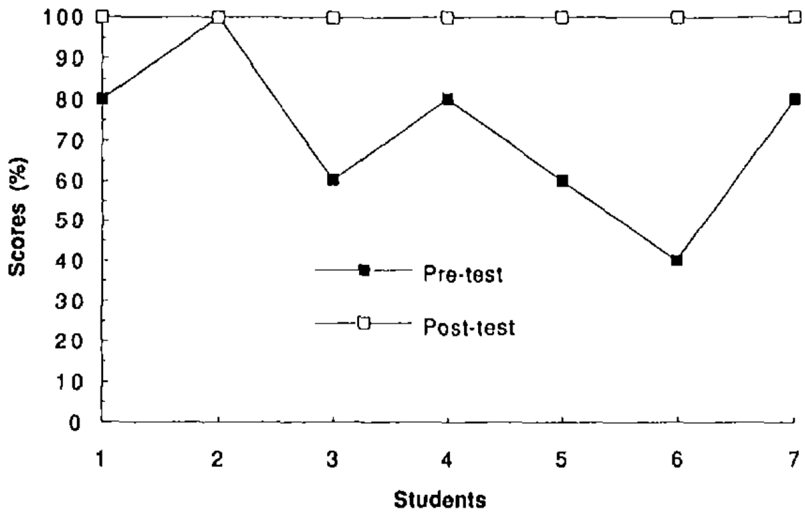

Figure 37. Bring to flat performance scores for the second iteration of the Timbral Ear Trainer I program (Quesnel & Woszczyk, 1994)

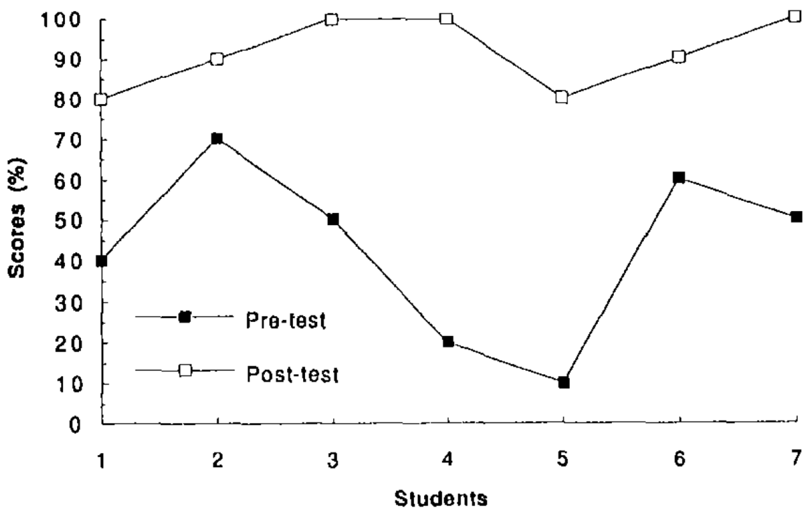

Figure 38. Absolute identification performance scores for the second iteration of the Timbral Ear Trainer I program (Quesnel & Woszczyk, 1994)

Next, Quesnel and Woszczyk (1994) investigated the correlation between performance and response times (Figures 39, 40 and 41); they conclude that 'no correlation seems to exist between practice time and final scores' (p. 6). Correlation between practice times and response times was also investigated, and the authors conclude that 'there is no correlation between practice time and response time for the' matching exercise (p. 6). However, for the absolute identification tasks, 'r values [were] significant at the 5 percent level using a two-tailed test' (p. 6). Finally, for the bring-to-flat task, the authors note that 'the [correlation] coefficient value is suggestive of a relation (r =0.732 and critical r value at .05 significance = 0.754)' (p. 6). The authors then investigated the relationship 'between total practice time and total number of problems answered', 'reveal[ing] a strong correlation of 0.851 at the 5 percent significance level' (p. 6). However, 'no significant correlation was found between total practice time and average response time per problem during the 8-month period' (p. 6). The final analysis concerned the relationship between the total practice time and average scores, and total practice time and response times. A correlation was found 'between total practice time and improvement in average score (0.789 p = 0.05)', but not found 'between practice time and improvement in mean response time' (p. 6).

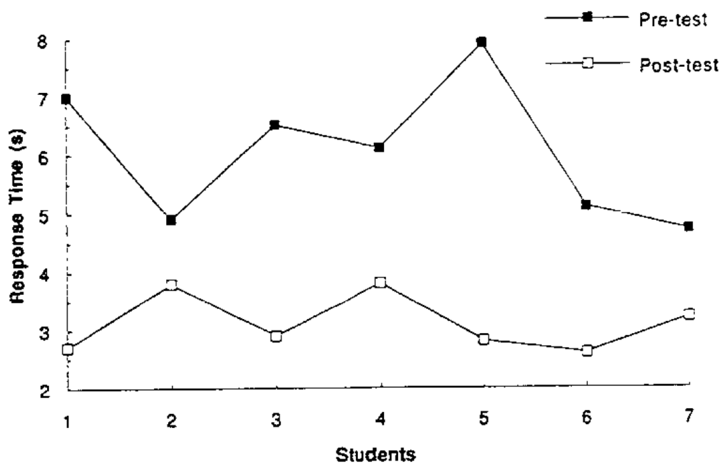

Figure 39. Response times for the matching task (Quesnel & Woszczyk, 1994)

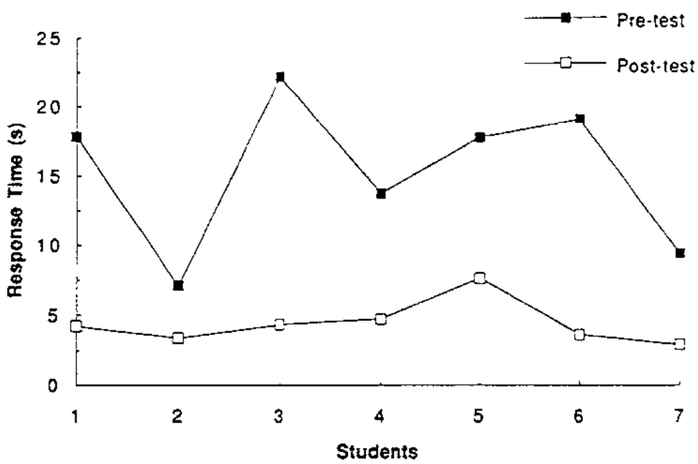

Figure 40. Response times for the bring to flat task (Quesnel & Woszczyk, 1994)

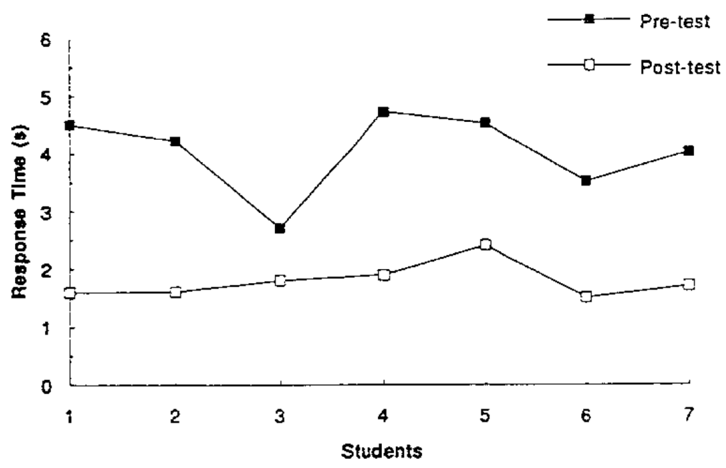

Figure 41. Response times for the Absolute identification task (Quesnel & Woszczyk, 1994)

In his 2001 paper, Quesnel presents performance results on a timbre discrimination task for five students who had undertaken the Timbral Ear Trainer II program, compared to five professional audio engineers who had not used the training program. The five students were postgraduate Masters students enrolled in McGill's TET course. On average, each student had 2.5 years of audio engineering training and 4.5 years of experience working in the field of audio. On average, the professional audio engineers possessed '19 years of experience working in the audio engineering field and 2.2 years of training in audio' (Quesnel, 2001, p. 71). Although Quesnel notes that the five professional audio engineers had not undertaken 'formal aural training', he contended that 'that the number of years of academic audio training was similar' in the student group and the professional group (ibid.). Quesnel (2001) notes that the professional engineers possessed 'the advantage of experience' and as such, the test was able to measure 'the effect of a short period of structured, theory-based training… against the effect of many years of hands-on experience' (p. 72).

The test required participants to perform a bring to flat task, which Quesnel (2001) categorises as a comparative listening task. Although this type of task involves the removal of modifications applied to the spectrum, it is in fact a type of matching task (p. 75). In each of the 15 trials in the test, participants were presented with three sounds – A, B and C – which they could switch between at any time. Sound A, which was unable to be modified by the user, had its spectrum modified using between one and three equalisers set to apply either a +6 dB boost or -6 dB cut, at one of one of eight octave band ISO-standard frequencies between 125 Hz and 16k Hz inclusive. Q was set to 2.0 and was made known to all participants. Participants were required to modify Sound B (initially identical to Sound A) so that it was perceptually identical to Sound C. Sound C was an unprocessed (flat) version of the signal. The stimulus was Steely Dan's 2000 track, 'Gaslighting Abbie', with which 'none of the subjects were familiar' (pp. 72-73). The test was conducted 'in the small listening room used for the Technical Ear Training course, using two Genelec 1030 loudspeakers arranged in a standard equilateral triangle stereo listening configuration' (p. 73). Participants were informed that they would be assessed based 'on the accuracy of their answers and their response time' and were told to answer as accurately and quickly as possible (p. 76). Participants had the option of undertaking up to six trial questions 'to familiarize themselves with the operation of the software, the test procedure, and the stimulus' (p. 76).

Participants' answers were assigned a mark from 0 to 1, with a different weighting given to various components. The mark was comprised of 65% frequency identification and 35% gain identification. The weighting was distributed evenly amongst the number of bands used, as shown in Table 2.

Table 2. Evaluation weightings – the weighting for frequency and gain is divided by the number of bands. (Quesnel, 2001, p. 78)

	1 Band	2 Bands	3 Bands
Frequency Weight (%)	0.65	0.325	0.217
Gain Weight (100%)	0.35	0.175	0.117

For responses to questions in which listeners 'modified more bands than there were in the question, a fixed value of 0.2 was subtracted from the mark for each extra band' (Quesnel, 2001, p. 78). Participants' response time was recorded but not used to calculate the mark for each question. It was, however, used to calculate a 'performance index'. Quesnel compared the performance of the two groups using several metrics.

The mean score and standard deviation for the control group (professional audio engineers) versus the student group are shown below in Figure 42. The five students outperformed the five professional engineers with a mean of 96% correct compared to 78.8% correct respectively. In addition, all students scored above 90%. The standard deviation for the student group was also considerably smaller, at 2.98, than the 16.8 for the professional audio engineers. Quesnel's (2001) analysis showed that 'the score difference between the two groups was statistically significant' based on a one-way analysis of variance ($F(1,9) = 5.08$, $p \leq 0.05$) (p. 81). Quesnel also examined the percentage of 100% correct trials for all 15 trials for each participant in each group. He found that students had significantly more ($F(1,9) = 6.58$, $p \leq 0.03$) perfect scores than the professionals on a trial-by-trial basis (p. 84).

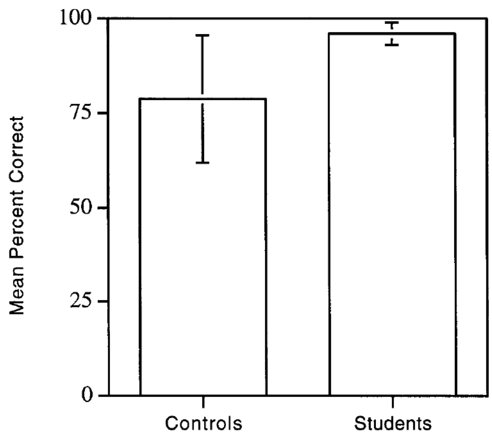

Figure 42. The mean and standard deviation scores for the professional audio engineers (Controls) and the students. (Quesnel, 2001, p. 81)

Quesnel (2001) had 'initially hypothesized that response time means would be lower for the student group than for the control group', but found no statistically significant difference in response time between the two groups, as shown in Figure 43 (p. 86). Quesnel did calculate the total time taken to complete the test for both groups, with students taking on average 40.4 minutes and the professionals 71.6 minutes; however, the total time spent listening was similar for both groups. This is discussed further in the Review section below.

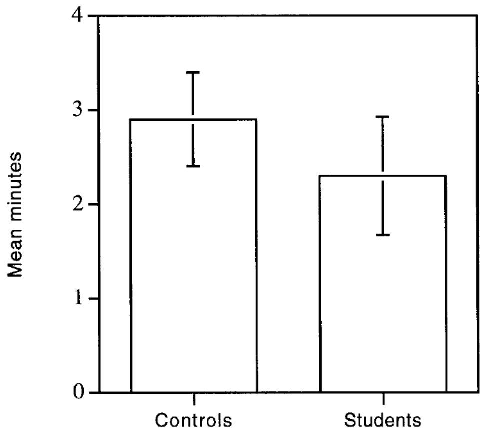

Figure 43. The mean and standard deviation response times for the professional audio engineers (Controls) and the students. (Quesnel, 2001, p. 86)

Quesnel had been an advocate of the use of vowels in TETPs for some time, first proposing the idea in his 1990 paper. Quesnel and Woszczyk (1994) state that a 'variant of [the absolute identification exercise] requires the students to answer using vowel categories instead of using center frequency values' (p. 4) in the first version of the Timbral Ear Trainer program. These absolute identification exercises require students to 'identify a particular timbre with a label (e.g. the center frequency value of a resonance or the vowel that can be associated with the resonance)' (Quesnel, 1994, p. 39). In Quesnel's 1996 paper, he details a dedicated vowel identification task in which students were trained 'to associate resonances with vowel-like sounds' and were required to answer 'by selecting [the] vowel that best corresponds to the resonance presented by the computer' (1996, '2.3.6 Vowel Identification Tasks'). The Timbre Solfège II test interface features vowel anchor point buttons for the English vowels 'a', 'e', 'i', 'o' and 'u'. During the test, instead of adjusting the centre frequency of an equaliser manually, participants could click on any of these five vowel buttons, which moved the centre frequency of the equaliser to the frequency that was predetermined to correspond with the chosen vowel sound. These vowel anchor points were only used for frequencies between 250 Hz and 4k Hz inclusive. Quesnel (2001) investigated the use of these vowel buttons in his analysis and found that 'the mean values obtained for the two groups [professional audio engineers and students] were similar and the difference between the two groups was not significant' (p. 91).

The criterion that was ultimately used to compare the performance of the professional audio engineers and students was a performance index (PI). This PI was calculated using both the response time an accuracy of the responses. Quesnel defined the PI 'as a weighted ratio of matching/identification accuracy over response time' as shown in Equation 2.

$$PI = \frac{S}{k(T_R)}$$

where PI is the performance index,
S is a measure of accuracy of the adjustment,
T_R is the response time in seconds, and
k is the weight applied to response time in the calculation of PI.

Equation 2. The performance index developed by Quesnel (2001, p. 37).

Quesnel compared the PI for the two groups, finding a statistically significant difference between the means for the student group (87.7) and the professional audio engineer group (55.8). Quesnel (2001) states that 'out of the 11 performance indicators under study, 4 were significant at the $p \leq 0.05$ level (score, number of steps, frequency adjustments and performance index)' (p. 96). He summarised as follows:

> The control group subjects had substantially more professional listening experience than the student subjects and this difference was highly significant ($p \leq 0.001$). Yet, the students outperformed them. In addition, the subjects in the two groups had received a similar number of years of audio training. This strongly suggests that the technical ear training course the student subjects took as part of their audio training contributed substantially to their higher performance level. (Quesnel, 2001, pp. 96-97)

The performance on the TETP itself is used as evidence of the effectiveness of the Timbral Ear Training program in Quesnel's initial study (1990). This clearly does not meet the third criterion proposed in the current study, as performance relating to the goals of the TETP was not measured separately from the TETP itself.

In their 1994 paper, Quesnel and Woszczyk claim that a 'pre-test/post-test procedure' was used to establish the effectiveness of the training program. This implies that a test separate from and different to the TETP was used at the beginning and end of the Timbral Ear Trainer program. However, the test used to establish the effectiveness of the TETP was not separate from the training program; the authors simply compared students' performance on the TETP tasks at the beginning of the program to their performance at the end (Quesnel & Woszczyk, 1994). This test does not meet the third criterion for the same reason as Quesnel's 1990 study.

The primary result from Quesnel's 2001 doctoral thesis, that is, students who trained using a TETP outperformed professional audio engineers on an equalisation task, is often cited by TET researchers as evidence of the effectiveness of TETPs (Corey, 2004, 2010, 2017; Rościszewska & Miśkiewicz, 2015). However, upon reviewing Quesnel's test, it becomes apparent that summarising the results as 'student subjects performed significantly better' is inadequate (Quesnel, 2001, p. 96; Rościszewska & Miśkiewicz, 2015, p. 6). Indeed, there are several reasons unrelated to the development of a 'memory for a set of octave resonance references', and aside from the small sample size, which may explain why the students outperformed professional audio engineers on the test (Quesnel, 2001, p. 45). The primary reason for this is that the students were familiar with many aspects of the test, largely because they had been trained to perform an identical task using identical software in an identical room during the Timbral Ear Trainer TETP. This provided students with a significant advantage over the professional audio engineers that was unrelated to any skill the students may have developed in identifying the properties of spectral irregularities introduced by a parametric equaliser.

The test that Quesnel administered contained a task identical to that which the student group was trained on. Indeed, when detailing the task participants undertook in the test, Quesnel (2001) refers the reader to Section 4.2 of his thesis (p. 73). This section details a task that 'illustrate[s] the training method' used in the Timbral Ear Trainer program (p. 50); Quesnel made no distinction between the task used as the test and the task used in the TETP, because they were identical. Quesnel also listed the task as one of several delivered within the Timbral Ear Trainer program in earlier papers (1996, 1994). Unlike the student group, the professional audio engineer group had no prior experience with the task. As students were intimately familiar with the task administered in the test, this does not meet the third criterion of TETP goals being measured on an independent task.

In addition to the task, the student group was familiar with the listening environment and playback system, as the test was conducted 'in the small listening room used for the Technical Ear Training course' (Quesnel, 2001, p. 73). Students were familiar with the overall space, the acoustic space and its idiosyncrasies, in addition to the response of the loudspeakers in the room. The professional audio engineers had no prior listening experience in the playback environment. It could be argued that professional audio engineers should possess the ability to walk into an unfamiliar acoustic space and perform their duties to a professional standard; if such a space has an appropriate frequency response and is relatively free from acoustic deficiencies such as standing waves, flutter echoes and speaker-boundary interference response, then this argument increases in validity. Although the specifics of the listening environment were not published, one group having familiarity with the space biases the comparison, and undoubtedly provided an advantage to the student group.

Quesnel's (2001) participants were given up to six sample questions 'to familiarize themselves with the operation of the software, the test procedure, and the stimulus' (p. 76). However, the students had already spent several months becoming familiar with the software interface, procedure, listening space and playback system. Six questions are clearly insufficient to bring the professional engineer's level of familiarity to parity with the students'. Quesnel partially acknowledges this, noting that the students 'were already familiar with the software and the task', but then dismissed this as a possible contributing factor, because the students were not 'familiar with the test procedure nor with the stimulus' (p. 76).

Quesnel (2001) notes that response times between the groups were not significantly different, but total test completion times on average differed by over half an hour in favour of the student group. As the total listening time was similar for both groups, the difference in completion times 'corresponds to time periods in between trials during which subjects could rest' (p. 87). Indeed, Quesnel notes that 'a common observation expressed by subjects in the control group was that the task was much more tiring than they expected' (ibid.). Students were already familiar with the test environment, but its specific elements were unfamiliar to the professional audio engineers. The professionals had to learn to use an unseen software interface, understand the testing procedure and familiarise themselves with the room by answering a maximum of six sample questions. The professionals' unfamiliarity with the software interface alone may have contributed to the tiredness that they experienced due to performing the task. In reference to this, Quesnel hypothesises that 'they [the professionals] needed more rest periods' because they did not train using the TETP (2001, p. 87). This is easily explained, not by the fact that the student group developed superior endurance on equalisation-related tasks as a result of training with the Timbral Ear Trainer program, but by the fact that the professionals were unfamiliar with the software interface. Conversely, the student groups' familiarity with the interface may have minimised the fatigue experienced when engaging with the interface.

Regardless of whether students were familiar with the interface, it is logical to assume that professional audio engineers with considerable experience should outperform novices on a simple test of audio engineering ability. Quesnel argues that the 'test task simulated a task common in sound recording practice'; based on this, he suggests that professional audio engineers should have been able to outperform students on the test (2001, p. 73). There was, however, one additional, crucial factor, aside from those already listed, that not only influenced the degree to which the test can be considered 'real-world', but which also casts doubt over the prediction that the professional engineers should outperform students: the use of vowels.

The test interface featured vowel buttons that Quesnel (2001) instructed participants 'to use… only if they were familiar with the relationship between vowels and resonance frequencies' (p. 91). As part of the Timbral Ear Training program, students were trained to use the vowel buttons and were familiar with the relationship between the buttons and the corresponding centre frequencies. Students were familiar with the interface and thus the presence of vowel buttons on the interface. Vowel buttons do not exist on professional audio equipment and are not used in commercial audio production. It is unsurprising, therefore, that every professional audio engineer stated that 'they were not familiar with the concept' of vowels as they pertain to equalisation (pp. 91–92). Although the professional audio engineers had received a 'number of years of audio training', the use of vowels to identify the centre frequency of a parametric equaliser is not taught outside a small number of schools, a fact demonstrated by the audio engineers' unawareness of the concept (pp. 96-97). The inclusion of these vowel labels not only makes the simulated task less like one 'common in sound recording practice' (2001, p. 73), but probably further biased the comparison in favour of the students.

Despite their unfamiliarity with the vowel buttons, the professionals did use them during the test. Quesnel notes that 'the mean values obtained for [the vowel button usage for] the two groups were similar and the difference between the two groups was not significant' (2001, p. 91). However, one group was not only familiar with this feature but was trained to use it on an identical task; the other group was entirely unfamiliar with the buttons, even conceptually. The data did show that the students used the vowel buttons approximately as often as the professional engineers, but the fact that the students were educated on the relationship between the vowel buttons and the centre frequencies and were previously trained to use the buttons on an identical task means that the very presence of these vowel buttons in the test diminishes the professional relevancy of the test itself and casts doubt over the results.

If the test was designed to mimic 'real conditions in which sound engineers shape the timbre of sound during the production of audio recordings', a primary feature of the test interface was not (Rościszewska & Miśkiewicz, 2014, 4. Conclusions, para. 1). The presence alone of the vowel buttons on a test that is meant to simulate a real-world audio engineering task weakens the methodology. The inclusion of interface features that are not used in professional audio practice or found on any piece of signal processing equipment, that one group was trained to use and the other was unfamiliar with, puts the results in a new context. These issues cast doubt over the legitimacy of claims that the students outperformed professional audio engineers on a real-world audio engineering task. Moreover, these issues, unrelated to the effectiveness of the TETP, also highlight why it was foreseeable that the students would outperform professional audio engineers on the test.

3.2.3.3 Beijing Union University Training

Liu et al. (2007a) reports the results of a study in which 57 students took part in a TETP delivered over 15 weeks at Beijing Union University. The students had never undertaken a TETP before and had minimal experience working in the field of audio. Based on this, the authors suggest that the results of their study could 'be applied in common cases' (p. 2). Students were divided into two groups of 31 and 26 respectively (why or how is not explained). Each student spent 30 minutes per week working on the TETP. One of the tasks that students trained on involved the discrimination of frequency response irregularities presented over loudspeakers. The frequency response irregularities were applied to pink noise using a parametric equaliser set to boost or cut 'about 3dB to 6dB' at one of seven octave band ISO-standard frequencies from 125 Hz to 8k Hz inclusive (p. 3). Students estimated the centre frequency of the equaliser and whether the frequency had been boosted or attenuated. Progress tests were administered throughout the TETP to 'let the subjects know their learning progress' (ibid.).

The test conducted at the end of the TETP required students to identify the centre frequency of an equaliser applied to pink noise at different frequencies using a 6 dB boost or cut. Liu et al. (2007a) do not state whether the same centre frequencies as the TETP were used, nor whether the same gain changes (boosts or cuts) were applied, instead, simply that '6dB [of] EQ processing' was used (p. 3). The test also contained other tasks including pure tone identification, musical scales and instrument identification. Test performance was assessed using 100 questions, with each correct response earning the student one mark. The authors compared the students' performance on this final test to that on an initial test.

The only information presented relating to students' performance on the initial test is the authors' statement that '[on] the initial test, the average mark [was] less than 50' (Liu et al., 2007a, p. 3). The mark refers to the total mark for all tasks, not only the frequency response discrimination task. With respect to the final test marked out of 100, the authors state that the first group averaged 84.5 and the second 86.2. The highest mark achieved within either group was 96, and the lowest 66. Based on this, the authors conclude that 'the two groups of subjects made great progress after [as a result of] training' (p. 3). Results for each task within the final test are presented in Table 3. As shown, Group 1's average score was 45.6% on the frequency response irregularity task, and Group 2's was 49.0% correct. Results for the frequency response irregularity task for the first test are not presented. Despite the lack of pre/post-training data, the authors conclude that 'the discrimination of the frequency response irregularity has been [*sic*] improved little' and is 'hard to improve, [and] it is difficult for most subjects' (p. 4). The authors hypothesise that this result occurred because 'the item is hard or the training methods need some improvement' (ibid.). As results from the first test are not presented, the magnitude of this claimed improvement is unknown.

Table 3. The percentage of correct responses for both groups on all tasks within the final test. (Liu et al., 2007a, p. 4)

Group	**Absolute Frequency**	**Relative level**	**Relative Frequency**
1	92.6%	91%	79.7%
2	96.4%	94.2%	55.4%
Average	94.3%	92.5%	68.6%
Group	**Irregularity**	**Instruments**	**Total Average**
1	45.6%	84.3%	84.5%
2	49%	88.1%	86.2%
Average	47.2%	86%	85.3%

Although Liu et al. (2007a) state that students' performance on the frequency response irregularity task 'improved little' (p. 4), the authors subsequently claim that 'after the ear training the majority of subjects can make great progress for all the training items' (p. 4). The authors also claimed 'nearly 85% average correctness rates for all the [test] items', which again is not supported by the results and contradicted within the paper.

Like several previous researchers, Liu et al. (2007a) attempted to use performance on the TETP as evidence of its effectiveness (although the results showed that students' performance on the frequency response irregularity training task did not improve markedly). As such, the third criterion proposed in the current research was not met, as performance relating to the goals of the TETP was not measured separately from the TETP itself.

3.2.3.4 Personalised Timbral Ear Trainer

In 2011, Kaniwa et al. published a paper detailing what the authors describe as an 'adaptive' and 'personalised' TETP. Aside from detailing their adaptive TETP, they presented the results of a study designed to establish the effectiveness of the adaptive TETP compared to a standard, non-adaptive TETP. The authors established the relative 'effectiveness' of the adaptive program by comparing the performance of one group of student' on the adaptive TETP to another groups' performance on a non-adaptive TETP. Specifically, the authors looked at the mean and standard deviation of the percentage of correct responses for the two groups and the improvement in students' performance on questions on which they initially scored poorly.

The eight students that participated in the study were divided into a Conventional group and a Proposal group. The Conventional group trained using a standard version of the authors' TETP that featured a non-weighted random question pool. The Proposal group trained with the same TETP, but their questions were chosen 'using [a] weighted random function [that] dynamically updated according to the trainee's previous training scores' (2011, '4. Experiment to evaluate proposed system').

Prior to undertaking the TETP, Kaniwa et al. (2011) conducted a preliminary study designed to ensure that the two student groups contained similarly skilled participants. In this preliminary study, the students performed an absolute identification task whereby they identified the centre frequency of a parametric filter applied to pink noise (+12 dB, Q not stated). Based on each student's performance on this preliminary study, the authors assigned them to one of the two groups. The authors then 'conducted a preliminary F-test to test the equal variance, followed by a two-tailed t-test (two independent samples with equal sample size and equal variance), which confirmed that there were no statistically significant differences between the two groups' (2011, '4. Experiment to evaluate proposed system').

Following the preliminary study, each group of students trained for 30 minutes twice a week with their respective TETP for four weeks (Kaniwa et al., 2011). The training regime differed each week, as shown below in Table 4. The students trained using three stimuli: pink noise, orchestral music and a piano sonata. The spectral aberrations were applied to one of seven octave band centre frequencies from 125 Hz to 8k Hz inclusive using a parametric equaliser set to one of the gain values shown in Table 4. Students performed an absolute identification of the centre frequency only. The TETP was delivered using headphones; the authors do not specify mono or stereo playback. In addition, students had the option of practising on the first day of training each week using the filter parameters and stimulus shown in Table 5.

Table 4. The four-week training regime undertaken by both groups. (Kaniwa et al., 2011, '4. Experiment to evaluate proposed system')

Week	dB	Sound File
1	+12 dB, +6 dB	Pink Noise, Orchestra, Piano
2	+6 dB, +3dB	Pink Noise, Orchestra, Piano
3	−12 dB	Pink Noise, Orchestra
4	±12 dB, −12 dB	Pink Noise, Orchestra

Table 5. The practice test filter parameters and stimulus. (Kaniwa et al., 2011, '4. Experiment to evaluate proposed system')

Week	dB	Sound File
1	+12 dB	Pink Noise
2	+6 dB	Pink Noise
3	−12 dB	Pink Noise
4	±12 dB	Pink Noise

Kaniwa et al. (2011) first examined the mean percentage of correct answers for each stimulus, for each group (Figure 44). As the pre-training average for the two groups was 'about the same' and the results indicated that the Proposal group's average correct answer rate was higher than that of the Conventional group, the authors conclude 'that the proposed [technical ear training] system raises the correct answer rate' in comparison to the standard training system (2011, '5.1 Analysis by Mean and Standard Deviation').

The results for each stimulus over the four-week training period were then presented (Figures 45–46), with results for only the first two weeks of training presented for Piano (Figure 47). From these results the authors concluded 'that the mean correct answer rate of the Proposal group [was] consistently higher than that of the Conventional group for all sounds' (Kaniwa et al., 2011, '5.1 Analysis by Mean and Standard Deviation'). The authors reported that an additional analysis of the data using a preliminary F-test and a one-tailed t-test showed a 'significant difference' between the performance of the two groups for the Piano and Orchestral music stimuli, but this difference was not apparent with Pink Noise (ibid.). Based on these results, the authors concluded 'that the proposed system is more efficient in terms of training for realistic and difficult tasks than the conventional training system' (ibid.).

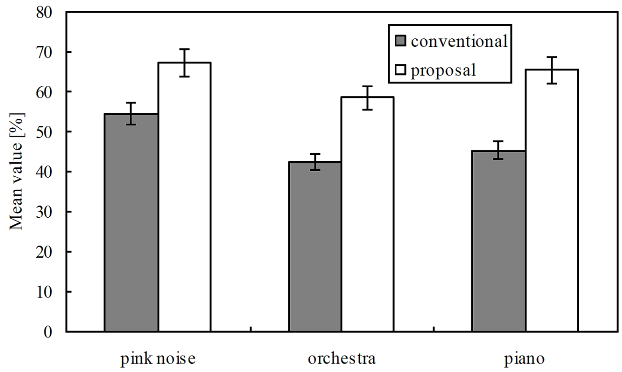

Figure 44. The mean and standard deviation of the percentage of correct answers for each stimulus for the two training groups. (Kaniwa et al., 2011, '5.1 Analysis by Mean and Standard Deviation').

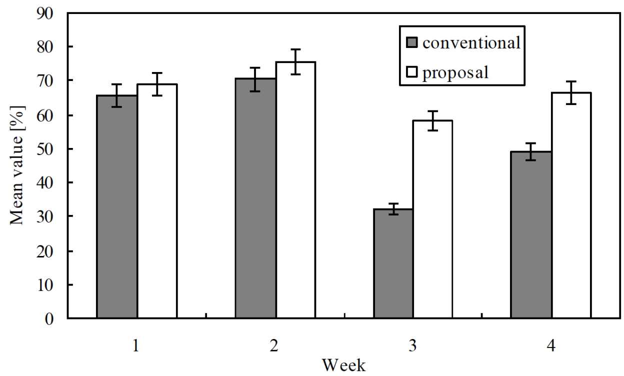

Figure 45. The mean and standard deviation of the percentage of correct answers for pink noise for each week of the training regime. (Kaniwa et al., 2011, '5.1 Analysis by Mean and Standard Deviation')

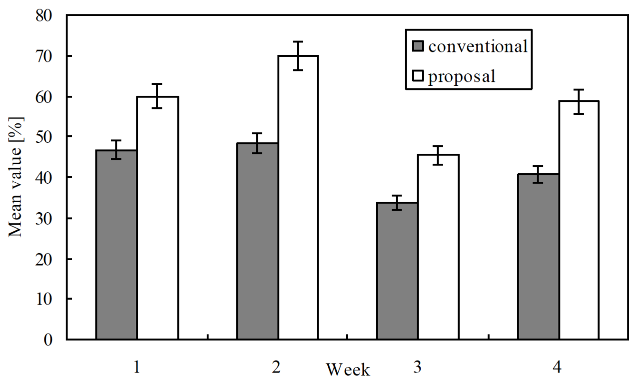

Figure 46. The mean and standard deviation of the percentage of correct answers for orchestral music for each week of the training regime. (Kaniwa et al., 2011, '5.1 Analysis by Mean and Standard Deviation')

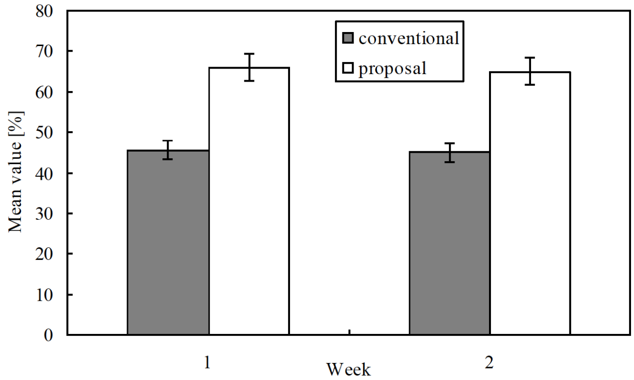

Figure 47. The mean and standard deviation of the percentage of correct answers for piano for the first two weeks of the training regime. (Kaniwa et al., 2011, '5.1 Analysis by Mean and Standard Deviation')

Students' performance in the 'low-score bands', that is, students' performance over time on questions which they initially answered incorrectly, was also analysed. For each student in the Proposal group, Kaniwa et al. (2011) identify the 'three low-score [frequency] bands that marked the lowest correct answer rates' for that week and note the percentage correct answer rate for each band (2011, '5.2 Analysis of Improvement for Low-Score Bands'). The authors then looked at how the students performed on those identified bands the following week. Relative to the previous weeks' percentage correct answer rate, students could either increase (assigned a '+'), decrease (assigned a '-'), or achieve the same percentage correct answer rate (assigned a '0') (Figure 48). The authors noted that the Proposal group had fewer occurrences of the '-' change value and more occurrences of the '+' change value than the Conventional group.

Kaniwa et al. (2011) also summed assigned change values (+1, -1, 0) for the two training groups, with the Proposal groups' sum equating to -3, compared to the Conventional groups' sum of -13. Based on this, the authors concluded that the adaptive system 'reveal[ed] greater improvements for the low-score band[s]' ('5.2 Analysis of Improvement for Low-Score Bands').

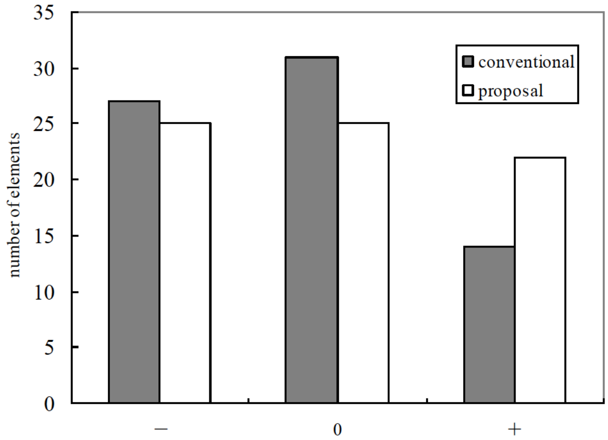

Figure 48. The distribution of the change in correct answer rate week to week for the three lowest scoring bands for the two training groups. (Kaniwa et al., 2011, '5.2 Analysis of Improvement for Low-Score Bands')

Finally, the frequency distributions for the week-to-week differences in the percentage correct answer rates on the low-scoring bands for the two groups were investigated (Figure 49). Kaniwa et al. (2011) note that, compared to the Conventional group, 'the Proposal group showed a higher frequency of positive score difference between 0% and 30%' and 'a lower frequency of negative score difference between -20% and -10%' ('5.2 Analysis of Improvement for Low-Score Bands'). Based on this, they conclude that 'the Proposal group showed a strong tendency to improve their listening in low-score bands' (ibid.). However, the results also show the Proposal group, compared to the Conventional group, exhibited an identical or lower frequency of correct answer rates for the two highest positive score differences (+50% and +100%) and an identical or higher frequency of correct answer rates for the two lowest negative score differences (-50% and -30%). There was no difference between the training groups in the greatest week-to-week differences in score (-50% and +100%). For the second greatest differences (-30% and +50%), the Conventional group outperformed the Proposal group.

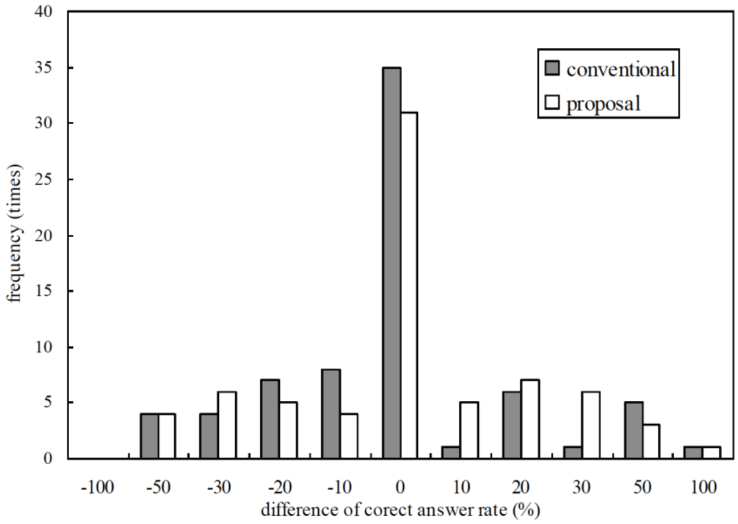

Figure 49. The distribution of the week-to-week difference in correct answer rates bands for the two training groups. (Kaniwa et al., 2011, '5.2 Analysis of Improvement for Low-Score Bands').

In the five negative difference score bands, the Proposal group scored the same (for the -50% and -100% scores), worse (for the -30% score) and better (for the -20% and -10% scores) than to the Conventional group. This equates to the Proposal group scoring better than the Conventional group in 40% of the low score bands. In the other 60% of the low score bands, the Proposal group did not outperform the Conventional group. Despite this, Kaniwa et al. (2011) conclude 'that the Proposal group exhibited better progress in the low-score bands, which indicates that they conducted self-training more efficiently than the Conventional group' ('5.2 Analysis of Improvement for Low-Score Bands').

The Proposal group outperformed the Conventional group in three out of five positive difference score bands. However, the Conventional group improved the most in the highest (+50% and +100%) difference score bands.

Ignoring the 0% (no change) difference score band, and the -100% band, for which no data was recorded for either group, the Proposal group outperformed the Conventional group on four of nine (44.4%) difference score bands. Put another way, the Proposal group did not outperform the Conventional group on the majority (55.6%)

of these nine difference score bands. Nevertheless, Kaniwa et al. (2011) conclude the paper by stating that their 'proposed system could assist the trainees to conduct more effective training by themselves, especially with realistic, therefore more difficult, identification tasks' (6. Conclusions, para. 1).

Kim et al. (2013) detailed a subsequent delivery of the Personalised Timbral Ear Training TETP and an associated test of its effectiveness. The TETP the authors describe was similar to the 2011 version in several ways, but differed in others. The earlier program contained '15–20 samples' per session, whereas the later program contained 25 samples. As well as orchestral music, piano and pink noise, the latter program contained drums as a stimulus. In addition, students had the option of training with their own sound files in the later program. As opposed to the four-week training regime featured in the first TETP, the second program doubled the duration of the training to eight weeks.

The first test Kim et al. (2013) administered was designed to investigate whether any 'inter-subject differences in identifying spectral modifications' were present (p. 425). Four professional audio engineers with more than 10 years' professional experience took part in this test. The participants undertook the Personalised Timbral Ear Training program as a group for an hour at a time, once a week for eight weeks, under the supervision of one of the authors. The eight-week training regime is shown in Table 6. The TETP was delivered via a single loudspeaker, but participants were also instructed to undertake practice sessions with the TETP using headphones.

Table 6. The eight week TETP training regime for the first experiment. (Kim et al., 2013, p. 427)

Week	Training Content
1	12 dB peak (pink noise)
2	12 dB peak (musical sources)
3	6 dB peak (pink noise)
4	6 dB peak (musical sources)
5	12 dB dip (pink noise)
6	12 dB dip (musical sources)
7	12 dB or 6 dB peak, or 12 dB dip (pink noise)
8	12 dB or 6 dB peak, or 12 dB dip (musical sources)

After completing the training, participants 'had to take a test that determined how consistently they remembered the spectral differences in terms of center frequency and gain' (Kim et al., 2013, p. 427). For each of the stimuli used in the test – pink noise, orchestral music, drums, and piano – 25 questions were presented. In each set of 25 questions, one of three possible gain settings (+9 dB, +3 dB, or -12 dB) was applied at the previously mentioned seven ISO octave band centre frequencies. Participants were required to identify the specifics of gain variation applied at each centre frequency.

Although four participants took part in the training and the test, Kim et al. (2013) present the results of only two participants. The authors used the percentage of correct answers averaged over the four stimuli for each of the two participants to determine if there were any inter-subject differences (Figure 50). The authors note that the 'test results highlight that the post-training ability to identify the given spectral modification was not the same for all trainees' (p. 427). As the percentage of correct answers for the two participants was not equal across the seven octave band frequencies, the authors conclude 'that the training would be more effective if it could provide trainees with more exercises in the areas where their identification performance is inconsistent' (pp. 427-428). In this statement, the authors appear to be arguing that an adaptive TETP is justified based on the results of only two participants. It is worth noting that the second participant (Sub2), a professional audio engineer with over 10 years of experience, failed to identify even a single boost at 4k Hz out of 100 questions presented.

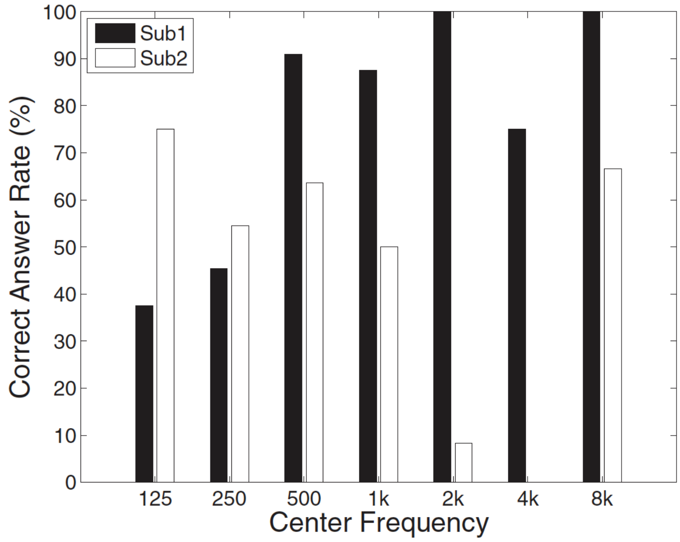

Figure 50. The percentage of correct answers averaged over the four stimuli for each of the two participants. (Kim et al., 2013, p. 428)

As the percentage of correct answers for the two participants was not equal, Kim et al. (2013) 'conducted an experiment that investigated whether such individual differences could be supported by an adaptive training sequence based on an analysis of each trainee's previous performance' (p. 428). This second experiment was previously detailed by Kaniwa et al. (2011).

Kim et al. (2013) stated that 'to fully validate the effectiveness' of their adaptive TETP (designed to replace a human instructor), an experiment needed to be undertaken to compare the performance of three groups of students: a group that trained using a standard non-adaptive TETP, a group that trained using an adaptive TETP, and a group trained by a human instructor (p. 430). The authors conclude the paper by reiterating that because the two participants in the initial study did not achieve equal percentage correct responses across all seven centre frequencies, an adaptive TETP was justified. The authors also reach the same conclusion as in the previous paper regarding the adaptive TETP – that 'personalized and adaptive feedback could assist trainees in conducting more effective training' (p. 430).

Indelicato, Hochgraf, and Kim (2014) reiterate the conclusions of the two previous papers, stating that the adaptive program 'helps students to acquire the critical listening ability more effectively than a traditional teaching method' (p. 4). The authors then detail the results of an experiment designed to 'evaluate the effectiveness of the program for the RIT [Rochester Institute of Technology] audio students' (p. 4).

The experiment utilised a pre/post-training test methodology in which students performed a matching-type task at the beginning and the end of a semester during which they trained using a TETP. The TETP required students to perform an absolute identification task, as previously detailed by Kim et al. (2013). Unlike in previous studies, however, the students were not divided into two training groups; all 36 students undertook the pre-training test, then trained using the adaptive TETP, then took the post-training test. In the pre/post-training tests, Indelicato et al. (2014) measured the students' response times and percentage of correct answers for 22 matching-type questions.

Indelicato et al. (2014) first present the percentage of correct answers averaged for all students, for the eight most difficult questions, as shown in Figure 51. The other questions 'seemed easy enough' so the authors decided to 'not show a difference between before and after the training' (p. 4). The authors summarise the results (Figure 51) by stating that 'that the students were able to complete the given mixing task more precisely after the training' (p. 4).

The response times for the eight most difficult questions are shown in Figure 52. Whilst Indelicato et al. (2014) claim these results show 'that the student completion time decreased in "post-test"', and then conclude that the TETP 'assisted the students in developing "aural sensitivity" that lead [sic] them to judge quickly and precisely', data showing the significance of the difference between the two groups is not presented. It is not clear if the performances of the pre-training and post-training groups were significantly different, as there is considerable overlap between the results.

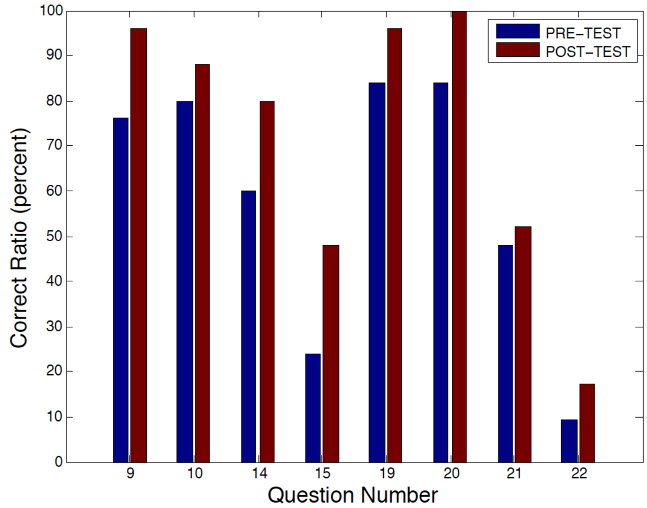

Figure 51. The pre-training and post-training percentage of correct answers for the eight most difficult questions averaged for all participants. (Indelicato et al., 2014, p. 4)

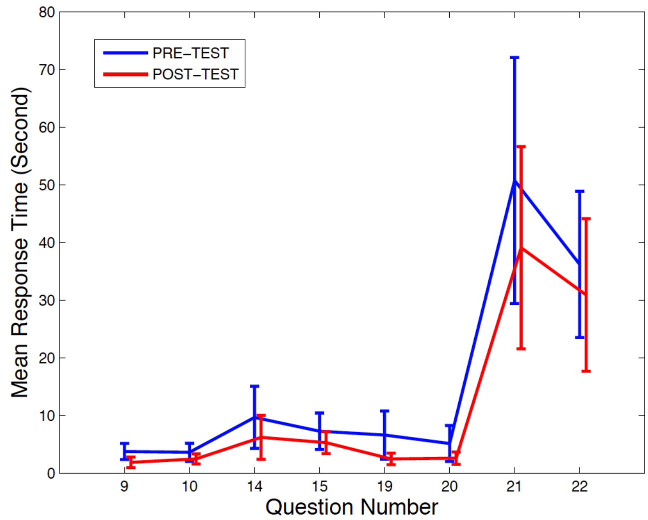

Figure 52. The pre-training and post-training mean response times for the eight most difficult questions averaged for all participants. (Indelicato et al., 2014, p. 4)

Kim (2015) presents the results of an experiment designed ‘to compare the efficacy of an individualized training sequence in the context of audio production’ (p. EL110). In a similar manner to previous experiments designed to investigate the efficacy of the Personalised Timbral Ear Trainer program, Kim implemented a pre/post-training test methodology. A matching-type test was delivered to students before and after they undertook a 15-week TETP. Prior to commencement, the students were divided into two groups of 10; one group trained using a standard, non-adaptive TETP (Control group) and the other trained using the author’s adaptive TETP (Treatment group). Kim then comparted the performance (percentage of correct responses, which he termed ‘accuracy’, and response time) of the two groups on the pre/post-training test. The TETP was identical to that previously detailed by Kim et al. (2013).

The pre/post-training test required students to perform a matching-type task in which they modified the centre frequency and gain of a parametric equaliser applied to pink noise and orchestral music respectively. The first 20 questions in each test featured a single gain modification between +12 dB and -12 dB in 3 dB steps, applied to either an octave or $1/3^{rd}$ octave band centre frequency chosen between 125 Hz and 8k Hz inclusive. The last two questions in each test featured three gain modifications of the type indicated above. Kim collected performance data for the TETP itself and for the pre/post-training tests.

The performance data results from the TETP are shown in Table 7. Kim (2015) notes that the number of questions undertaken in the TETP was not the same for each group, because students were allowed to repeat the training if desired. Kim performed a Chi-square analysis that revealed ‘the grouping did not differentiate the correct answers for both pink noise (p = 0.2819) and orchestra (p = 0.4842)’ (p. EL111). Kim suggests that this result 'impl[ied] that the individualized training method did not result in a different proportion of correct answers over the training period' (ibid.).

Table 7. The number of correct responses and total questions per stimulus for the Treatment and Control groups on the TETP. (Kim, 2015, p. EL111)

	Pink noise		Orchestra	
	Treatment	Control	Treatment	Control
Correct answer	1907	1736	792	890
All question	2512	2329	1488	1714

Kim (2015) also investigated the amount of time each group spent training with the program and the correlation between an individual's training duration and their performance on the TETP. A *t*-test was performed between the two training groups which 'could not reject the null hypothesis that two group means were same (p = 0.8975)' (p. EL111). There was also no significant correlation between the amount of time spent training with the TETP and performance '(r = 0.178, p = 0.4519)' (ibid.).

As with previous studies, Kim (2015) investigated students' performance in the low-scoring frequency bands. However, unlike the results detailed in Kim et al. (2013), the results in this study 'show[ed] that participants with longer training eventually acquired higher correct-answer ratios regardless of whether or not they received individualized training' (Kim, 2015, p. EL112). Kim's analysis of each group's TETP performance data 'did not reveal significant differences' or support the 'efficacy of individualized training' (ibid.).

In order 'to determine whether (1) the participants had improved their high-level sound engineering ability through the given training, and (2) if a significant difference between the two groups existed or not', Kim (2015) compared the mean correct answer ratios (% of correct responses) (Figure 53) and response times (Figure 54) for the two groups on the pre-training test and post-training test (p. EL112).

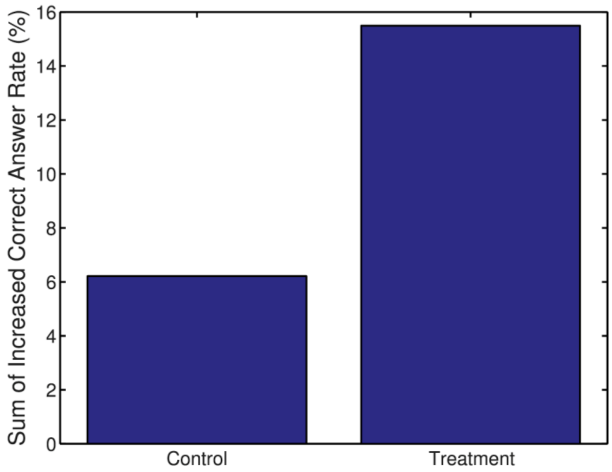

Figure 53. The difference in mean correct answer ratios between the pre-test and post-test for the Control and Treatment groups' (Y-axis is mislabelled 'Sum of…'). (Kim, 2015, p. EL112)

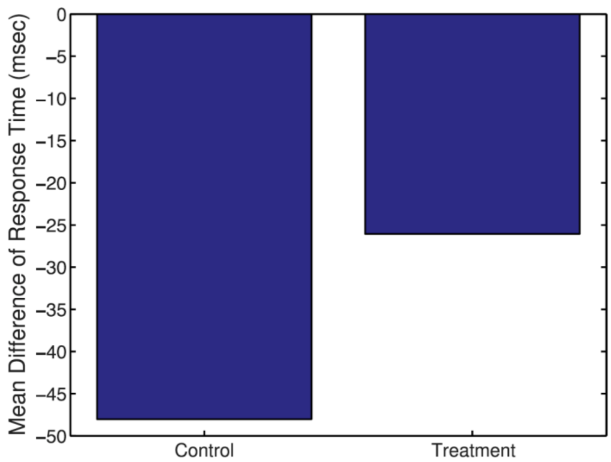

Figure 54. The difference in mean response times between the pre-test and post-test for the Control and Treatment groups' (Y-axis is mislabelled 'Sum of…'). (Kim, 2015, p. EL112)

In reference to these results, Kim (2015) notes that all students appeared to improve in performance, with the data showing 'that both groups scored higher correct answer ratios with shorter response times' (p. EL112). Kim conducted a '2-way mixed-design ANOVA' (Table 8) to investigate differences in performance and found a statistically significant interaction between the two training groups and the correct answer ratio '($F_{1,18}$ = 6.079, p = 0.024)' (ibid.).

Table 8. 2-way mixed-design ANOVA pre/post-training test results. (Kim, 2015, p. EL113)

Source	SS	dF	MS	F	Prob > F
Correct answer ratios					
Pre- and post-test	106.9	1	106.9	1.435	0.246
Interaction	452.9	1	452.9	6.079	0.024
Residuals	1341.1	18	74.5		
Response times					
Pre- and post-test	8599	1	8599	11.206	0.00358
Interaction	758	1	758	0.988	0.333338
Residuals	13811	18	767		

In additional *t*-tests, the increase in correct answer ratios between pre/post-training tests for the Treatment group was found to be significant '(t = -2.539, dF = 9, two-tailed p = 0.031)', but this was not the case for the Control group '(t = 0.91526, dF = 9, two tailed p = 0.3839)' (Kim, 2015, p. EL112). Kim (2013) suggests that 'the results thus support [the claim] that the treatment group participants improved their spectral-matching performance more than the control group' and concludes that 'participants with individualized training had a higher ratio of correct responses than trainees who did not received [*sic*] individualized training' (p. EL112).

Although the TETP detailed by Kaniwa et al. (2011) and Kim et al. (2013) was adaptive in nature, these authors use students' performance on the TETP as evidence of the training program's effectiveness. Specifically, they use the superior performance of students who trained with the adaptive program as evidence that the adaptive TETP was more effective than a standard, non-adaptive TETP. The authors present data which supports their claims that training with an adaptive TETP leads to superior performance on the tasks within the TETP itself, compared to a non-adaptive program. The underlying assumption is that an increase in performance on a TETP task is beneficial in and of itself and has therefore 'improve[d] the listening of sound engineers' (Kaniwa et al., 2011, 'Abstract'). However, the results the authors present only support the claim that an adaptive TETP leads to an *increase in performance on a technical ear training program* and nothing more. The causal link between improved performance on a TETP over time and an increase in a student's critical listening ability cannot be taken as a given; empirical data must be provided to support such a claim. The third criterion has not been met, as the authors measured performance relating to the goals of the training program using the TETP itself.

Recognising that performance on a TETP isn't necessarily evidence of anything other than an improved performance on the TETP, Indelicato et al. (2014) deployed a pre/post-training test methodology to establish the effectiveness of their adaptive TETP. The authors did not divide the students into two training groups, so performance relating to the adaptive TETP cannot be compared to that on a traditional, non-adaptive TETP. Instead, the authors compared all students' performance on a test before and after under undertaking the adaptive TETP. During the TETP, students were trained to perform absolute identification tasks, whereas the pre/post-training test required students to perform a matching-type task. The authors chose absolute identification training because 'it is easier for inexperienced trainees to achieve success' and because they wanted the test 'to be independent from the training' (p. 4). However, the assertion that absolute identification is easier than passive tasks such as matching or removing task, is contradicted by Case (2007, p. 117), and Corey (2010), who declared the absolute identification 'practice mode [as] the most difficult' (p. 40).

Students did not train using a matching task in the TETP (although matching tasks are part of the TETP previously detailed by Kim et al. (2013)); even so, the third criterion is not met as the pre/post-training test was not *significantly different* from a task with which students were familiar. Matching and absolute identification are both examples of identification tasks. Absolute identification tasks are passive identification tasks and do not provide students with aural feedback. They present a flat signal in addition to a modified signal, and students compare the two sounds and identify the modification(s). Matching tasks are active identification tasks, providing students with aural feedback associated with parameter changes. Absolute identification tasks require students to reference sounds stored in their long-term memories and compare these reference frequency anchor points to the presented sounds. Matching tasks require students to compare two presented sounds, switching at any time, and modify one of the sounds until the two are perceived to be identical, and do not require students to recall sounds stored in long-term memory. However, both types of tasks require students to identify the modifications applied to a flat reference signal using an equaliser; this identification is either done via demonstration (matching) or via indication (absolute identification). Instead of *demonstrating* (mimicking) the modification that has been applied to the signal, students *indicate* the modification that has been applied to the signal. Thus, it is argued that the difference between absolute identification tasks and matching tasks detailed by Indelicato et al. (2014) and Kim (2015) is not significant enough to meet the third criterion.

3.2.3.5 Summary

For a TETP to be considered effective, performance relating to the goals of the program must be measured separately from the TETP. In the sections above, the manner in which the authors of the four spectral-based student-targeted TETPs established the effectiveness of their programs based on students' performance is assessed against the third criterion. Despite implementing a pre/post-training test methodology, the three reviewed papers that aimed to establish the effectiveness of the

Timbre Solfège TETP all used performance on an identical or similar task to that on which students trained (Rościszewska, 2011; Rościszewska & Miśkiewicz, 2014, 2015). As such, while meeting the first and second criteria, the Timbre Solfège TETP does not meet the third criterion.

In papers on earlier versions of the Timbral Ear Trainer TETP, Quesnel uses performance on the TETP itself as evidence of the effectiveness of the training program, then in a later publication, uses performance an identical task to one delivered within the TETP. The research described in Quesnel's 2001 paper aimed to establish the effectiveness of the training program by comparing the performance of students who had previously trained with the TETP to that of professional audio engineers who had not. Quesnel uses the fact that the students outperformed the professional engineers on the test as evidence of the TETP's effectiveness. However, the test contained a task which was identical to that on which students had trained previously. In addition to knowing the task, the students (unlike the professional engineers) were familiar with the software interface, the playback system and the acoustic environment. Finally, the software interface used in the test contained a set of unique vowel buttons with which the students had trained, but with which the professional engineers were entirely unfamiliar because such buttons do not exist on professional audio equipment. When these details are taken into account, use of the results of the experiment as evidence for the TETP's effectiveness is invalid. Moreover, even if the differences in performance between the two groups are set aside, the fact remains that the test used to establish effectiveness contained a task on which the students had trained previously. Therefore, the various iterations of Quesnel's Timbral Ear Trainer program do not meet the third criterion, in that the task used to establish the TETP's effectiveness is not sufficiently different from the tasks contained in the TETP.

Authors studying the Beijing University TETP only present results for the post-training test, thus is it not possible to ascertain any improvement in performance as the pre-training results are not public. Nevertheless, the spectral-based task that the authors used to test the effectiveness of the TETP was identical to a task delivered within the TETP itself. Therefore, the Beijing University TETP research does not meet the third criterion, as the authors used a TETP task to establish the effectiveness of the TETP.

In seeking to establish the relative effectiveness of an adaptive TETP in comparison to a non-adaptive training program, Kaniwa et al. (2011) present students' performance (mean and standard deviation of the percentage of correct responses) on two versions of the TETP. Indelicato et al. (2014) and Kim (2015) use students' results on a pre/post-training matching task as evidence of the effectiveness of the TETP, which trained students using an absolute identification task. The Personalised Timbral Ear Trainer program does not meet the third criterion as performance relating to the goal of the TETP was measured on an identical or similar task to that on which students trained.

3.3 Pre/Post-training Test Methodology

Berchtold (2016) makes a distinction between two concepts that the pre/post-training test, or 'test-retest' (comparison between successive measurements) methodology commonly covers; reliability and agreement. He defines reliability as the ability of a test to 'replicate the same ordering between respondents when measured twice' (p. 1). Agreement, however, he defines as the ability of a test to, under the same conditions, provide strictly identical results when applied twice. Berchtold argues that the measurement tools researchers use for longitudinal studies must be evaluated for both reliability and agreement. A test needs to show agreement, he asserts, to be certain that a difference between successive measurements is the result of ‘a real change in the examined person’ and not ‘a random variation or a systematic deviation due to the test itself or the conditions of the testing’ (p. 2). A test needs reliability, on the other hand, so that comparisons can be made between different respondents. Reliability and agreement can be ascertained by administering the test twice to the same group of people under the same conditions.

Berchtold cautions that a change over time in such a test-retest setup ‘does not necessarily imply a real evolution of the respondents’ (p. 2) because the change could also be the result of a measurement error, caused by respondents learning from the first test to the second one. This is especially pertinent in studies in which researchers use an increase in performance over time on a TETP as evidence that the student has improved their critical listening skills. Demonstrated improvements over time on a TETP could be explained by students getting better at using the training interface, as opposed to improving their critical listening abilities.

Ideally a control group that does not undertake the TETP should be included in a pre/post-training test methodology to establish what change, if any, occurs between tests as a result of other endeavours outside of the TETP. It is reasonable to assume that students enrolled in an audio engineering education program are unlikely to be undertaking a TETP in isolation, as students commonly enrol in multiple subjects simultaneously. As such, students can be expected to be undertaking a variety of other critical listening tasks, such as mixing and microphone placement, concurrently to the TETP. A control group in this scenario aids in establishing the effect of these non-TETP-based tasks on the students’ critical listening abilities.

3.4 SUMMARY

It has been demonstrated that there is a lack of consensus between published papers, authors and professionals regarding the most effective method to train students to operate equalisers. However, there are no published studies specifically investigating the effectiveness of either the ICA or ISA training method in general or in the context of a TETP. Although Quesnel (2001) encouraged students in the Timbral Ear Trainer I program to use the ICA method and Rościszewska and Miśkiewicz (2015) encouraged Timbre Solfège students to use the ISA method, these authors did not specifically investigate the difference between the two methods. Quesnel did examine the overall number of parameter adjustments, however, as participants' frequency adjustments were not examined in isolation, is not possible to ascertain whether the professionals made fewer adjustments to the frequency control than the students. The absence of research relating to the topic stands in contrast to the large number of texts promoting the ICA method and those that caution against its use. This review shows that there is no published evidence that the ICA method lacks 'accuracy and cohesiveness' (Stavrou, 2003, pp. 166-167), for example, or that the ISA method will 'strengthen your ear while maintaining your perspective and sensitivity' (ibid.).

Many authors recognise the usefulness of pre/post-training tests in endeavouring to demonstrate an improvement in performance on a task relating to the goals of the TETP (Indelicato et al., 2014; Kim, 2015; Liu et al., 2007a; Quesnel & Woszczyk, 1994; Rościszewska, 2011; Rościszewska & Miśkiewicz, 2014, 2015). The pre-training test serves to establish a baseline level of performance on a task prior to undertaking the TETP, which is then compared to the post-training test results. A change in performance is then often cited as evidence supporting claims that the TETP has caused a change in the critical listening abilities of the student. If the pre/post-training test requires the student to perform a task that is relevant to their future career and relates to the goals of the TETP, an improvement in performance after training may be used as evidence of the effectiveness of the TETP. Best practice suggests that the reliability and agreement of the test should also be established in order to ensure such conclusions are valid. However, it is also evident that the TETP should not train students in how to take the pre/post-training test. The tasks contained in the TETP must be independent and sufficiently different from those contained in the pre/post-training tests, as recognised by Indelicato (2014) for example, who 'wanted to measure the students' performance in unexpected situations' (p. 4). The review of the literature presented in this chapter shows that although researchers favour the pre/post-training test methodology in seeking to establish the effectiveness of their TETPs, the tests they use do not support their claims due to the similarity between the tasks in the TETP and the task in the pre/post-training test.

This chapter highlights the need to develop an appropriately independent test of skill transfer to provide more rigorous testing of the effectiveness of TETPs. Chapters 4, 5 and 6 detail the development of such independent tests in isolation, then in a pilot study within the context of a TETP, followed by a full-scale study.

4 Developing Tone Colour Discrimination Tests

As detailed in the literature review, one of the criteria required to establish the effectiveness of a TETP is that performance relating to the goals of the program must be measured on a task that is significantly different to any task within the TETP. This chapter details the development of three tone colour discrimination tests that were deployed separately from a TETP. The tests were trialled to gauge their effectiveness in showing differences in participants' tone colour discrimination ability, with the ultimate goal of deploying them in a pre/post-training scenario. The following experiments were carried out with Human Research Ethics Committee approval from the University of Sydney (Appendix A).

4.1 Tic Tac® Study

The ability to discriminate between sounds differing in tone colour is partly dependent on the ability to discriminate between perceptual attributes that vary with the physical attribute of sound frequency. The general hypothesis is that, if listeners, through training, can develop and increase their ability to discriminate frequencies, they can also develop and increase their ability to discriminate tone colour, which as previously argued, is a professionally relevant skill. The Tic Tac® study was designed to determine a listener's ability to discriminate between stimulus frequencies without asking them to specifically identify those frequencies. Its design was founded on the assumption that the ability to accurately determine how many pellets were being shaken in a container depended on the listener's analysis of the spectral energy distribution. For a review of the use of event-based percussive timbres in listening tests see Martens, McKinnon-Bassett and Cabrera (2012). By manipulating the container size but leaving the number of pellets constant, the test effectively manipulates spectro-temporal patterns in the lower portion of the stimulus spectrum. It was therefore expected that the box size would significantly influence the listener's ability to accurately determine the number of pellets being shaken. The Tic Tac® study's specific hypothesis was that if a TETP increases a listener's frequency discrimination abilities, their performance on the post-training test should improve in comparison to their pre-training scores.

4.1.1 METHOD

4.1.1.1 DATES, PARTICIPANTS

Eighteen students enrolled in the Master of Design Science (Audio and Acoustics) 'Digital Audio' class at the University of Sydney in 2011 voluntarily participated in the experiment. The researcher recruited students at the commencement of the class, with the experiment taking place approximately halfway through the semester. The class had 24 students, so the participation rate was 75%. All students were at least 18 years of age. No students reported any hearing loss.

The experiment took place during the Digital Audio class on the 13th April 2011 in a computer lab at the University. The Digital Audio class took place between 6pm and pm on a weeknight. Students performed the test individually and concurrently with other students in the class.

4.1.1.2 STIMULUS

Fifty recordings were made as Tic Tac®[8] pellets were shaken within two different-sized plastic containers, referred to here as 'small' and 'large'. The large container measures 80.7 mm x 42.5 mm x 19.5mm, the small container measures 62mm x 39.4mm x 15.1mm. The recordings were made in an anechoic chamber using an omnidirectional microphone positioned approximately 0.5m from the container being shaken. The author's research supervisor shook the containers. Separate recordings were made as each of the two containers was shaken with the number of pellets inside varying in increments of one from one to 25 (hence the 50 recordings). The maximum of 25 pellets was chosen as this was the capacity of the small container. For each recording, the container was shaken for approximately one second. Figure 55 shows the temporal envelopes (normalised power over time) that resulted when each container was shaken five times within one second when filled with 10 pellets.

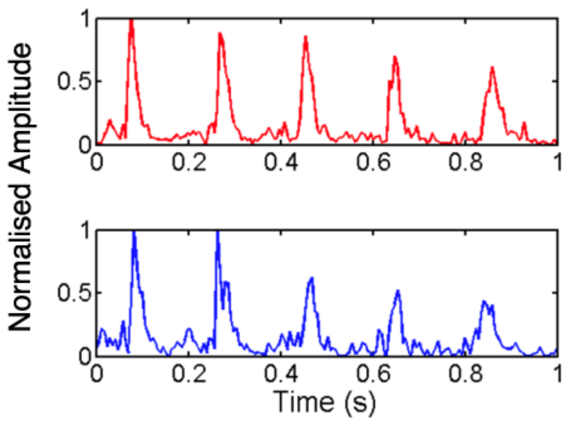

Figure 55. Normalised temporal envelopes for the small (red) and large (blue) and containers when filled with 10 pellets and shaken five times.

[8] A Tic Tac® is a small, hard, pill-shaped candy mint.

The frequency response of the sound generated when shaking pellets within a small plastic box is largely dependent on the number of pellets and the size of the box. Other factors affecting the frequency response of the generated sound include the velocity of the pellets when they impact the box and other pellets (how hard the box is being shaken), how the box is held and whether or not the hand shaking the box is attenuating certain frequencies. Each box has a fundamental resonant structure, which does not change with the number of pellets. The smaller box has two relatively narrow spectral peaks at approximately 1k Hz and 4k Hz, whereas the larger box has one broad spectral peak from approximately 1.5k Hz to 2.5k Hz. As the number of pellets change, a second, higher-frequency resonance also changes. In this way, the frequency response of the shaken box of pellets is analogous to the voicing of vowel sounds, whereby a low frequency formant (resonant boost) is combined with a higher formant and applied to a glottal pulse in order to generate the various vowel sounds of human speech.

A linear predictive coding (LPC) analysis of the recordings was performed to document the spectral characteristics of the two containers. As the LPC spectra for all 25 recordings per container were similar, the resulting filter coefficients were averaged for the small and large containers. Figure 56 shows the average LPC spectra for the small and large containers based on the average filter coefficients for each container.

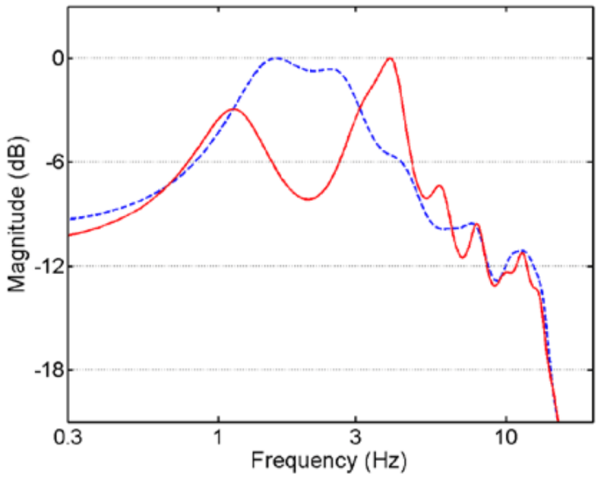

Figure 56. The average LPC spectra for all pellets for the small (red) and large (blue) containers.

Although the loudness varied over time throughout the duration of each recording due to shaking, each of the 50 recordings was approximately matched in loudness using the Loudness Meter Comparison Utility (Australian Broadcasting Corporation, 2006). Fifty sounds were presented at the standard level detailed above, 50 were presented at a level of +3 dB relative to the standard level, and 50 were presented at a level of -3 dB relative to the standard level, totalling 150 sounds (Tables 9-10) to test for the effects that small increments in reproduction level might have on estimates. The presentation order of sounds was randomly chosen from the pool of 150 sounds without replacement.

Table 9. Specifics of the 75 small container stimuli.

Scenario No.	Box Size	No. of Pellets	Gain (dB)	Scenario No.	Box Size	No. of Pellets	Gain (dB)	Scenario No.	Box Size	No. of Pellets	Gain (dB)
1	Small	1	0	**26**	Small	1	-3	**51**	Small	1	3
2	Small	2	0	**27**	Small	2	-3	**52**	Small	2	3
3	Small	3	0	**28**	Small	3	-3	**53**	Small	3	3
4	Small	4	0	**29**	Small	4	-3	**54**	Small	4	3
5	Small	5	0	**30**	Small	5	-3	**55**	Small	5	3
6	Small	6	0	**31**	Small	6	-3	**56**	Small	6	3

7	Small	7	0	**32**	Small	7	-3	**57**	Small	7	3
8	Small	8	0	**33**	Small	8	-3	**58**	Small	8	3
9	Small	9	0	**34**	Small	9	-3	**59**	Small	9	3
10	Small	10	0	**35**	Small	10	-3	**60**	Small	10	3
11	Small	11	0	**36**	Small	11	-3	**61**	Small	11	3
12	Small	12	0	**37**	Small	12	-3	**62**	Small	12	3
13	Small	13	0	**38**	Small	13	-3	**63**	Small	13	3
14	Small	14	0	**39**	Small	14	-3	**64**	Small	14	3
15	Small	15	0	**40**	Small	15	-3	**65**	Small	15	3
16	Small	16	0	**41**	Small	16	-3	**66**	Small	16	3
17	Small	17	0	**42**	Small	17	-3	**67**	Small	17	3
18	Small	18	0	**43**	Small	18	-3	**68**	Small	18	3
19	Small	19	0	**44**	Small	19	-3	**69**	Small	19	3
20	Small	20	0	**45**	Small	20	-3	**70**	Small	20	3
21	Small	21	0	**46**	Small	21	-3	**71**	Small	21	3
22	Small	22	0	**47**	Small	22	-3	**72**	Small	22	3
23	Small	23	0	**48**	Small	23	-3	**73**	Small	23	3
24	Small	24	0	**49**	Small	24	-3	**74**	Small	24	3
25	Small	25	0	**50**	Small	25	-3	**75**	Small	25	3

Table 10. Specifics of the 75 large container stimuli.

Scenario No.	Box Size	No. of Pellets	Gain (dB)	Scenario No.	Box Size	No. of Pellets	Gain (dB)	Scenario No.	Box Size	No. of Pellets	Gain (dB)
76	Large	1	0	**101**	Large	1	-3	**126**	Large	1	3
77	Large	2	0	**102**	Large	2	-3	**127**	Large	2	3
78	Large	3	0	**103**	Large	3	-3	**128**	Large	3	3
79	Large	4	0	**104**	Large	4	-3	**129**	Large	4	3
80	Large	5	0	**105**	Large	5	-3	**130**	Large	5	3
81	Large	6	0	**106**	Large	6	-3	**131**	Large	6	3
82	Large	7	0	**107**	Large	7	-3	**132**	Large	7	3
83	Large	8	0	**108**	Large	8	-3	**133**	Large	8	3
84	Large	9	0	**109**	Large	9	-3	**134**	Large	9	3

85	Large	10	0	**110**	Large	10	-3	**135**	Large	10	3
86	Large	11	0	**111**	Large	11	-3	**136**	Large	11	3
87	Large	12	0	**112**	Large	12	-3	**137**	Large	12	3
88	Large	13	0	**113**	Large	13	-3	**138**	Large	13	3
89	Large	14	0	**114**	Large	14	-3	**139**	Large	14	3
90	Large	15	0	**115**	Large	15	-3	**140**	Large	15	3
91	Large	16	0	**116**	Large	16	-3	**141**	Large	16	3
92	Large	17	0	**117**	Large	17	-3	**142**	Large	17	3
93	Large	18	0	**118**	Large	18	-3	**143**	Large	18	3
94	Large	19	0	**119**	Large	19	-3	**144**	Large	19	3
95	Large	20	0	**120**	Large	20	-3	**145**	Large	20	3
96	Large	21	0	**121**	Large	21	-3	**146**	Large	21	3
97	Large	22	0	**122**	Large	22	-3	**147**	Large	22	3
98	Large	23	0	**123**	Large	23	-3	**148**	Large	23	3
99	Large	24	0	**124**	Large	24	-3	**149**	Large	24	3
100	Large	25	0	**125**	Large	25	-3	**150**	Large	25	3

4.1.1.3 Procedure

For each of the 150 sounds presented, subjects were required to indicate how many pellets they thought were in the container. Subjects indicated their estimates of the number of pellets by adjusting a slider on the graphical user interface (GUI), which permitted responses from 1 to 25 inclusive (Figure 57). Each presented sound could be replayed by the subject during the trial if desired. Prior to commencing the experiment, each subject completed a practice session that contained 10 randomly sampled questions from the 150 stimuli. The GUI and underlying software program stored each subject’s responses in a text file that was saved at the conclusion of the experiment. The 0dB sounds were reproduced over Sennheiser HD428 headphones at approximately 78 dB Sound Pressure Level (A-weighted) – approximately the same as the level that the listener would have been exposed to had they occupied the position of the microphone during the recordings.

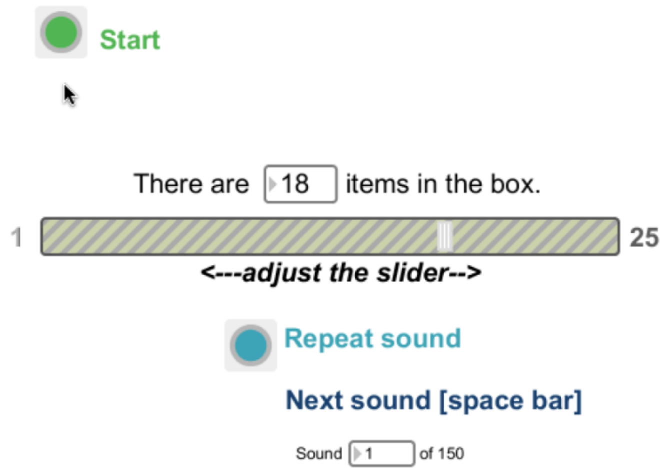

Figure 57. The GUI used to store each subject's estimates of pellet numbers.

4.1.2 Results

The mean estimated number of pellets for the 18 subjects was plotted as a function of the actual number of pellets for the standard +3 dB and +6 dB levels for both containers (Figure 58). Responses located on the dashed line indicate that the mean estimate was identical to the actual number of pellets for that relative reproduction level and container size. As shown, the mean estimated number of pellets generally increased as the actual number of pellets increased, for all reproduction levels and container sizes. Estimates for 1–5 pellets were relatively accurate, differing from the actual number by a maximum of one. Estimates with an actual number of pellets over 10 grew increasingly inaccurate. It is evident by the lack of symbol patterns that the variation in relative reproduction levels yielded no significant difference in the mean estimates. Similarly, container size was not associated with a statistically significant difference in mean estimates.

A single regression analysis was performed to detail the relationship between the mean estimated number of pellets and the actual number of pellets for all 150 stimuli. The best-fitting linear function accounted for 61.5% of the estimated variance, but applying a compressive power function with an exponent of b = 0.89 allowed the model to account for 84.8% of the variance. The shape of this compressive power function is shown by the solid line in Figure 58, and its formula is shown in Equation 3.

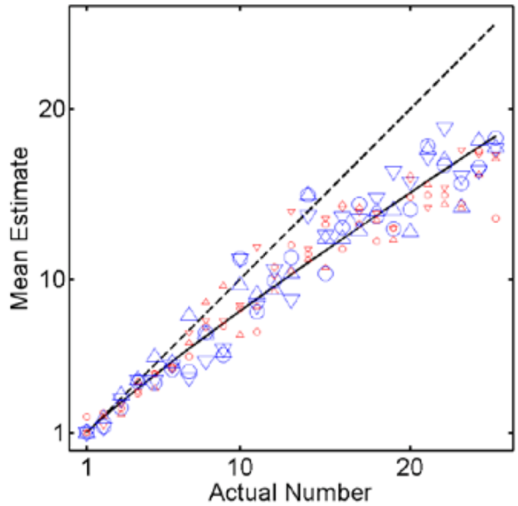

Figure 58. The mean estimates of pellets versus the actual number of pellets reproduced at standard (Δ), +3 dB (O) and =6 dB (∇) levels for the small (red) and large (blue) containers.

$$y = 1.05x^{0.89}$$

Equation 3. The compressive power function that best fit the data, where x is the actual number of items and y is the estimated number of items.

4.2 Revised Tic Tac® Study

A revised version of the Tic Tac® study was implemented using a similar procedure to the first study, with fewer recordings and a filter introduced to modify the spectral energy distribution. This revised study was designed to investigate what effect, if any, varying the spectral energy distribution would have on estimates of the number of pellets being shaken inside various-sized containers.

4.2.1 Method

4.2.1.1 Participants

Thirty-five students enrolled in an undergraduate Sound Design class at the University of Sydney in 2011 voluntarily participated in the experiment. The researcher recruited students at the commencement of the class, with the experiment taking place approximately halfway through the semester. The class had 38 students, so the participation rate was 92%. All students were at least 18 years of age. No students reported any hearing loss.

The experiment took place during the class on the 12th September 2011 in a computer lab at the University, at approximately 11am on a weekday. Students performed the test individually and concurrently with other students in the class.

4.2.1.2 Stimulus

A subset of the 50 recordings from the first Tic Tac® experiment was used in this experiment. Recordings of 5, 10, 15, and 20 pellets being shaken inside both the small and large containers were used, resulting in a total of eight recordings. Each of these eight recordings was reproduced in the experiment in one of three states: original recording ('original spectrum'), +12 dB high shelf applied at 1k Hz ('high-frequency boost'), or a -12 dB high shelf applied at 1k Hz ('high-frequency attenuation'). This processing resulted in 24 stimuli. For each stimulus, the loudness was matched before the application of the filter. A 12 dB trial-to-trial variation in overall level was also introduced for the 'high-frequency boost' and 'high-frequency attenuation' sounds, resulting in 40 stimuli sounds. The stimuli were each presented twice, resulting in a total of 80 stimuli for the experiment.

4.2.1.3 Procedure

The procedure for this experiment was identical to that for experiment 1.

4.2.2 Results

The mean estimated number of pellets for each of the three filter types for 5, 10, 15 and 20 pellets for the small container is shown in Figure 59, and in Figure 60 for the large container. The 12 dB trial-to-trial variation in overall level was found to have no significant effect on the estimated number of pellets and was therefore not included in the presented figures. The mean estimated number of pellets for the small container, shown in Figure 59, increased linearly as a function of the actual number of pellets. However, the mean estimated number of pellets for the large container shown in Figure 60 increased in a non-linear manner. The results for the large and small containers show that the three spectral modifications introduced resulted in different mean estimates. The high frequency boost results in an increase in the mean estimated number of pellets, and conversely, the high frequency attenuation causes a decrease in mean estimates. A t-test confirmed that the means were significantly different between the three spectral categories for each number of actual pellets as shown in Table 11. These results suggest that this test is a suitable assessment of tone colour discrimination ability that is independent of the TETP.

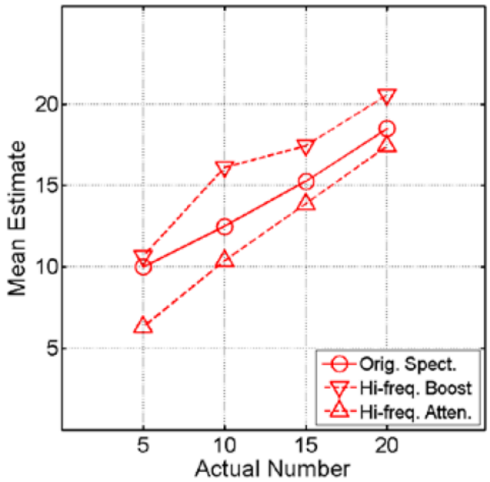

Figure 59. The mean estimates of pellets versus the actual number of pellets for the small container.

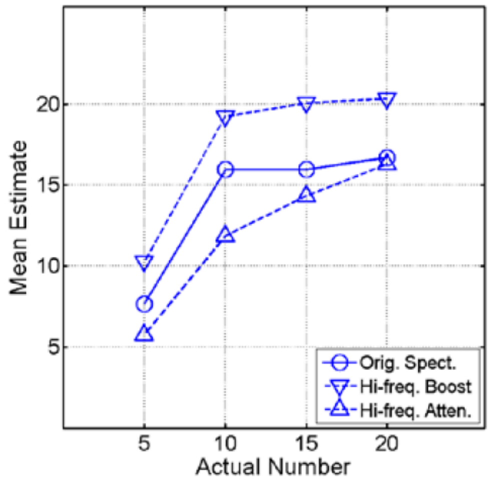

Figure 60. The mean estimates of pellets versus the actual number of pellets for the large container.

Table 11 - Significance of mean differences.

Box Size	**Actual No. of Pellets**	***p***
Small	5	0.019
Small	10	0.002
Small	15	0.000
Small	20	0.000
Large	5	0.003
Large	10	0.000
Large	15	0.000
Large	20	0.000

4.3 VOWEL CATEGORISATION STUDY

This study explored students' ability to discriminate and subsequently categorise artificial vowel sounds on two separate continuums. The vowel categorisation study presented students with a series of sounds that they assigned to one of two categories in a 2AFC scenario. The artificially generated vowel sounds mimic the sound of various vowels in human speech. The hypothesis was that the ability to discriminate between perceived resonances is dependent upon the ability to discriminate the physical parameter frequency and as previously argued, tone colour discrimination is a professionally relevant skill.

4.3.1 METHOD

4.3.1.1 PARTICIPANTS

Fourteen students enrolled at the University of Sydney voluntarily participated in two related experiments in June 2011. All students were enrolled in the Design Science undergraduate program and were at least 18 years of age at the time of the experiment. The researcher recruited students at the commencement of the class, with the experiment taking place approximately halfway through the semester. The total number of students in the class was 16, so the participation rate was 88%. No students reported any hearing loss.

The experiment took place during the class on the 1st June 2011 in a computer lab at the University at approximately 11am on a weekday. Students performed the test individually and concurrently with other students in the class.

4.3.1.2 STIMULUS

For the first experiment, a series of synthetic vowel sounds were generated by applying a 300 Hz bandpass filter (F1) to an artificial glottal pulse. A secondary bandpass filter was then applied (F2), generating a series of vowel-like sounds that ranged from /u/ as in 'hoot' (F1=300 Hz, F2=800 Hz) through to /i/ as in 'heat' (F1=300 Hz, F2=2400 Hz). The F1 filtering of the glottal pulse was applied in Matlab and the F2 filtering was applied in real time using a program developed in Max/MSP[9]. The nine frequencies between the /u/ and /i/ sounds were generated at an even logarithmic spacing as shown in Figure 61, producing a total of 11 sounds.

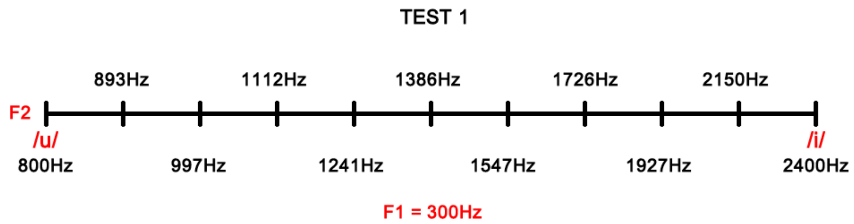

Figure 61. The 11 sounds in the /u/ - /i/ range.

Each of the 11 sounds was presented in two states – normal gain and +3 dB gain – resulting in 22 presentation scenarios (Table 12) to test for the effects that small increments in reproduction level might have on estimates. Each set of 22 sounds was presented three times, totalling 66 sounds. The presentation order of the 66 sounds was randomised without replacement.

[9] Max (formerly Max/MSP) is visual programming language developed by Cycling '74 in San Francisco, CA.

Table 12. The parameters of the 22 stimuli used in the '/u/ versus /i/' experiment.

Stimulus no.	Filter 1 (F1)	Filter 2 (F2)	Gain
1	300 Hz	800 Hz	0 dB
2	300 Hz	893 Hz	0 dB
3	300 Hz	997 Hz	0 dB
4	300 Hz	1112 Hz	0 dB
5	300 Hz	1241 Hz	0 dB
6	300 Hz	1386 Hz	0 dB
7	300 Hz	1547 Hz	0 dB
8	300 Hz	1726 Hz	0 dB
9	300 Hz	1927 Hz	0 dB
10	300 Hz	2150 Hz	0 dB
11	300 Hz	2400 Hz	0 dB
12	300 Hz	800 Hz	+3 dB
13	300 Hz	893 Hz	+3 dB
14	300 Hz	997 Hz	+3 dB
15	300 Hz	1112 Hz	+3 dB
16	300 Hz	1241 Hz	+3 dB
17	300 Hz	1386 Hz	+3 dB
18	300 Hz	1547 Hz	+3 dB
19	300 Hz	1726 Hz	+3 dB
20	300 Hz	1927 Hz	+3 dB
21	300 Hz	2150 Hz	+3 dB
22	300 Hz	2400 Hz	+3 dB

For the second experiment, a series of synthetic vowel sounds were generated by applying a 700 Hz bandpass filter (F1) to an artificial glottal pulse. A secondary bandpass filter was applied (F2) at one of 11 F2 filters that ranged from /ah/ as in 'hot' (F1=700 Hz, F2=900 Hz) to /ae/ as in 'hat' (F1=700 Hz, F2=1800 Hz). The nine frequencies between the /ah/ and /ae/ sounds were generated at an even logarithmic spacing as shown in Figure 62, producing a total of 11 sounds. The gain processing was identical to that used in the first experiment (Table 13). The presentation of stimuli was again randomised.

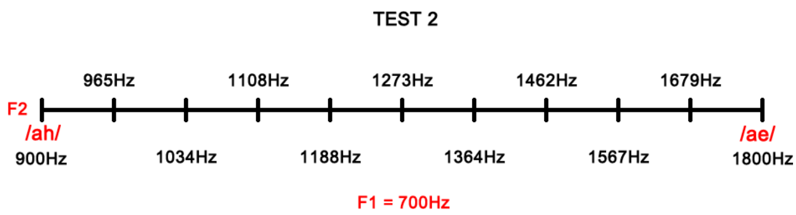

Figure 62. The 11 sounds in the /ah/ - /ae/ range.

4.3.1.3 Procedure

Prior to commencing each experiment, participants undertook an orientation session designed to familiarise them with the task and the user interface. In the first orientation session, the /u/ (F1=300 Hz, F2=800 Hz) and /i/ (F1=300 Hz, F2=2400 Hz) sounds were presented separately and for each sound the user was required to respond accordingly using the GUI. The orientation sessions for the /ah/ (F1=700 Hz, F2=900 Hz) and /ae/ (F1=700 Hz, F2=1800 Hz) were identical except for the sounds presented. The sounds were reproduced over Sennheiser HD428 headphones at approximately 78 dB Sound Pressure Level (A-weighted).

Table 13. The parameters of the 22 stimuli used in the '/a/ versus /ae/' experiment.

Stimulus no.	Filter 1 (F1)	Filter 2 (F2)	Gain
1	700 Hz	900 Hz	0 dB
2	700 Hz	965 Hz	0 dB
3	700 Hz	1034 Hz	0 dB
4	700 Hz	1108 Hz	0 dB
5	700 Hz	1188 Hz	0 dB
6	700 Hz	1273 Hz	0 dB
7	700 Hz	1364 Hz	0 dB
8	700 Hz	1462 Hz	0 dB
9	700 Hz	1567 Hz	0 dB
10	700 Hz	1679 Hz	0 dB
11	700 Hz	1800 Hz	0 dB
12	700 Hz	900 Hz	+3 dB
13	700 Hz	965 Hz	+3 dB
14	700 Hz	1034 Hz	+3 dB
15	700 Hz	1108 Hz	+3 dB
16	700 Hz	1188 Hz	+3 dB
17	700 Hz	1273 Hz	+3 dB
18	700 Hz	1364 Hz	+3 dB
19	700 Hz	1462 Hz	+3 dB
20	700 Hz	1567 Hz	+3 dB
21	700 Hz	1679 Hz	+3 dB
22	700 Hz	1800 Hz	+3 dB

The first experiment featured a 2AFC test in which participants were required to categorise each of the 66 vowel-like sounds as either perceptually more similar to /u/ or to /i/. Participants indicated their choice by clicking on either the /u/ or the /i/ button on the GUI (Figure 63). Participants were only permitted to hear each sound once. The GUI, developed using Max/MSP software, stored each subject's responses in a text file that was saved at the conclusion of the experiment. After completing the first experiment, the participants then immediately undertook the second experiment, which was identical to the first with the exception of the stimuli and the user interface text, shown in Figure 64.

Figure 63. The GUI used in the '/u/ versus /i/' experiment.

Figure 64. The GUI used in the '/ah/ versus /ae/' experiment.

4.3.2 RESULTS

As participants were asked to categorise the various sounds within the 'vowel scale' for each test (vowel pairing), the proportion of correct responses can be plotted on a graph with F2 represented by the x-axis and the categorisation choices represented by the y-axis. The results of the first experiment for one participant are shown in Figure 65. This plot shows the average proportion of /i/ responses versus F2 values. The results of the second experiment – the average proportion of /ae/ responses versus F2 values for the same participant – are shown in Figure 66. Logistic regression analysis (LRA) was then employed to fit a curve to the subject's responses for each experiment. LRA produces an 's'-shaped function called a logit function that best fits the data

(Equation 4). The 0.5 response proportion point (μ) was calculated by taking the ratio of the two parameters of the equation.

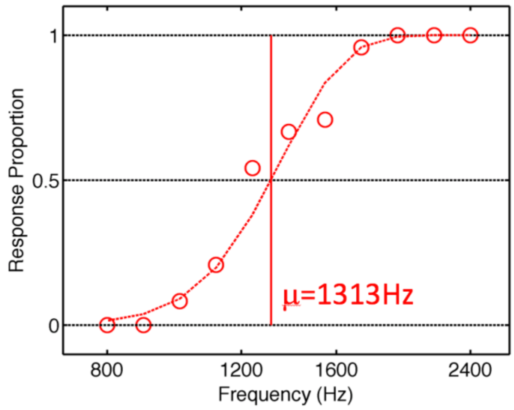

Figure 65. The /i/ response proportions for one subject (F1 = 300 Hz) with the 0.5 proportion point (μ) represented by the vertical red line.

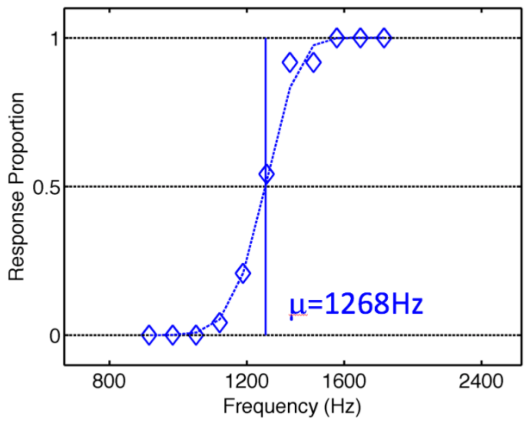

Figure 66. The /ae/ response proportions for one subject (F1 = 700 Hz) with the 0.5 proportion point (μ) represented by the vertical blue line.

$$y = \frac{1}{1 + e^{-(b_0 + b_1 x)}}$$

Equation 4. The logit function. x = a continuous independent variable, y = observed binary dependent variable

The /i/ response proportions for the single participant for the lower F2 frequencies were distributed evenly along the fitted logit function, with the lowest two

F2 frequencies assigned to the /u/ category (Figure 65). The highest three F2 frequencies were assigned to the /i/ category. In contrast, the /ae/ response proportions for the participant are less evenly distributed. The responses for the five lowest and five highest F2 frequencies are each assigned to the /ah/ and /ae/ limits of the scale respectively, with a single response located near the 0.5 response proportion point.

The /i/ and /ae/ response proportions for all 14 participants are shown in Figure 67. Gain levels of 0 dB and +3 dB are indicated by the red and blue markers respectively. It is evident that the introduction of the variations in gain (0 dB and +3 dB) did not affect the response proportions in any significant way.

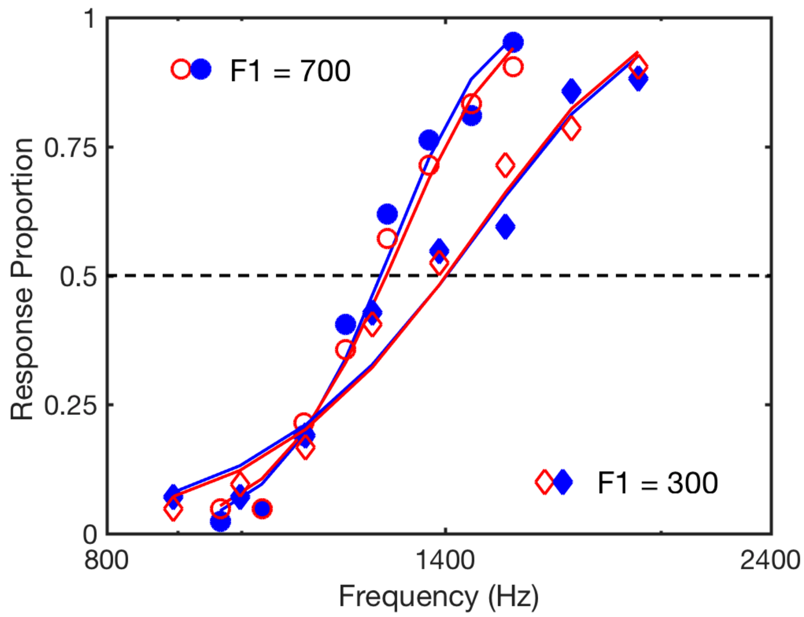

Figure 67. The /i/ (diamond) versus /ae/ (circle)response proportions averaged for all subjects for 0 dB (red) and +3 dB (blue).

It is commonplace within TET research to associate a parametric boost in the spectrum with a single vowel sound. This approach assumes that a spectrally flat input signal is sent to the parametric filter. When such a flat input signal is present, associating a single parametric boost with a single vowel sound may yield a consistent association as shown in Table 14. However, these results indicate that if the input signal contains a resonance, the resulting sound is similar to a different set of vowel sounds. For example, in the Timbre Solfege TETP, Letowski and Miśkiewicz associated a parametric boost at 250 Hz with the vowel /u/ (as in tool). However, these results show that users associated a boost at 800 Hz applied to a stimulus with an existing resonance at 300 Hz with the /u/ vowel sound. Changing the F2 value slightly generates a completely different vowel sound if the existing resonance is changed to 700 Hz. It is evident that the association of a single parametric filter boost with a single vowel sound is applicable only when the input signal is spectrally flat. If the input signal contains existing resonances, which musical material does, the perceived vowel sound may change, rendering the previously learned associations shown in Table 14 inaccurate.

Table 14. Relationships between formant frequencies and corresponding vowel sounds. (Letowski & Miśkiewicz, 1995)

Centre Frequency	**Vowel sound**	**Example**
250 Hz	/u/	Boot
500 Hz	/o/	Boat
1000 Hz	/a/	Bottle
2000 Hz	/ε/	Bet
4000 Hz	/i/	Beat

4.4 Conga Tone Dissimilarity Study

The conga tone dissimilarity study presented participants with a series of pairwise comparisons of nine synthetic percussive timbres (conga drum sounds). For each trial, participants were required to first discriminate between the two presented sounds, then provide pairwise dissimilarity judgements. The dissimilarity data was combined and submitted for individual differences scaling (INDSCAL) analysis, 'an implementation of MDS (Multi-Dimensional Scaling) that takes subjects' individual differences into account; it rotates the derived stimulus space to best fit individual weights on each of the dimensions' (Martens & Marui, 2005). This produced a two-dimensional common timbre space that indicated the relative weight each subject placed on each dimension. It is suggested that after undertaking a TETP participants' individual differences (subject weights) in the percussive timbre dimensions will change, based on their ability to discriminate tone colour, which as previously argued, is a professionally relevant skill.

4.4.1 METHOD

4.4.1.1 PARTICIPANTS

Seventeen students enrolled at the SAE Institute in Sydney voluntarily participated in the experiment in December 2011. All students were enrolled in the undergraduate Bachelor of Audio Production degree and were at least 18 years of age at the time of the experiment. The researcher recruited students at the commencement of the class, with the experiment taking place at the end of the semester. The total number of students in the class was 18, so the participation rate was 94%. No students reported any hearing loss.

The experiment took place during the class on the 2nd December 2011 in a computer lab at the SAE Institute at approximately 3pm on a weekday. Students performed the test individually and concurrently with other students in the class.

4.4.1.2 STIMULUS

To generate the nine synthetic conga tones used in the experiment, a monophonic recording of a conga drum being struck was first analysed using LPC. The LPC spectra that resulted from this analysis are indicated by the blue lines in Figures 68 and 69. An artificial strike transient was then generated by applying an exponential decay to a burst of white noise. This artificial transient was filtered using two 2-pole filters to produce resonant peaks at two frequencies, F1 and F2. Varying the parameters of the applied filters provided control over the resultant tone colour of the strike sounds. For all sounds, F1 was fixed at 276 Hz, but the Q of the F1 filter was varied (4, 10, 20). The F2 frequency value was also varied (731 Hz, 1032 Hz, 1683 Hz), with a fixed Q of 4. The combination of these parameters resulted in the generation of nine synthetic conga-like tones, whose spectra are shown in Figures 68, 69 and 70. The parameters of each strike tone are shown in Table 15. The loudness of the tones was matched using the Loudness Meter Comparison Utility (Australian Broadcasting Corporation, 2006). Each triad of sounds [1 2 3], [4 5 6], and [7, 8, 9] (Figures 68, 69, and 70) had the same F1 values but differing F2 values. Within each triad, the F1 and F2 values were identical, but the Q was varied (Table 15).

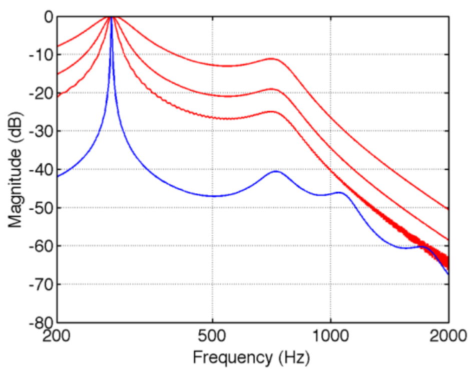

Figure 68. The red lines represent the low-order filters applied to sounds 1–3, the blue line represents the LPC spectra that resulted from the analysis of the monophonic conga recording.

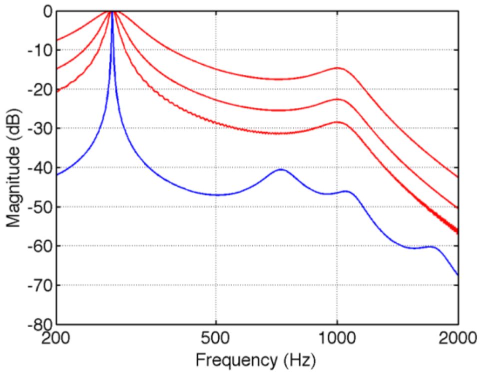

Figure 69. The red lines represent the low-order filters applied to sounds 4–6.

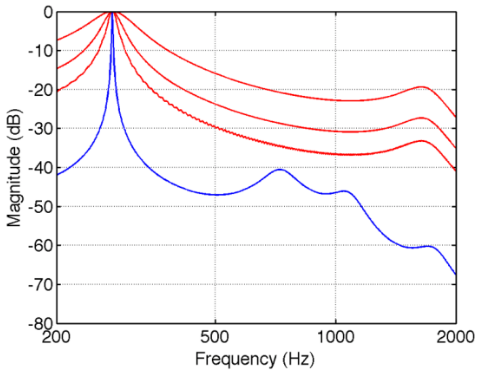

Figure 70. The red lines represent the low-order filters applied to sounds 7–9.

Table 15. The low-order filters applied to the artificial strike tone.

Sound No.	F1 (Hz)	Q	F2 (Hz)	Q
1	276	4	731	4
2	276	10	731	4
3	276	20	731	4
4	276	4	1032	4
5	276	10	1032	4
6	276	20	1032	4
7	276	4	1683	4
8	276	10	1683	4
9	276	20	1683	4

4.4.1.3 Procedure

Using the GUI shown in Figure 71, participants provided global dissimilarity ratings for the nine synthetic conga sounds presented pairwise. The experiment involved 72 pairwise presentations that resulted from the nine-by-nine matrix that excluded comparisons between identical stimuli. For each presented pair, participants were instructed to give a dissimilarity rating between 0 (identical) and 10 (totally dissimilar) inclusive, using an on-screen slider. Each sound could be replayed if required. The software program, developed in Max/MSP, stored each user's dissimilarity data as a text file once the experiment ended. The presented sounds were chosen randomly from the pool of 72 pairs. The software prevented participants from pressing the 'Store Answer / Next Pair' button until they had adjusted the slider. The sounds were reproduced monophonically over Sennheiser HD428 headphones.

Figure 71. The GUI used in the synthetic conga tones experiment.

4.4.2 RESULTS

The 72 dissimilarity-rating scores generated by the 17 participants were analysed using INDSCAL, which generated a two-dimensional spatial configuration, shown in Figure 72.

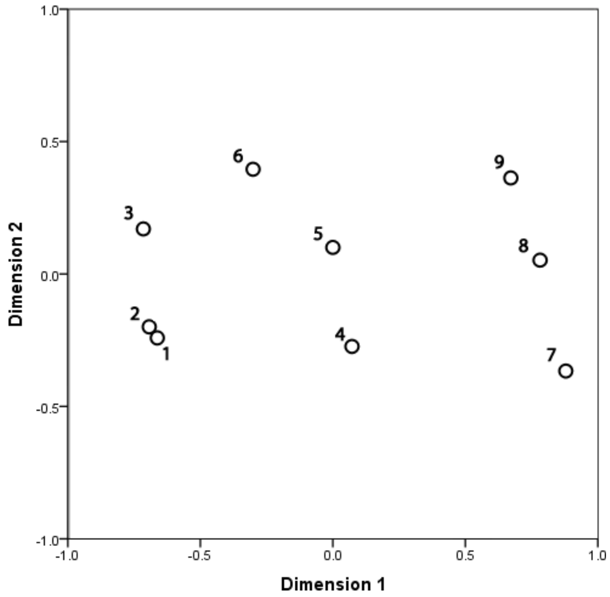

Figure 72. The two-dimensional spatial configuration resulting from the INDSCAL analysis of the dissimilarity ratings for the 17 participants

4.5 SUMMARY

The results of the Tic Tac® study suggest that the 18 participants were generally able to estimate the number of pellets being shaken inside a plastic container as the mean estimated number of pellets generally increased as the actual number of pellets increased. These estimates were achieved regardless of the introduced variation in levels and the size of the containers. Unlike the first Tic Tac® experiment in which mean estimates appeared to be impervious to reproduction level, the results of the revised Tic Tac® experiment suggest that mean estimates are dependent on spectrum. The results of the vowel categorisation study show that participants were sensitive to subtle changes in tone colour, as evidenced by their ability to categorise the presented sounds along a continuum. The results of the conga tone dissimilarity study show that students were able to discriminate between sounds that have varying F2 values, as evidenced by the spacing of the three triads (1-3, 4-6, 7-9) of sounds along the horizontal axis (Dimension 1) of Figure 72. Students were also able to discriminate between sounds with varying Q values, but to a lesser extent than the F2 values, as shown by the vertical grouping of sounds 1 and 2 in Figure 72.

The preliminary work detailed in this chapter investigated the efficacy of three tests of tone colour discrimination in isolation. The following chapter details the implementation of these three tests in conjunction with a newly-developed TETP in a pre/post-training test scenario.

5 PILOT STUDY

The previous chapter detailed the development and initial testing in isolation of three tests designed to be sensitive to students' tone colour discrimination abilities. This chapter details a pilot study in which the three tests – the Tic Tac® test, the conga tone dissimilarity test, and the vowel categorisation test – were implemented in a pre/post-training scenario to ascertain their sensitivity to changes in students' tone colour discrimination abilities resulting from undertaking a TETP.

Students were divided into two groups prior to the commencement of the experiment. Students in the first group trained using an ICA version of the TETP while students in the second group trained using an ISA version. The pre/post-training test results for the two groups were then compared to investigate the relative influence on training method on the students' ability to discriminate tone colour. The pilot study was carried out with Human Research Ethics Committee approval from the University of Sydney (Appendix A).

5.1 METHOD

5.1.1 PARTICIPANTS

The 52 students that participated in the experiment were undergraduate students enrolled in the Bachelor of Audio Production degree program at the SAE Institute in Sydney in 2012. The students voluntarily took part in the experiment each week as part of the tutorial component of the BAP120 Signal Processing and Aural Perception unit, delivered by the researcher in the second trimester of the first year of the degree program. The researcher recruited students at the commencement of the class, with the experiment taking place at the end of the semester. The class had 66 students, giving a participation rate of 79%. All participants were at least 18 years of age and none reported any hearing loss.

The experiment took place throughout the second semester over June–August 2012 in a computer lab at the SAE Institute. Students performed the TETP individually in their own time using their own computers, and completed the pre/post-training tests concurrently with other students in the class.

5.1.2 PROCEDURE

Prior to the commencement of the experiment, the 66 students in the class were randomly assigned to Group 1 (ICA) or Group 2 (ISA), with 33 participants assigned to each group respectively. Students used Sennheiser HD428 headphones to complete all pre/post-training tests. All students were instructed to undertake the three pre-training tasks by the end of the second week of the trimester as the TETP was scheduled to begin in the third week. The number of students that completed each pre-training test by the end of the second week is shown in Table 16.

Table 16. Number of participants per training group to complete each pre-training test.

Test 1	Test 2	Test 3	
		/u/ /i/	/ah/ /ae/
ICA=28 ISA=24	ICA=28 ISA=24	ICA=32 ISA=24	ICA=30 ISA=25

5.1.2.1 TIC TAC® TEST

The Tic Tac® test administered in this experiment was similar in method to the Revised Tic Tac® study detailed in Chapter 4. The test presented 40 stimuli, comprising five random presentations (without replacement) of the eight stimulus-filter combinations shown in Table 17. After first entering their student number, students were instructed to indicate how many pellets they thought were being shaken in the container by adjusting the on-screen slider of the GUI shown in Figure 73. The slider range was set to vary from two to 25 inclusive. The option to respond with a value of '1' using the slider was removed as the sound of five pellets (the smallest number of pellets) clearly did not sound like one pellet being shaken.

Table 17. The eight stimuli filter combinations used in the Tic Tac® test.

Scenario	No of Pellets	Container Size	Filter
1	5	Small	+12 dB High-shelf @ 1k Hz
2	5	Small	None
3	10	Small	+12 dB High-shelf @ 1k Hz
4	10	Small	None
5	15	Small	+12 dB High-shelf @ 1k Hz
6	15	Small	None
7	20	Small	+12 dB High-shelf @ 1k Hz
8	20	Small	None

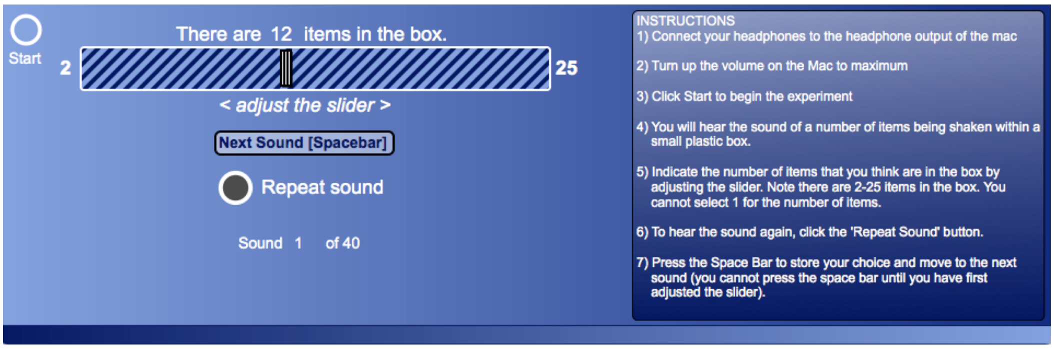

Figure 73. The Tic Tac® pre/post-training test GUI

5.1.2.2 Conga Tone Dissimilarity Test

The conga tone dissimilarity test administered in this experiment was identical to the 'synthetic conga tones' study previously detailed in Chapter 4, with the exception of the modified GUI shown in Figure 74.

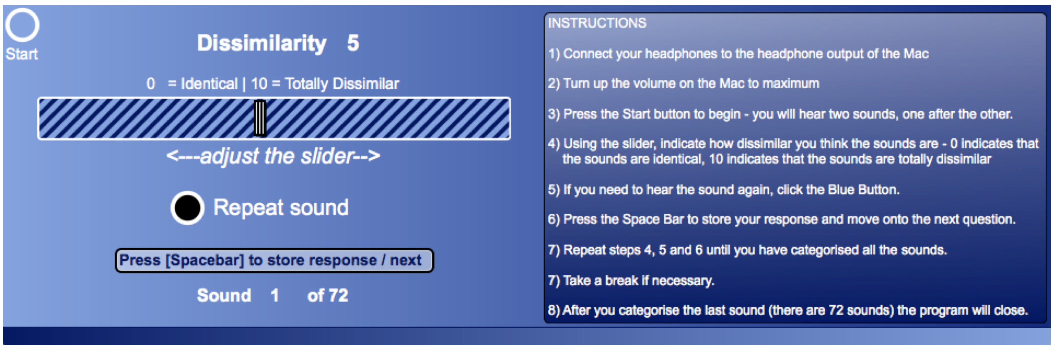

Figure 74. The conga tone dissimilarity pre-training test GUI

5.1.2.3 Vowel Categorisation Test

The vowel categorisation tests administered in this experiment were identical to the vowel categorisation study detailed in Chapter 4, with the exception of the orientation and test GUIs, shown in Figures 75–78.

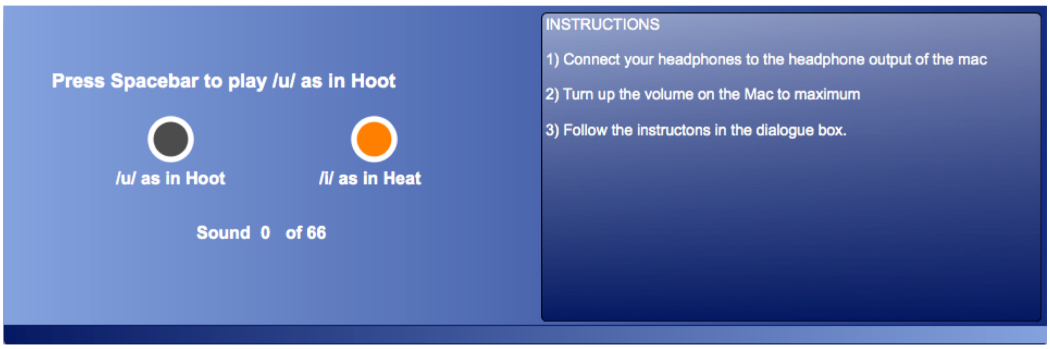

Figure 75. The /u/ versus /i/ vowel categorisation study pre-training orientation GUI

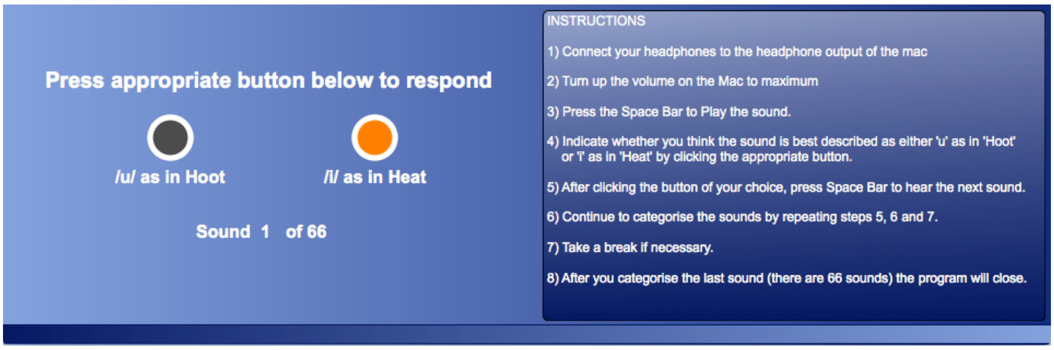

Figure 76. The /u/ versus /i/ vowel categorisation study Pre-training test GUI

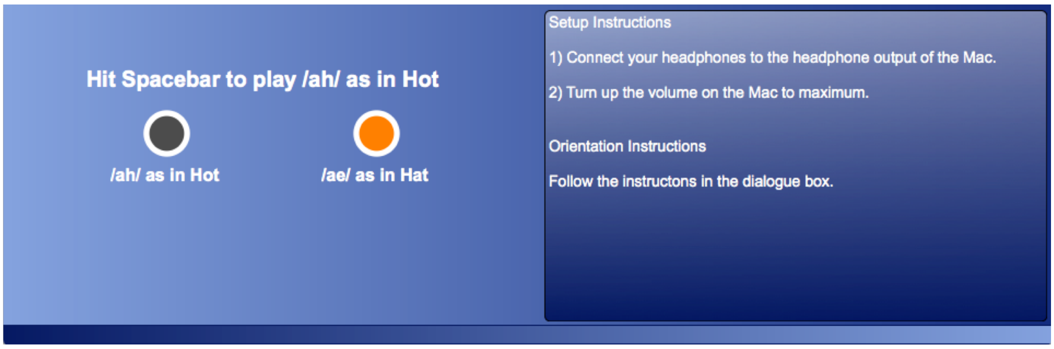

Figure 77. The /ah/ versus /ae/ vowel categorisation study Pre-training orientation GUI

Figure 78. The /ah/ versus /ae/ vowel categorisation study Pre-training test GUI

5.1.2.4 Technical Ear Training Program

After completing the three pre-training tests in the first two weeks of the trimester, students commenced a TETP entitled 'Frequency Identification Training'. The students were instructed to undertake the 63-question TETP a minimum of three times per week over six weeks. Students could train at home or at the SAE campus, where a computer lab was made available with the training program installed. Of the 44 students who completed the pre-training tests and commenced the TETP, 25 (57%) successfully completed the six-week training program. At the conclusion of each run through the training program, two text files were generated that stored the student's performance data. The students were instructed to submit these data files at the conclusion of the experiment as evidence of completing the training (although the TETP performance data was reserved for subsequent analysis and not examined in this study). In contrast to the pre/post-training tests, students were instructed that they could use their own headphones in the TETP, but they were asked to use the same headphones throughout the experiment.

Two distinct versions of a TETP were developed for the pilot study using Max/MSP software. The TETP administered in this study is described in further detail in Chapter 7. Students assigned to Group 1 were provided with an ICA version of the training program and students in Group 2 were provided with an ISA version. The students were not told that there were two different versions of the training program, and the GUIs for each group were identical. However, the students were given the following instructions:

> *You have been provided with a specific training program that may differ from the program your classmates have. It is vital that you DO NOT SWAP, COPY OR USE ANOTHER STUDENT'S PROGRAM. Use the program you were provided with.*

5.1.3 Post-Training Tests

In the week immediately following the final week of the TETP regime, students were instructed to repeat the three tests detailed above.

5.2 Results

For each of the three pre/post-training tests, in order for a student's performance data to be examined, the student needed to 1) complete the pre-training test, 2) complete the six-week TETP regime, and 3) complete the post-training test. As noted earlier, students' performances on the TETP, although recorded, were not the focus of the study. As such, TETP performance data is not presented here.

5.2.1 Tic Tac® Pre/Post-Training Test

The Tic Tac® test results for eight ICA and seven ISA students were analysed. The mean estimated number of pellets was plotted against the actual number of pellets for the pre-training and post-training tests for each training group (Figures 79–80). The pre-training results differed between groups, with the ICA group demonstrating a difference in pellet estimates between normal and high-frequency boost sounds pre-training. Post-training, both groups were able to largely ignore the high-frequency boost.

5.2.2 Conga Tone Dissimilarity Pre/Post Training Test

The conga tone dissimilarity test results for nine ICA and seven ISA students were analysed as detailed in Section 2 of Chapter 6, which describes the full-scale study. The INDSCAL-derived stimulus spaces for the pre-training and post-training tests for each training group are shown below (Figures 81–82).

ICA Pre-Training ISA Pre-Training

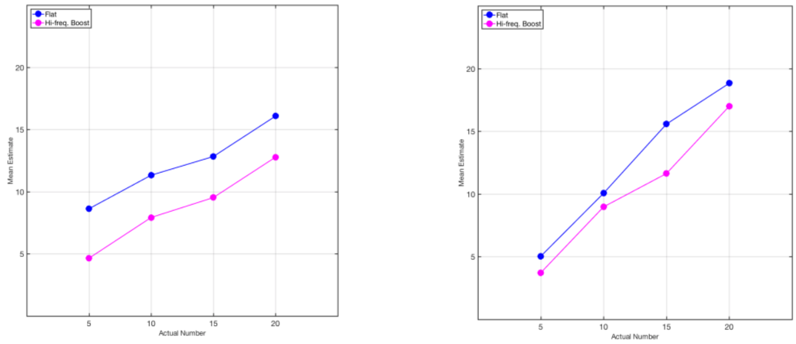

Figure 79. The pre-training mean estimated versus actual number of pellets for the ICA and ISA groups.

ICA Post-Training ISA Post-Training

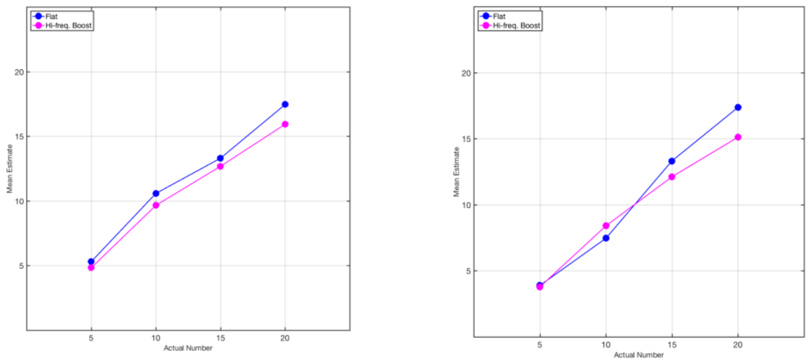

Figure 80. The pre-training mean estimated versus actual number of pellets for the ICA and ISA groups.

For both groups, the first dimension of the Stimulus Spaces derived from the dissimilarity ratings was highly correlated with the frequency of the second formant used in synthesising the nine conga timbres (r = 0.9723 for ICA; r = 0.9633 for ISA). However, the second dimension coordinates of the Stimulus Spaces derived separately for each of the two groups (receiving either ICA or ISA training) were not so similar between groups. Whereas the correlation between coordinates on dimension two of the ISA Stimulus Space and the Q value of the first formant was relatively high (r = 0.7775), this relationship was not strong in the ICA Stimulus Space (r = 0. 3482).

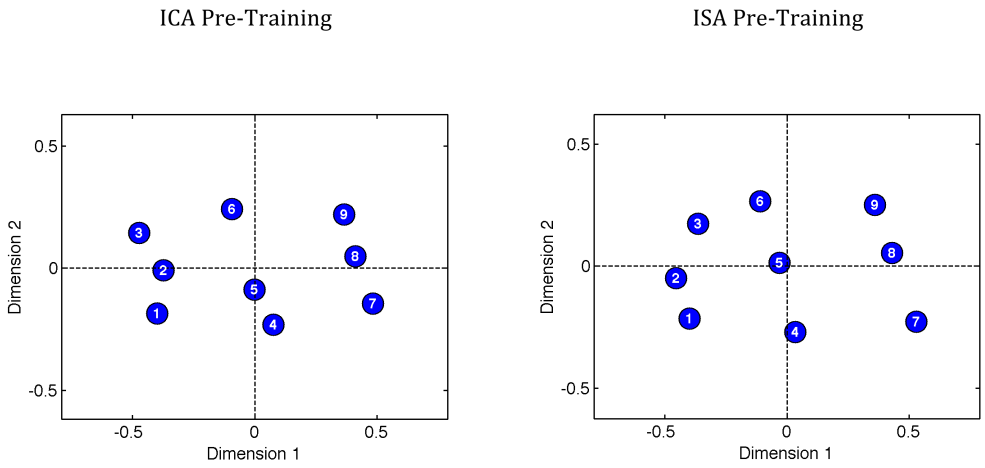

Figure 81. The pre-training INDSCAL-derived stimulus spaces for the ICA and ISA groups.

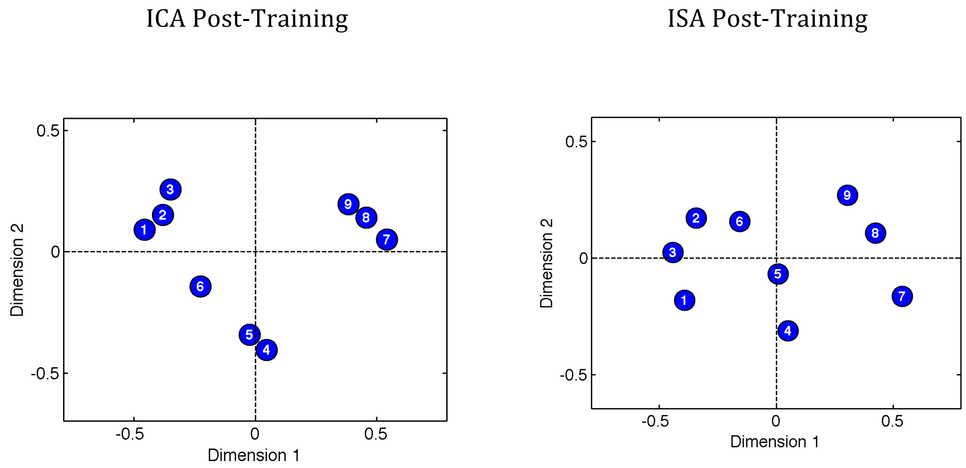

Figure 82. The post-training INDSCAL-derived stimulus spaces for the ICA and ISA groups.

5.2.3 Vowel Categorisation Test

Only four participants completed each of the vowel categorisation tests. In addition, the number of students that completed the '/u/ versus /i/' pre/post-training tests is different to the number that completed the '/ah/ versus /ae/' test. Therefore, the data resulting from the vowel categorisation tests is not presented due to the low number of participants.

5.3 SUMMARY

Following the appraisal of the three tests of tone colour discrimination in isolation, the pilot study deployed the tests in a pre/post-training scenario. The Tic Tac® test results yielded similar post-training results for both training groups. The results of the vowel categorisation tests were unable to be analysed due to the small sample size. The results of the conga tone dissimilarity test, however, showed a difference in the Stimulus Space between the ICA and ISA training groups. Based on this result, the conga tone dissimilarity test was chosen for deployment within the full-scale study detailed in the following chapter.

6 FULL-SCALE STUDY

The previous chapter detailed the design and results of the pilot study conducted to determine the suitability of the three proposed pre/post-training tests as measures of students' tone colour discrimination ability. Based on the results of the pilot study, the conga tone dissimilarity test was chosen as the most suitable for showing a change in participants' ability to discriminate tone colour when deployed in a pre/post-training methodology. This section details the full-scale implementation of the pre/post-training test methodology featuring the conga tone dissimilarity test. The full-scale study was carried out with Human Research Ethics Committee approval from the University of Sydney (Appendix A).

6.1 METHOD

6.1.1 PARTICIPANTS

The participants in the full-scale study were first-year undergraduate students enrolled in the Bachelor of Audio Production degree program at SAE Institute campuses in Australia (Sydney, Melbourne, Byron Bay, Brisbane) and the United Arab Emirates (Dubai). All participants were at least 18 years of age at the time of the experiment. Some students reported hearing loss in the user questionnaire, as detailed below. The pre/post-training tests and TETP were included as a single assessment item in the BAP120 Signal Processing and Aural Perception unit delivered in the second trimester of the first year of the degree. Completion of the tests and the TETP was mandatory for students if they wanted to receive a passing grade for the assessment. However, participation in the research was voluntary; students were given the opportunity to opt out of the study at the sign-up stage of the training program. If a student chose to opt out of the study, data relating to their performance on the TETP and the pre/post-training tests was not recorded.

6.1.2 Procedure

6.1.2.1 Conga Tone Dissimilarity Test

As in the pilot study described in the previous chapter, the conga tone dissimilarity test featured a pre/post-training test methodology whereby students undertook a series of tone colour discrimination tasks before and immediately after a TETP. This study involved a similar procedure in that a conga tone dissimilarity test was administered prior to the commencement of the TETP. However, the test was then administered at various regular intervals throughout the TETP, as opposed to adhering to a strictly pre/post-training test methodology. In addition to the pre-training test, the conga tone dissimilarity test was administered to students after they completed each group of five TETP training sessions. The discrimination task employed in this study was identical to that administered in the pilot study (chapter 6), except for cosmetic changes to the GUI.

The Unit Outline for the BAP120 Signal Processing and Aural Perception unit instructed students to complete three TETP sessions per week over 10 weeks. In order to begin the first TETP session, students had to complete the pre-training test, in addition to the orientation and warm-up sessions.

Five hundred and ten participants signed up for the training program. Although not all registered participants were enrolled at SAE, only the data from participants that were also enrolled as SAE students in the BAP120 Signal Processing and Aural Perception unit were included in this study. The total number of participants from SAE was 223, with 97 assigned to the ICA group and 126 assigned to the ISA group. The difference in group numbers was due to the test software alternatively assigning users to each training group; it assigned 50% of users to each group, but after non-SAE participants were removed prior to data analysis by the researcher, the percentage of participants in each group changed. Only 140 of the 223 (63%) participants who agreed to participate in the research actually commenced the TETP (completed the pre-training test), and of those, only 45 (32%) completed all five tests. As the students could complete the tests and TETP anywhere they had an internet connection, they were permitted to use their own headphones, but were instructed to use the same headphones throughout the study.

As part of the TETP sign-up process, students completed a user questionnaire (Figure 83). A summary of the data from the questionnaire is shown in Figures 84–88. The relationship between users' responses to these questions and their performance on the pre/post-training test and TETP was beyond the scope of this study and was not assessed, as the focus of this study was on students' performance on the pre/post-training tests and not on the TETP.

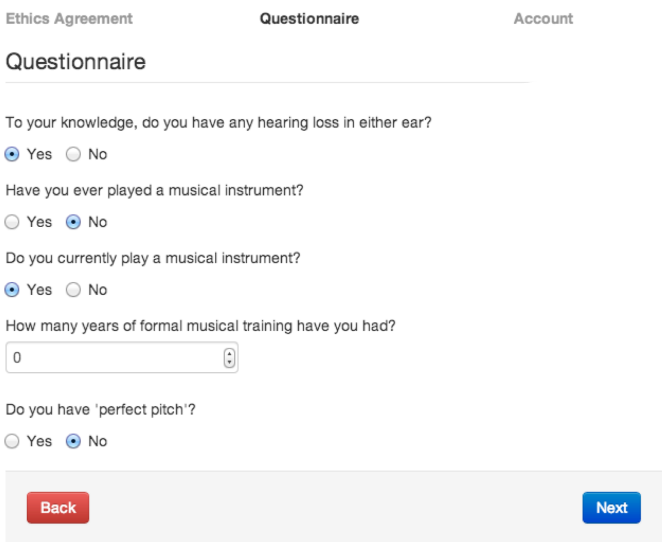

Figure 83. The user questionnaire.

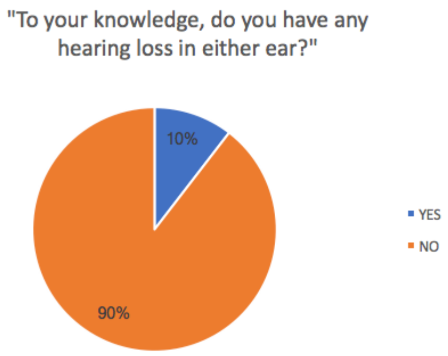

Figure 84. User questionnaire response percentages for the hearing loss question.

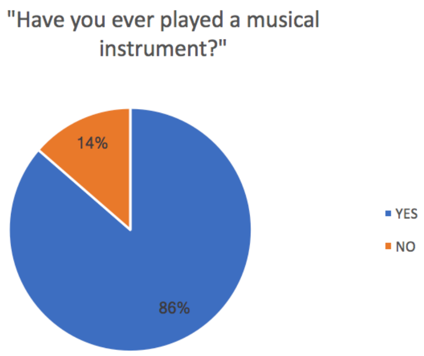

Figure 85. User questionnaire response percentages for the musical instrument question.

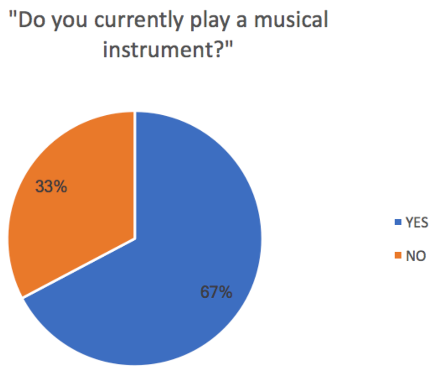

Figure 86. User questionnaire response percentages for the current musical instrument question.

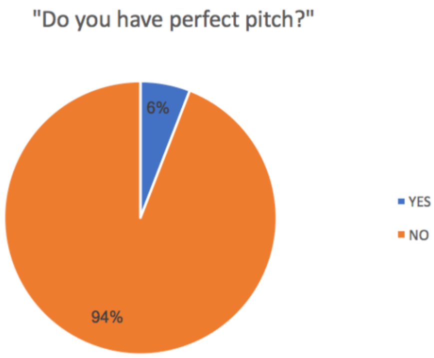

Figure 87. User questionnaire response percentages for the perfect pitch question.

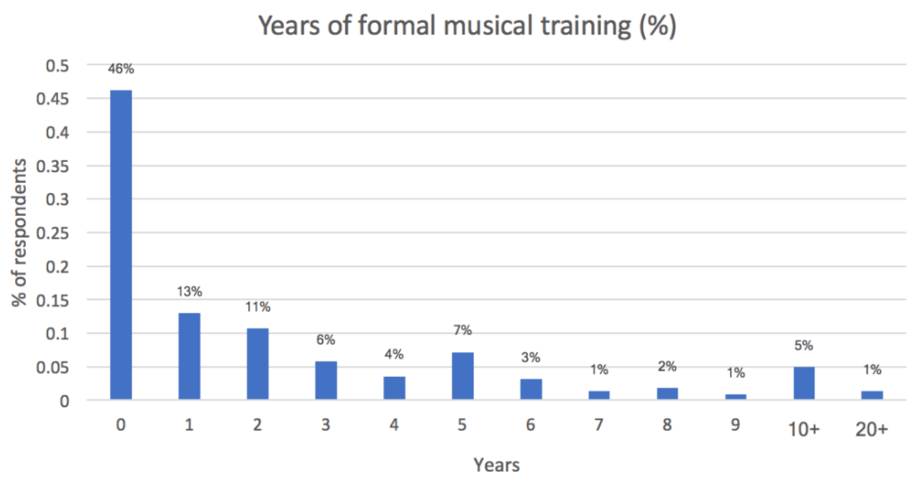

Figure 88. User questionnaire response percentages for years of formal musical training question.

6.1.2.2 Technical Ear Training Program – 'Tonal Ear Training'

As in the pilot study (chapter 5), two versions of the TETP were developed for this experiment. Students assigned to Group 1 were trained with an ICA version of the program, while students in Group 2 trained with an ISA version. Students were not told that there were two different versions of the training program, and the GUIs for each group were identical. The TETP administered in this study is presented in detail in Chapter 7.

6.2 Results

The results of five iterations of the conga tone dissimilarity test were examined separately for each iteration and training group by submitting the dissimilarity-rating scores to a multi-dimensional scaling (MDS) analysis. The INDSCAL-derived stimulus spaces for each group and test iteration were then examined for differences in discrimination ability between training groups and test iterations. The number of participants for each training group and test iteration prior to data filtering is shown in Table 18.

Table 18. The number of participants in each training group prior to data filtering.

	ICA	ISA
Pre-training test	62	78
Post-training test 1	39	54
Post-training test 2	28	43
Post-training test 3	23	37
Post-training test 4	17	28
Totals	169	240

Prior to the MDS analysis, the data was examined for the presence of outlier responses that indicated that a user did not legitimately attempt the test. A low standard deviation, for example, suggests that the user responded with a relatively small number of the 11 available response options (see Figure 90 for an example). For each training group and test iteration, the standard deviation of each user's 72 dissimilarity-rating responses was first calculated. The first (Q_1) and third (Q_3) quartiles were then calculated to ascertain the interquartile range (IQR). The data for users with a standard deviation score outside of +/-1.5 IQR was removed. To ensure the same users were in each training iteration group, if a user did not meet the outlier conditions for any test iteration, their data was deleted from all test iterations, although some dropped out (and therefore completed fewer tests) over time. The resultant number of participants for each training group and test iteration post-outlier filtering is shown in Table 19.

Table 19. The number of participants in each training group after outliers were removed.

	ICA	**ISA**
Pre-training test	55	75
Post-training test 1	34	50
Post-training test 2	24	39
Post-training test 3	20	35
Post-training test 4	14	27
Totals	147	226

Following the removal of outliers using the IQR, the data was visually examined by plotting the nine-by-nine dissimilarity rating matrix for each for each group and test iteration and conditionally formatting the data using a heat map-style graded colour scale (Figure 89). Identical pairs of sounds were not presented in the test, so the top-left to bottom-right diagonal (the dissimilarity ratings for these identical pairs) was populated with '0's when the data was stored after each test. Note the general tendency for higher scores in the bottom-left and top-right corners of the matrix.

Examining the data for all users in this manner, it became evident that the outlier filtering had not removed data for all users who did not legitimately attempt the test. Figure 90 shows the nine-by-nine dissimilarity rating matrix for one such user. It is clear that this user did not attempt the test in good faith, as he or she responded with a dissimilarity rating of '5' for every pair of sounds.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	**0**	0	3	3	5	8	10	8	8
2	0	**0**	0	1	3	0	10	7	7
3	1	0	**0**	2	1	0	3	1	7
4	4	2	2	**0**	0	1	2	6	8
5	3	1	6	3	**0**	0	10	3	5
6	5	4	4	4	0	**0**	7	7	9
7	6	7	2	6	1	8	**0**	0	2
8	9	3	8	7	7	7	0	**0**	5
9	8	5	5	8	8	1	1	0	**0**

0	1	2	3	4	5	6	7	8	9	10

Figure 89. The nine-by-nine dissimilarity-rating matrix for a user, conditionally formatted using a heat map-style graded colour scale.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	0	5	5	5	5	5	5	5	5
2	5	0	5	5	5	5	5	5	5
3	5	5	0	5	5	5	5	5	5
4	5	5	5	0	5	5	5	5	5
5	5	5	5	5	0	5	5	5	5
6	5	5	5	5	5	0	5	5	5
7	5	5	5	5	5	5	0	5	5
8	5	5	5	5	5	5	5	0	5
9	5	5	5	5	5	5	5	5	0

Figure 90. The nine-by-nine dissimilarity-rating matrix for a user who did not attempt the test in good faith.

A rule was devised to highlight such cases so that the data of users who did not attempt the test in good faith was removed from the analysis. The number of times each response (from '0' through to '10') was submitted by the user for each row and column on the matrix was counted. For each user's row of the matrix in each test, any user who responded with the same response five or more times (out of nine) was deemed a nuisance user and had their entire dataset removed from all test iterations. The same rule was applied for each user's column in each test's matrix.

It also become apparent that several users did not use all possible responses in the entire dissimilarity-rating scale provided. The user's responses shown in Figure 91 met the requirements detailed in the previous paragraph, but in 72 responses, this user never responded with a '2', '5', '6' or '9' rating. This response pattern suggests that the user was consciously selecting only 1, 3, 4, 7, 8, 10 and not responding with a rating in the centre of the scale. Another example of this type of response pattern is shown in Figure 92. This user also met the IQR requirements and those detailed above, but did not use the full range of the scale, never responding with a '7', '8', '9' or '10' rating. Participants who used 7 or fewer of the 11 possible rating responses in their set of 72 dissimilarity ratings per test iteration also had their dataset removed from all test iterations.

Following data filtering, the number of users in the ICA group for the fourth post-training session was considered too small to be useful. As such, the pre-training test results were compared to the third set of post-training test results.

	1	2	3	4	5	6	7	8	9
1	0	3	3	10	10	10	10	3	3
2	7	0	10	8	3	8	8	3	10
3	3	4	0	8	10	1	3	1	8
4	3	3	1	0	8	7	3	10	8
5	10	7	4	1	0	10	10	1	10
6	7	8	3	3	1	0	7	3	1
7	10	3	10	8	3	8	0	1	10
8	7	8	3	8	3	4	1	0	8
9	10	1	1	3	8	8	1	1	0

Figure 91. The nine-by-nine dissimilarity-rating matrix for a user who did not respond with '2', '5', '6' or '9'.

	1	2	3	4	5	6	7	8	9
1	0	1	1	1	1	5	5	6	6
2	1	0	0	2	1	1	5	6	4
3	1	0	0	1	1	1	6	4	5
4	1	2	2	0	1	3	6	4	4
5	2	1	3	1	0	1	5	3	3
6	2	2	3	1	1	0	6	4	4
7	2	4	5	3	4	4	0	1	1
8	5	3	4	1	3	2	3	0	0
9	5	3	4	4	3	2	1	0	0

Figure 92. The 9x9 dissimilarity-rating matrix for a user who did not respond with '7', '8', '9' or '10'.

At this stage, a final round of data filtering took place. Only the pre-training dissimilarity rating data for users who completed at least the first, second and third post-training tests was included in the final analysis. This ensured that for each of the ICA and ISA training groups, the dissimilarity rating data for the same number of users and the same set of users was compared pre- and post-training. The number of participants in each training group and test iteration resulting from the data filtering detailed in this section is shown in Table 20.

Table 20. The number of participants in each training group after all data filtering.

	ICA	ISA
Pre-training test	18	30
Post-training test 3	18	30

The INDSCAL-derived stimulus spaces for each group and test iteration were then examined for the pre-training and post-training tests. The pre-training test and third post-training test were separated by 15 TET sessions and approximately five weeks. The performance of both groups (ICA and ISA) on the pre-training test is shown in Figure 93. The INDSCAL-derived stimulus spaces for both groups were similar prior to undertaking the TETP.

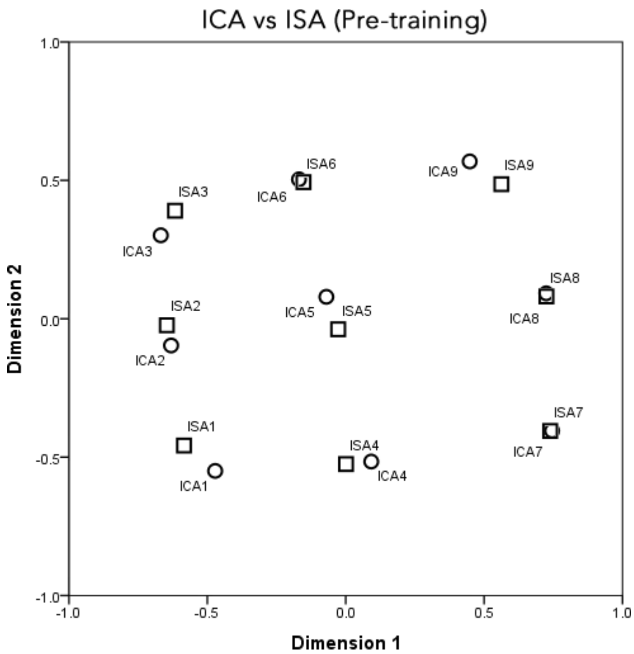

Figure 93. The pre-training INDSCAL-derived stimulus spaces for ICA (circles) and ISA (squares) groups.

The pre/post-training MDS analysis results for both groups are shown in Figure 94. The post-training results for the ISA group show that the users' performance on the task was almost identical to their pre-training performance. The post-training results for the ICA group show that performance for the 4th, 7th, 8th and 9th sounds was also almost identical to pre-training performance, with minor differences in the MDS analysis results for the remainder of the sounds observed. For the sounds that did exhibit a pre/post-training difference, the magnitude of the change was minor, and the order of the sounds along both dimensions did not change. As such, the overall difference between the ICA group's pre-training performance and their post-training performance is not considered significant.

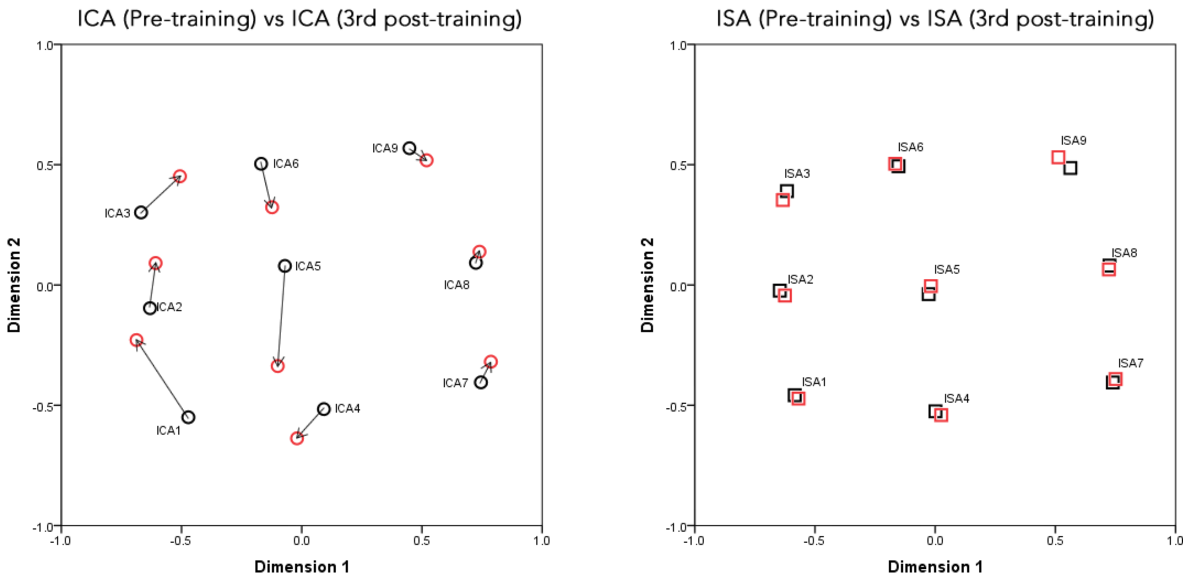

Figure 94. The INDSCAL-derived stimulus space for the ICA (left - circles) and ISA (right - squares) training groups pre-training (black) and post-training test (red).

6.3 Summary

Following the pilot-study evaluation of the three tests of tone colour discrimination in a pre/post-training scenario, the conga tone dissimilarity test was deployed in a full-scale study. In contrast to the pilot study, the full-scale study results for the conga tone dissimilarity test did not reveal a meaningful difference in the Stimulus Space between the ICA and ISA training groups.

The following chapter details the TETP developed for the pilot study and the full-scale study, in addition to a newly developed program titled the SAE Parametric Equaliser Training application.

7 New Technical Ear Training Programs

Deployment of the pilot and full-scale studies each required the development of two separate elements. Firstly, suitable tests of tone colour discrimination were developed and deployed in isolation, separately from any TET program. Following this early work, the tests were deployed in a pre/post-training scenario in a pilot study. Based on the results of the pilot study, one of the tests was selected for use in the full-scale study, where it was again deployed in a pre/post-training arrangement.

In addition to the pre/post-training tests, each of these studies required the development of a TETP. For the pilot study and the full-scale study, custom TETPs were developed that facilitated training students with either the ICA or ISA training methods.

Following the full-scale study, the SAE Parametric Equaliser Training application was developed. This TETP differs from earlier programs as it was not deployed as part of a pre/post-training test methodology and was not designed to gather user performance data. In addition, this TETP is not limited to students; it was developed to provide a free, publicly available training program that is easy to use and accessible to anyone with an Apple Mac computer.

This chapter provides an overview of the TETPs developed for the training stages of the pre/post-training scenarios employed in the pilot and the full-scale studies. In addition, the chapter provides an overview of the SAE Parametric Equaliser Training application.

7.1 SAE Frequency Identification Training

This section details the TETP developed for the training component of the pilot study conducted at the SAE Institute in Sydney in 2012, herein referred to as the SAE TETP.

7.1.1 Design Considerations

The SAE TETP had simple design requirements relative to the other TETPs detailed later in this chapter. The program needed to facilitate a matching-type task by applying a parametric equaliser with fixed gain and Q to monophonic pink noise. The orientation session needed to allow students to return to the bypass and filtered signals, but the training session was to automatically step the user through these sounds and prevent the user from auditioning them more than once per question.

Both ICA and ISA versions of the TETP needed to be developed. The ICA version was to allow the user to hear the centre frequency of the parametric filter 'sweep' throughout the spectrum as they adjusted the on-screen slider, and the ISA version was to mute the audio output, preventing the user from 'sweeping' as they adjusted the on-screen slider.

7.1.2 Platform / Operating Environment

As the researcher had previous experience in programming in the Max environment, the TETP was developed in Max. The program was developed in the Mac OSX operating environment due to the researcher's familiarity with it and the fact that the SAE Institute used iMac computers in the computer labs where students would undertake the training.

7.1.3 WALKTHROUGH

Each time students loaded the TETP they were required to complete a 10-question orientation session (Figure 95). The orientation session was embedded into each version of the TETP; Group 1 students undertook an ICA version of the orientation session and Group 2 undertook an ISA version. The orientation task was a matching task as it required students to adjust the centre frequency of a filter applied to the 'Response' so that it matched the 'Filtered' signal.

After entering their student number, students were presented with monophonic unfiltered continuous pink noise, labelled as the 'Bypass' signal. Students were then instructed via the GUI to click on the 'Filtered' button, which applied a parametric filter (+18.06dB, Q=4.0) to the noise signal at a centre frequency randomly chosen from one of 22 1/3rd octave frequencies between 100 Hz and 12k Hz inclusive. The GUI then instructed students to click on the 'Response' button, which initially played unprocessed monophonic pink noise. As soon as students adjusted the on-screen slider, a filter with the chosen centre frequency was applied. The on-screen slider could be set to one of 85 1/12th octave frequencies between 100 Hz and 12.8k Hz inclusive. This set of frequencies contained all 22 1/3rd octave frequencies that could be chosen from the question pool. In the orientation sessions, students could go back and listen to the 'Bypass' or 'Filtered' signal at any time by clicking on the corresponding button on the GUI. Once the student had decided on the centre frequency, they pressed the spacebar to submit their answer. If they were correct, the GUI displayed 'Correct'. If they were not correct, the GUI displayed 'Incorrect', then 'the correct answer was [answer] Hz'.

Students using the ICA (Group 1) version of the orientation session received continuous aural feedback when they adjusted the on-screen slider, allowing them to hear the filter ‘sweep’ throughout the spectrum. The orientation session provided to the ISA (Group 2) students, however, muted the audio output whenever the slider was adjusted; only once the user ceased adjusting the slider did the audio output unmute. As such, the duration of signal muting was dependent on the user’s adjustment of the slider. The program did not save orientation performance data.

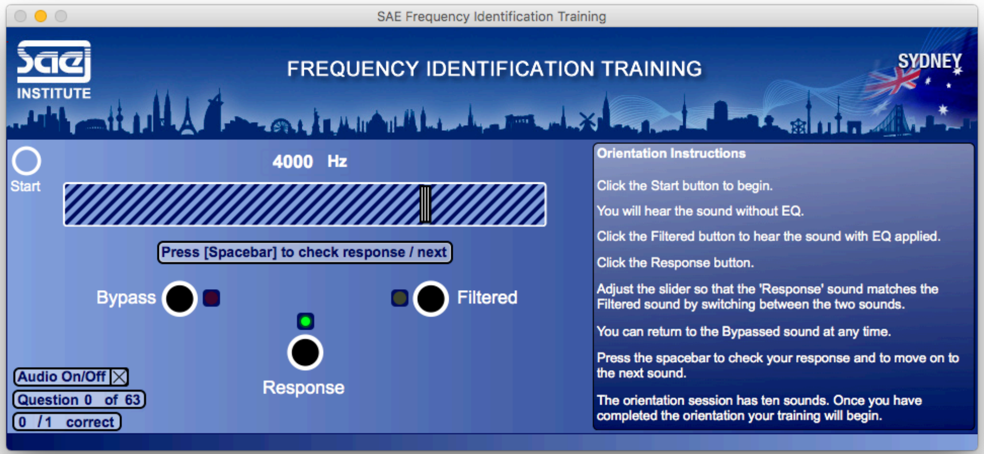

Figure 95. The orientation screen of the SAE TETP (identical for ICA and ISA training groups).

Once the 10 orientation questions were completed, the GUI changed to the training screen (Figure 96). Tasks in the training session were similar to those in the orientation session in that students were required to perform matching tasks. However, in the training session, all buttons were removed and the student was automatically presented with the ‘Bypass’, ‘Filtered’ and ‘Response’ sounds. For each question, the ‘Bypass’ signal (identical to that used in the orientation session) was presented automatically by the program. After three seconds, the program switched to the ‘Filtered’ signal. After another three seconds, the program switched to the ‘Response’ signal, which was initially set to unfiltered pink noise. The TETP contained 63 questions relating to three sets of 21 randomly selected centre frequencies chosen from the 22

1/3rd octave frequency question pool detailed above for the orientation session. The question pool, filter type and slider parameters used in the training session were identical to those used in the orientation session. For each question, students did not have the option of returning to the 'Bypass' or 'Filtered' signals once presented, as these buttons were removed from the training session interface. The ICA and ISA features used in the orientation session were implemented in an identical manner in the training session. At the conclusion of each TETP training session, two text files with the student's performance data were generated and saved. Students submitted these text files at the conclusion of the training regime.

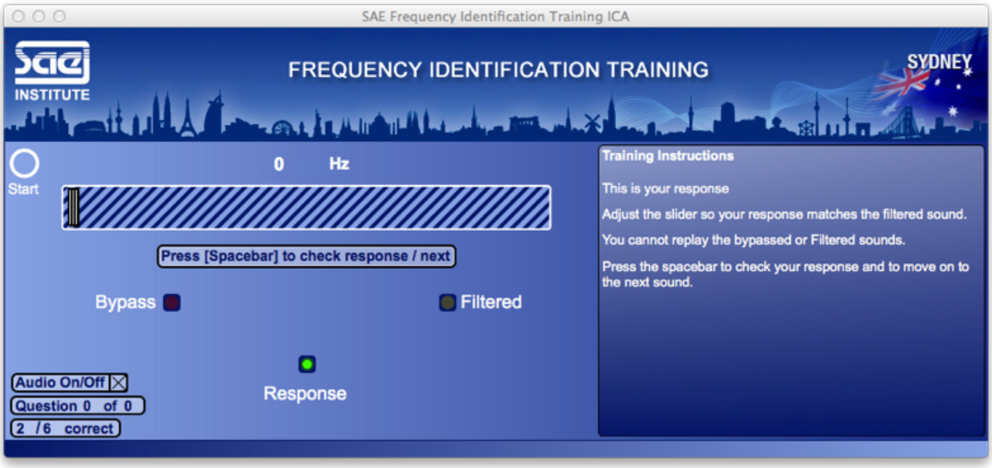

Figure 96. The training screen of the SAE TETP – the GUIs for both training groups were identical.

7.1.4 User Performance Indicators

The GUI provided minimal feedback to students regarding their performance on the Orientation and Training phases of the TETP. The current question number, total number of questions, number of attempted questions and the number of questions answered correctly were displayed.

7.1.5 DEMONSTRATION

To view a short video demonstration of the ICA and ISA versions of the TETP used in the pilot study, go to http://audio.education/phd.html and enter the following credentials:

username: phd

password: MarkBassett

This password gives access to all videos on the website.

7.2 TONAL EAR TRAINING

This section details the Tonal Ear Training TETP that was developed for the training component of the full-scale study, which implemented a pre/post-training methodology. Unlike the other TETPs detailed in this chapter, the Tonal Ear Training TETP was delivered online via the user's web browser. The tone colour discrimination pre/post-training tasks embedded within the Tonal Ear Training TETP are detailed in Chapter 6 and are not addressed here. The Tonal Ear Training TETP utilises a matching task as in the pilot study.

Tonal Ear Training was designed in collaboration with David Major (then an audio lecturer at the SAE Institute Dubai). Mr Major provided advice on the implementation of features of the GUI and overall guidance on the implementation of a TETP running within a web browser. Mr Major was solely responsible for the programming (coding) of the training program using a combination of Ruby on Rails, HTML5 and JavaScript.

7.2.1 Design Considerations

There were two choices with respect to how to approach the storage and presentation of audio material within the Tonal Ear Training program: real-time processing and off-line processing. Off-line processing involves rendering all possible stimulus and filter combinations as audio files ahead of time and storing them on the server. The user's web browser then downloads and plays these individual files as the program requires. Because this method requires the user browser to download 1323 separate .wav files, (21 stimuli multiplied by 63 filters = 4.08GB), interruptions in playback could occur (depending on the user's internet connection speed). Therefore real-time processing was chosen, which involved downloading only 21 stimuli to the user's browser and performing all signal processing natively within the browser.

7.2.2 Platform / Operating Environment

To ensure consistency in user experience, users were required to engage with the TETP using Google's Chrome web browser only. As the TETP was deployed within a browser, users could access the training program on either a Windows or Mac computer.

7.2.3 Walkthrough

Upon visiting the Tonal Ear Training website for the first time, users were required to open an account. After completing an Ethics agreement, users completed a questionnaire (Figure 97), then chose an account username and password.

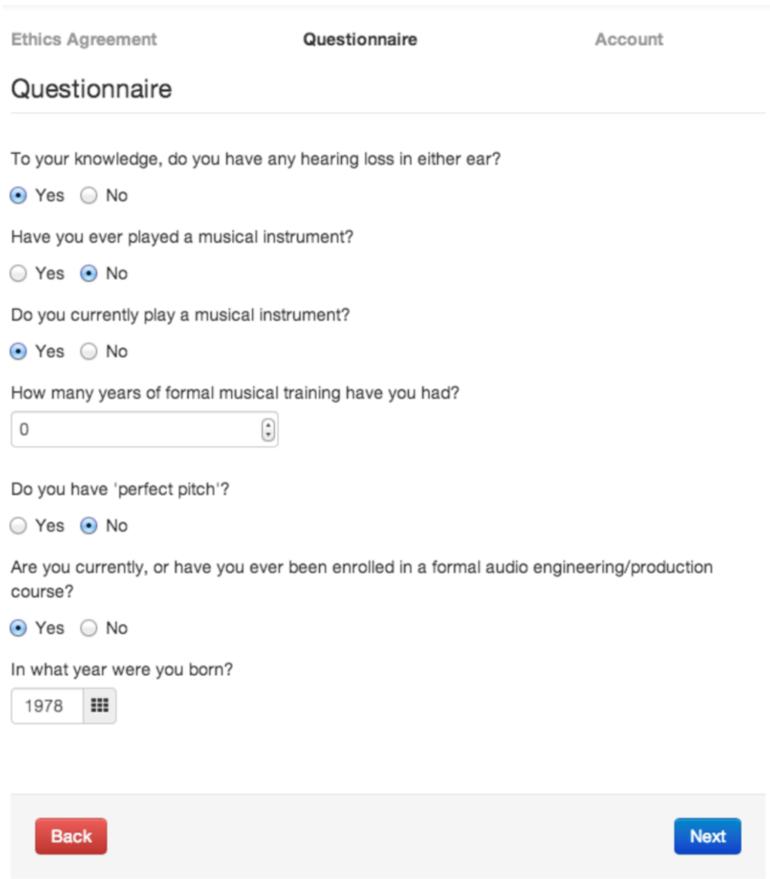

Figure 97. The user questionnaire page.

After completing the sign-up process, users were taken to the Home page (Figure 98). Returning users were taken to the Home page directly after signing in. The Home page displays the user's progress through the Practice (Orientation and Warm-up) and Training (Training and Testing) components of the TETP.

During the sign-up process, users were alternatively assigned to be in either the ICA or ISA training group. The Practice and Training components of the TETP were programmed to deliver either an ICA or ISA version of the training interface. Users were not told that there were two different groups or the group to which they were assigned.

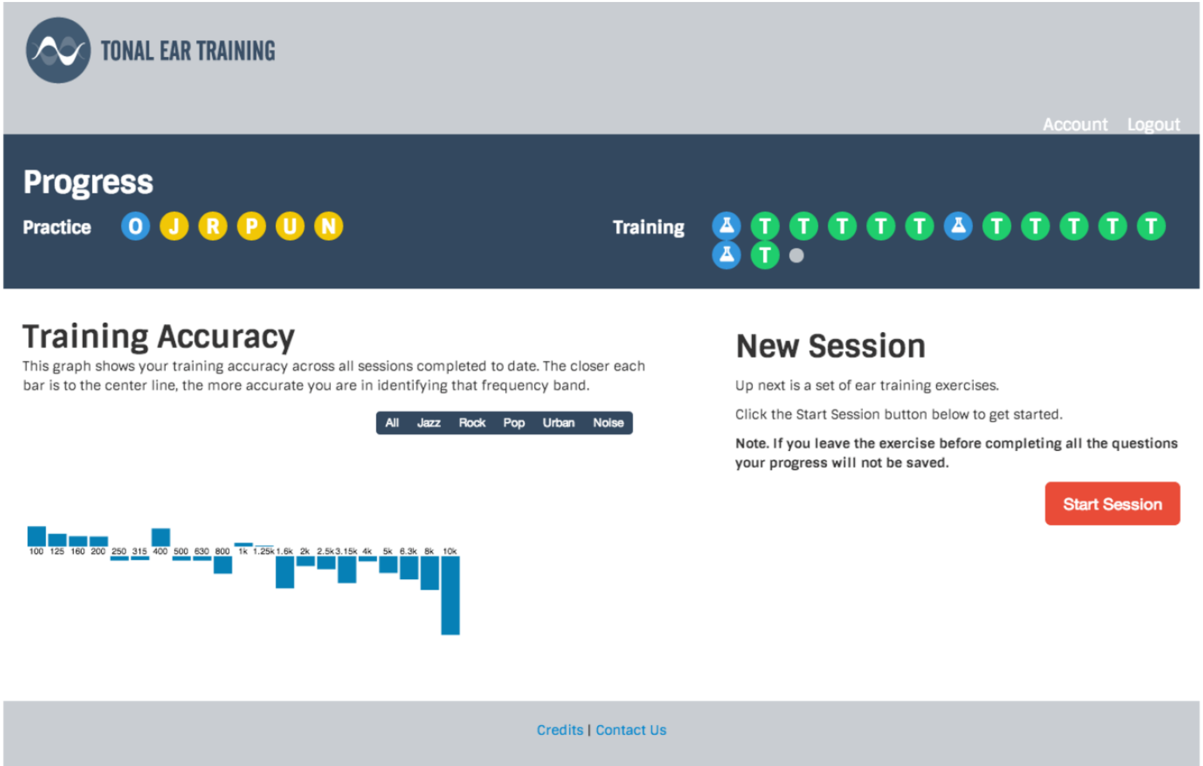

Figure 98. The home page of the Tonal Ear Training Program.

Before undertaking the orientation and training sessions, students were presented with a headphone calibration page (Figure 99). On this page students were asked to indicate the model of headphones they were using. They were also presented with filtered pink noise, which is indicative of the loudest level signal that they would experience in the training program, and advised to set the playback level of the noise to a comfortable level.

New users were first required to complete an orientation session that walked them through the training interface and explained the features. After completing the orientation session, the user was again returned to the home page. Before commencing the training regime, the user had to complete five warm-up sessions, one for each genre (Jazz 'J', Rock 'R', Pop 'P', Urban 'U', and Pink Noise 'N'). The warm-up sessions each consisted of 21 questions. For each musical genre, five 15–20-second music tracks

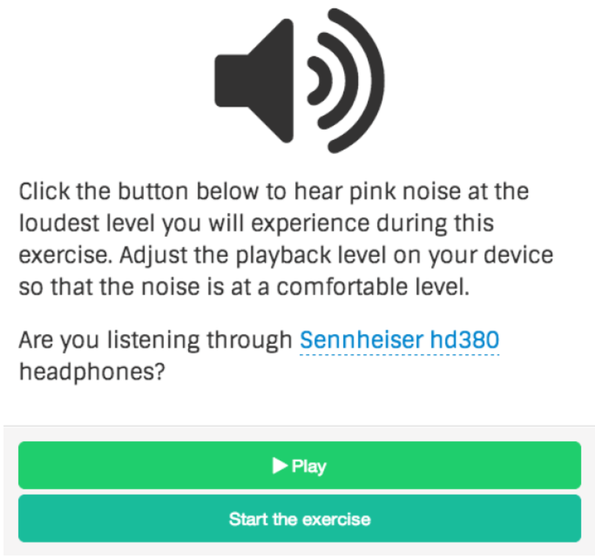

Figure 99. The headphone calibration page.

were selected as stimuli. For the Pink Noise sessions, a single pink noise sound file was used. All stimuli were presented in mono. For the musical genres, the stimulus was randomly chosen (with replacement) from the five music tracks for that genre for each presented question. First, the user was presented with the looped stimulus signal without signal processing (Figure 100). Three seconds later, the program automatically switched to the 'Filtered' version of the signal (Figure 101). The filter applied was a parametric equaliser (+6 dB, Q=4) with a centre frequency randomly chosen (without replacement) from one of 21 1/3rd-octave band frequencies between 100 Hz and 10k Hz inclusive. After another three seconds, the program switched to the 'Response' signal (Figure 102), which initially presented an unprocessed signal. Once the user adjusted the on-screen frequency slider, a parametric equaliser was applied (+6 dB, Q=4) with a centre frequency equal to that selected by the user. The frequency slider allowed the user to select one of 21 1/3rd octave band frequencies between 100 Hz and 10k Hz inclusive. For each question, the user could not return to the 'Bypassed' or 'Filtered' signals once presented. The user had an unlimited amount of time to select their answer and click 'Submit'.

Students assigned to the ICA group could hear the centre frequency of the filter 'sweep' throughout the spectrum as they adjusted the on-screen frequency selector. When students in the ISA group adjusted the frequency selector, stimulus playback was muted until one second after the slider was released.

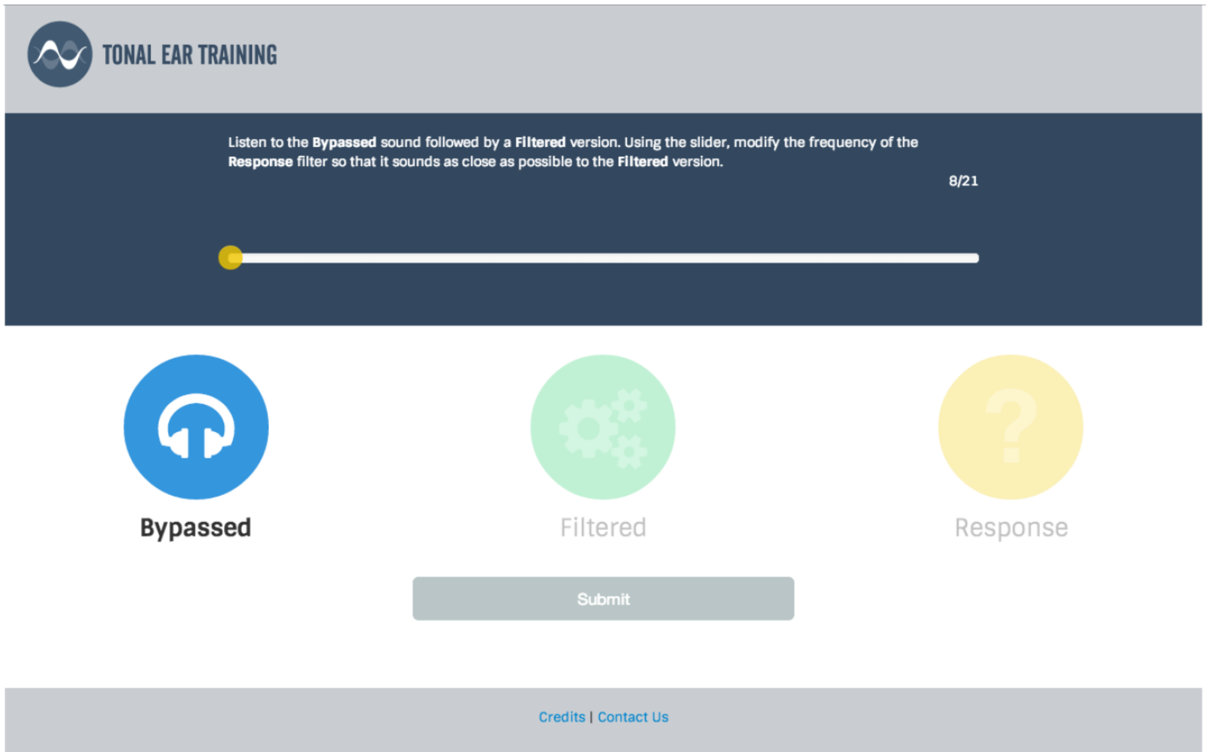

Figure 100. The warm-up / training page with the Bypassed (unprocessed) stimulus presented.

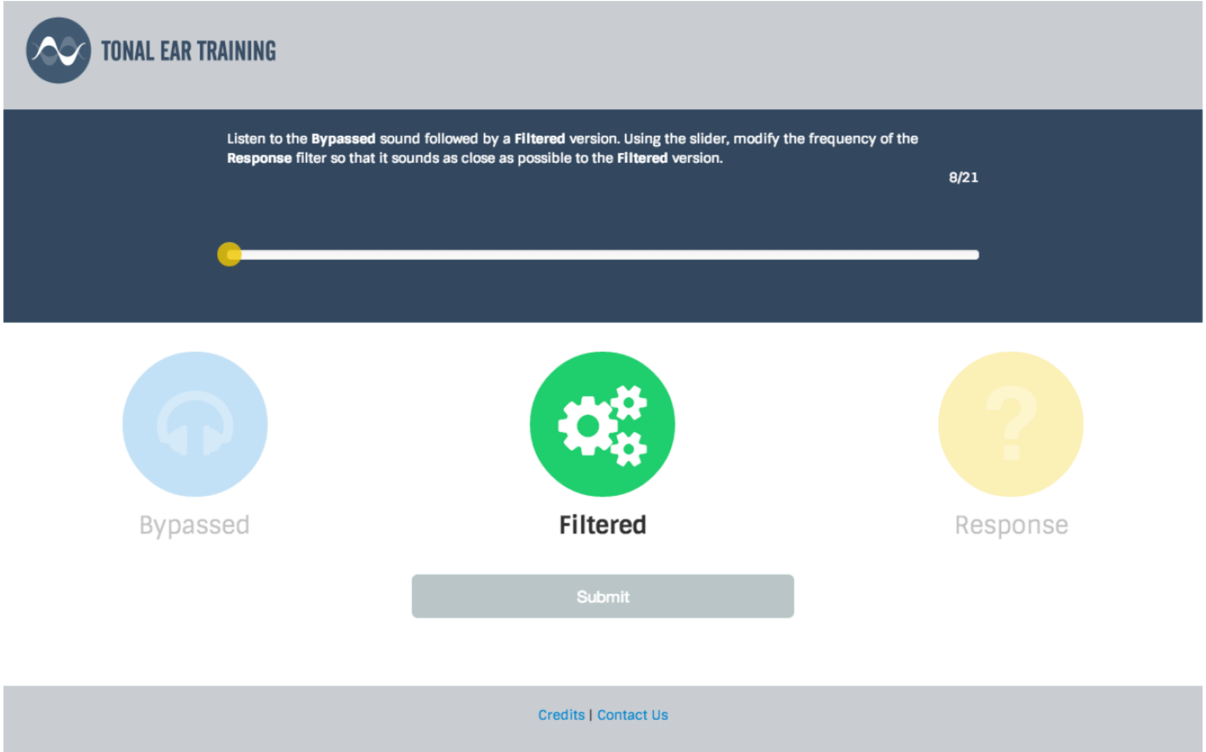

Figure 101. The warm-up / training page with the Filtered (processed) stimulus presented.

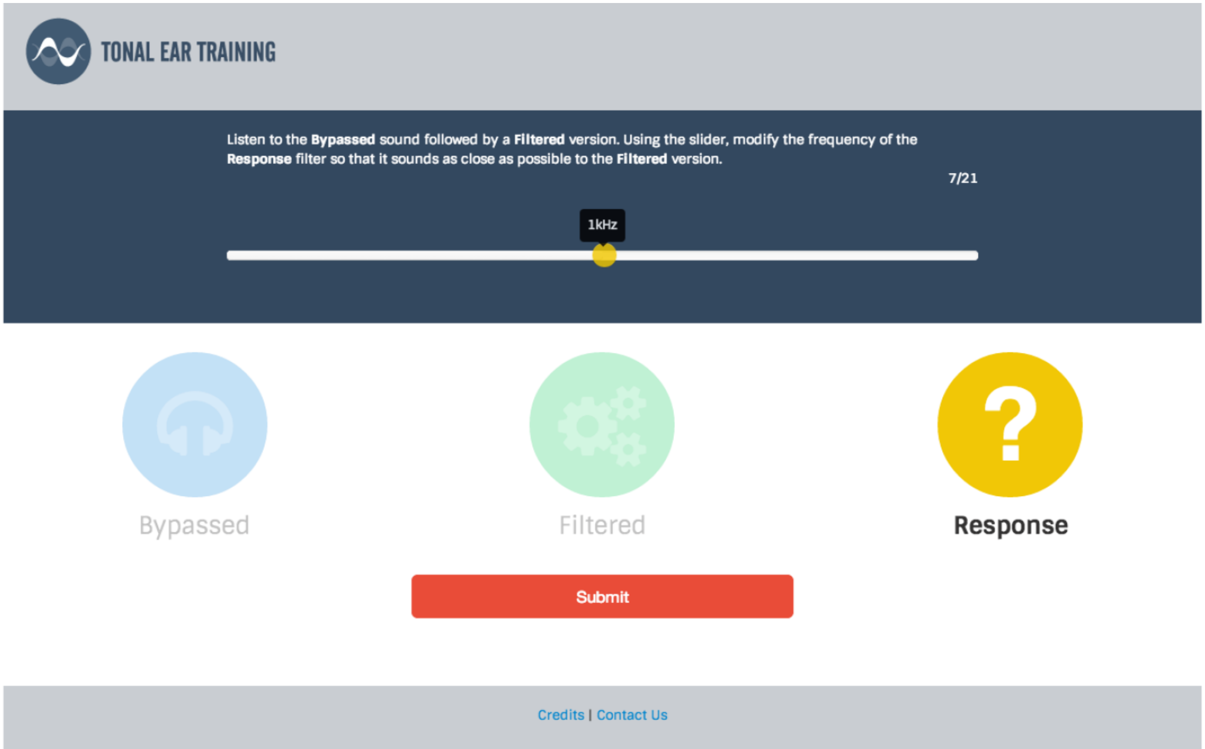

Figure 102. The warm-up / training page with the Response (user-processed) stimulus presented.

If the user responded correctly, the frequency slider control changed colour to green and the 'Response' icon changed from a yellow question mark to a green tick. If the user responded incorrectly, the frequency slider control changed colour to red and a second green-coloured indicator appeared on the frequency slider showing the correct answer (Figure 103). In addition, the 'Response' icon changed from a yellow question mark to a red cross. Finally, the filter applied at the user-selected centre frequency changed in real-time to match the on-screen green indicator. This allowed the user to first hear their response, then when they clicked 'Submit' and were incorrect, hear the applied filter change centre frequency as it moved throughout the spectrum to the correct frequency. This permitted the user to not only see what the correct answer was, but to hear what the correct centre frequency sounded like. The next question was presented after a short delay.

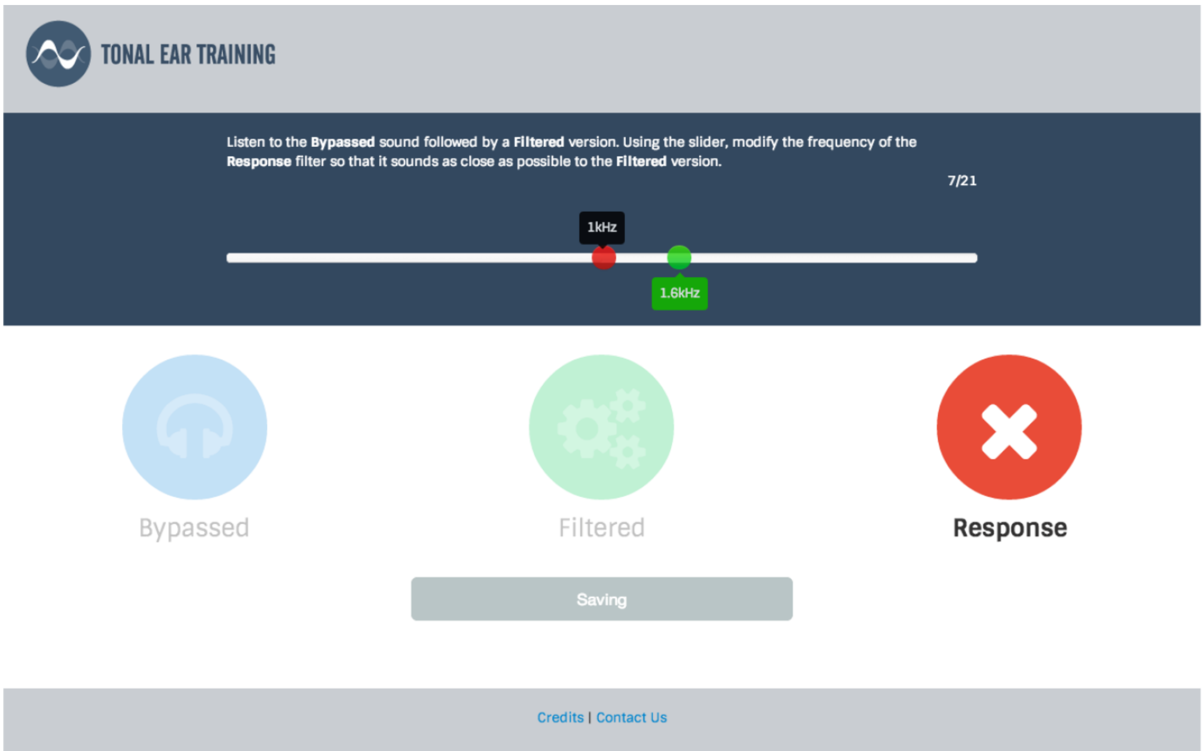

Figure 103. The warm-up / training page after the user has responded incorrectly.

Users were required to complete at least one run through each practice set (five sets corresponding to the five stimulus types). To being the training session, users had to obtain a score of 100% correct on each of the practice sets, taking as many attempts as required.

Prior to undertaking the training session, users had to complete the pre-training tone colour discrimination task detailed in Chapter 6, indicated by the blue chemistry flask icons in Figure 98. After completing the pre-training task, users began the training session.

The training session tasks were identical to that used in the orientation task, with two exceptions. First, for each training session of 21 questions, the stimuli were randomly selected (with replacement) from a pool of 21 (five by each of four genres and one pink noise signal) sound files. Second, a pool of 63 1/9th octave frequencies between 90 Hz and 10.8k Hz inclusive (Table 21) were used for the centre frequency of the applied response filter. Each 1/3rd octave frequency shown in red in Table 21 had two associated 1/9th octave frequencies, one higher and one lower. These three frequencies formed a set, referenced by the 1/3rd octave 'centre' frequency of the set. The centre frequency of the applied parametric equaliser was chosen randomly (without replacement) from one of the 21 'sets' of frequencies. As such, only one frequency from each set was presented in a given training session of 21 questions. However, the frequency slider only permitted the user to select 1/3rd octave centre frequencies. As opposed to the practice sets in which for each question there was a direct corresponding answer, in the training phase of the program students indicated which 1/3rd octave frequency (shown in red in Table 21) was closest to the filter frequency that was applied.

After completing each training session of 21 questions, students returned to the Home page. After completing five training sets, students were again presented with the pre/post-training test, then five more training sets, and so on.

7.2.4 User Performance Indicators

At the completion of a training session, users were presented with their performance data for that session via the Session Summary page (Figure 104). The percentage of correct responses was shown as text and also via a donut chart. The user's mean accuracy for all 21 questions in the preceding training session was shown in the same manner as the percentage correct score. Accuracy was calculated by comparing the correct response to the actual response. If these two frequencies matched, the accuracy of the answer is 100%. If the response was within

Table 21. The 63 1/9th-octave centre frequencies (in Hz) used in the training program.

1	90	**22**	467	**43**	2333
2	100	**23**	500	**44**	2500
3	108	**24**	540	**45**	2700
4	117	**25**	583	**46**	2916
5	125	**26**	630	**47**	3150
6	135	**27**	680	**48**	3402
7	146	**28**	735	**49**	3675
8	160	**29**	800	**50**	4000
9	173	**30**	864	**51**	4320
10	187	**31**	933	**52**	4666
11	200	**32**	1000	**53**	5000
12	215	**33**	1080	**54**	5400
13	232	**34**	1167	**55**	5833
14	250	**35**	1250	**56**	6300

15	270	**36**	1350	**57**	6804
16	292	**37**	1458	**58**	7349
17	315	**38**	1500	**59**	8000
18	340	**39**	1620	**60**	8640
19	367	**40**	1750	**61**	9332
20	400	**41**	2000	**62**	10000
21	432	**42**	2160	**63**	10800

one 1/3rd-octave band frequency either side of the correct response, the accuracy of the answer was 90%, within two 1/3rd octave band frequencies either side, the accuracy of the answer was 80%, and so on. Responses that were 11 or more 1/3rd octave band frequencies away from the correct answer in either direction scored an accuracy of 0%.

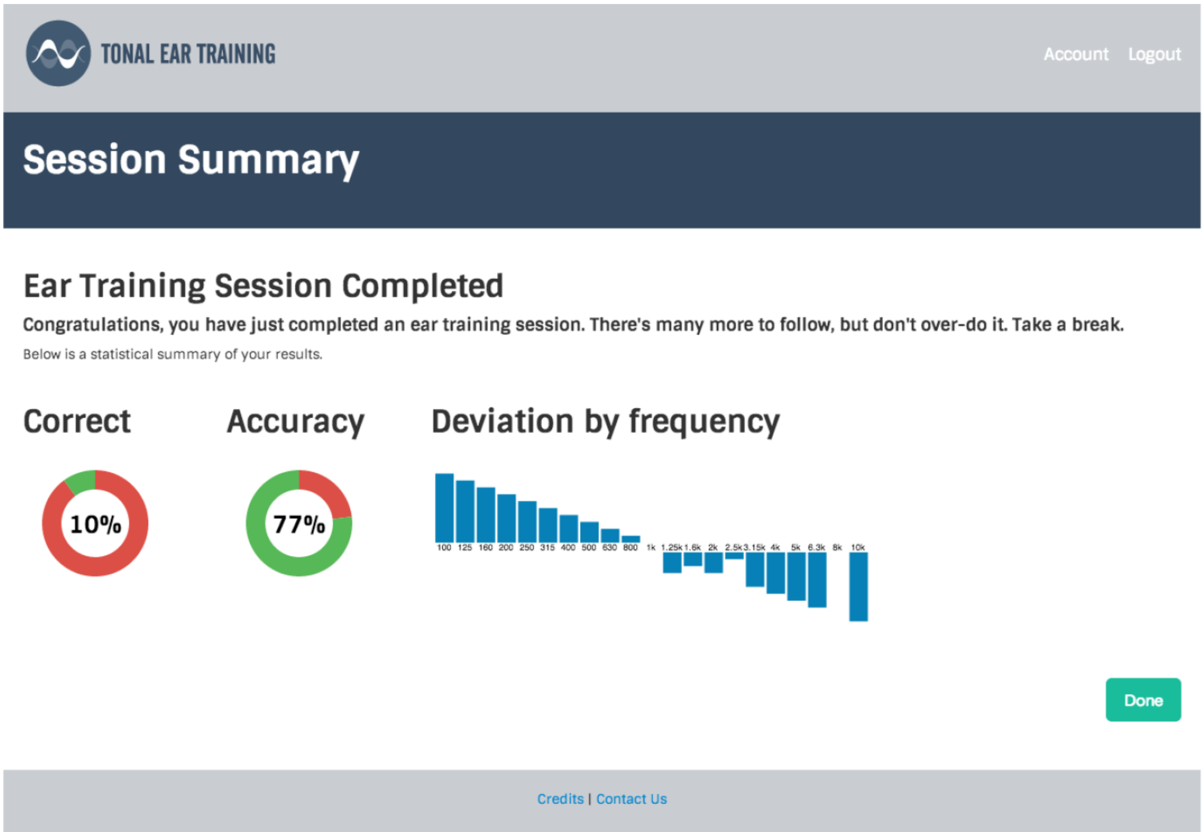

Figure 104. The training Session Summary page displayed summarising the user's performance.

The user's 'Deviation by Frequency' was also summarised (Figure 104). For each of the 21 1/3rd octave band response frequencies that were presented, the direction and magnitude of the user's response deviation from the correct frequency was shown. A blue column above a frequency indicated that for the presented frequency, the user responded incorrectly and higher than the correct answer. If the column was displayed below the frequency, this indicated that the user responded incorrectly and lower than the correct answer. The absence of a blue column indicated that the user answered correctly. Frequencies at the limits of the spectrum allowed for greater deviation from the correct response.

Upon returning to the Home page (Figure 98), the user's overall performance was summarised. The Deviation by Frequency display presented the mean deviation for all training sessions for each 1/3rd octave frequency. The user could set the display to show data for all stimuli, for each genre of music, or for pink noise. The Deviation by Frequency display is essentially a display of the magnitude and direction of the user's accuracy.

7.2.5 Administration Control Panel

The Administration Control Panel (Figure 105) allows users with administration privileges to export to a .csv file all training session data (including the tone colour discrimination test), user data or data for an individual participant.

7.2.6 Deployment within SAE Curricula

Tonal Ear Training was included as an assessment within the BAP120 Signal Processing and Aural Perception module offered as part of the Bachelor of Audio Production degree at the SAE Institute's Australian degree centres (including Dubai) in the second trimester commencing in May 2013. Although engagement and completion

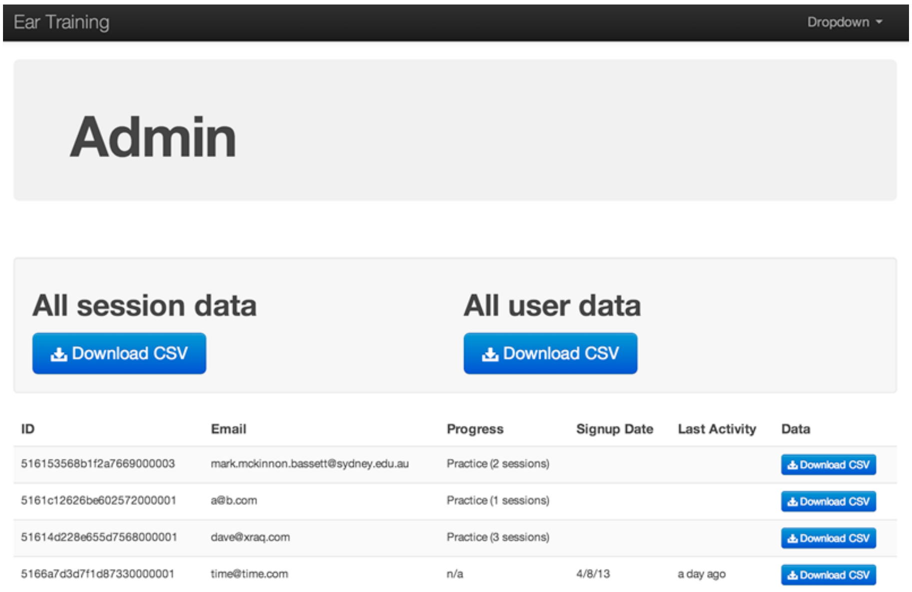

Figure 105. The user administration page.

of the training program over the course of the trimester was required in order to fulfil the requirements of the module, students had the option of opting out of the research (opting out of data collection) during the account sign-up phase. Therefore, it was mandatory for students enrolled in the BAP120 module to complete the TETP (if they wanted to fulfil the module requirements), but it was not mandatory to participate in the research.

7.2.7 Demonstration

To view a short video demonstration of the ICA version of Tonal Ear Training, visit http://audio.education/phd.html and enter the following credentials:

username: phd

password: MarkBassett

The password gives access to all videos on the website

7.3 SAE Parametric Equaliser Training

The SAE Parametric Equaliser Training application was not developed as part of a pre/post-training study or designed as a research tool. Rather, it was developed as a free, publicly available TETP designed to train users to identify the centre frequency of a parametric equaliser applied to any imported audio files. The TETP meets the first criterion detailed in Chapter 3, as the goal of the program is professionally relevant. However, it does not meet the second or third criterion, as performance relating to the goals of the TEPT was not measured. However, as discussed, this does not necessarily mean that the program is ineffective, rather, the lack of recorded performance data simply prevents claims regarding the effectiveness of the training program.

7.3.1 Design Considerations

Several requirements for the TETP were decided upon before development commenced. The most important decision was which type of TETP to develop. It was decided to implement a frequency spectrum-based TETP that utilised absolute identification tasks. The program needed to train the user to identify the centre frequency of a single parametric equaliser applied to any imported stereo audio files. The user was not required to identify the gain or the Q of the filter, and these parameters had to be user-modifiable throughout the training program. Depending on the version of the task, absolute identification tasks can allow the user to switch between the filtered sound and a flat version of the sound at their leisure, or present the flat and/or filtered sound once per trial. The decision was made to allow users to switch between flat and filtered signals as required.

As the application was designed for users with limited experience in audio engineering, the difficulty of the listening task required considerable attention. The task had to be easy enough initially for a novice engineer to be able to complete (answer correctly) most the time, otherwise, the user might simply disengage with the application through frustration. However, the task also needed to increase in difficulty as the user's skill increased, to continue to provide a challenging but achievable task. Based on this, three stages of training were implemented: 2-octave, octave and 1/3rd octave.

7.3.2 Platform / Operating Environment

Having developed numerous pre/post-training tests and a TETP using Max software, the researcher decided to use Max to create this application. Max runs on both Windows and Mac operating systems, but due to the researcher's familiarity with Mac systems and their extensive use in the audio industry, a Mac-only application was the obvious course. The application needed to be distributed via the Apple App Store, so the researcher had to become an Apple Developer. Development of the application for deployment on mobile devices was ruled out early in the planning phase as mobile devices are commonly used in environments with high levels of background noise and in conjunction with low-quality headphones.

7.3.3 Branding

Although the researcher was an employee of the SAE Institute throughout the duration of his PhD candidature, the SAE Parametric Equalisation Training application was not commissioned or developed by the SAE Institute. The application was developed independent of SAE, outside of working hours; the researcher received no compensation of any kind for the development of the application from SAE or any other party. During the development of the application the researcher approached national management at the SAE Institute in Australia, inquiring as to the possibility of including the SAE logo on the application's interface and naming the application 'SAE Parametric Equaliser Training'. Permission was obtained from the Dean of the SAE Institute Southern to use the SAE logo on the interface and to name the interface using SAE. Rights were not assigned or licensed to the SAE Institute; the application remains the property of the author. Branding the application with the SAE logo and name provided advantages for internal and public distribution and promotion, as discussed below.

7.3.4 Distribution and Marketing

The decision was made in the planning stage to distribute the application free of charge and without 'in-app' purchases. This was to ensure that anyone with access to a Mac computer could use the application regardless of their financial circumstances, and that the application would reach the largest possible audience. Inclusion in the Apple Store (Figure 106) allowed global distribution of the application and automatic distribution of future updates. Upon completion, the application was promoted internally via email to all audio staff and faculty at SAE Institute campuses around the world. Staff members were asked to promote the application to their students and use it in their classes if they saw it fit to do so. As the application is branded with the SAE logo, the expectation was that faculty and students at SAE would be more likely to download and use the application. The use of the SAE logo was also expected to provide advantages in promoting the application outside of SAE, because the SAE brand is well known in the audio education industry due to its 50-plus campuses worldwide. Potential users' probable familiarity with the brand name was anticipated to result in the application being accessed and promoted more widely. A website (Figure 107) was created to promote the application, provide a link to the Apple App Store page, and to provide more information about the application.

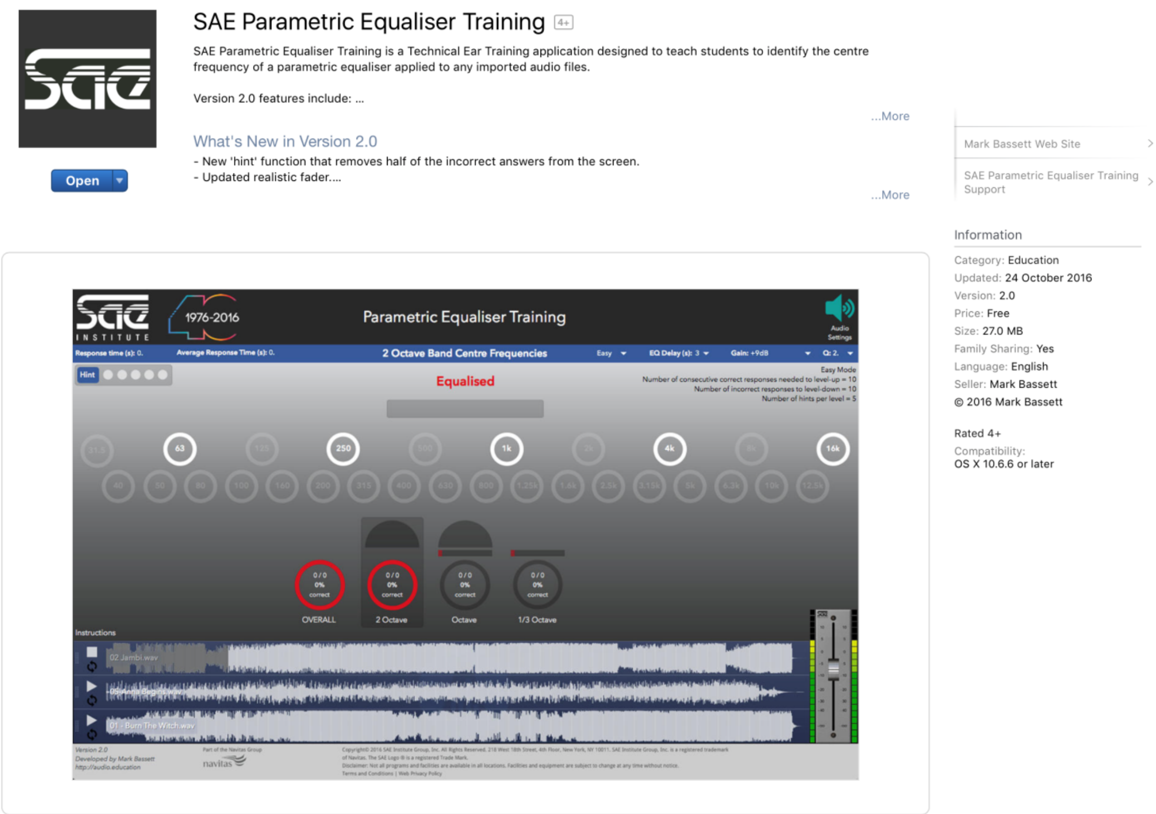

Figure 106. The SAE Parametric Equaliser Training application listed in the Apple App Store (version 2.0 shown).

Figure 107. The SAE Parametric Equaliser Training website.

7.3.5 Target Market

The application was designed to be used by any person interested in learning to identify the centre frequency of a parametric equaliser. It was thought that students studying audio engineering, individuals interested in audio engineering and novice engineers would be particularly interested in using the application. In addition, it was hoped that SAE Institute campuses around the world would install the application on their local machines and promote the application to their students. As the application is branded with the SAE logo, it was not expected that educational institutions other than SAE would install the application on their campus machines.

7.3.6 Release Versions

The initial released version of the application (Figure 108) was later updated to the version shown in Figures 109–116. The update included the addition of a hint function, an update of the fader, and cosmetic changes.

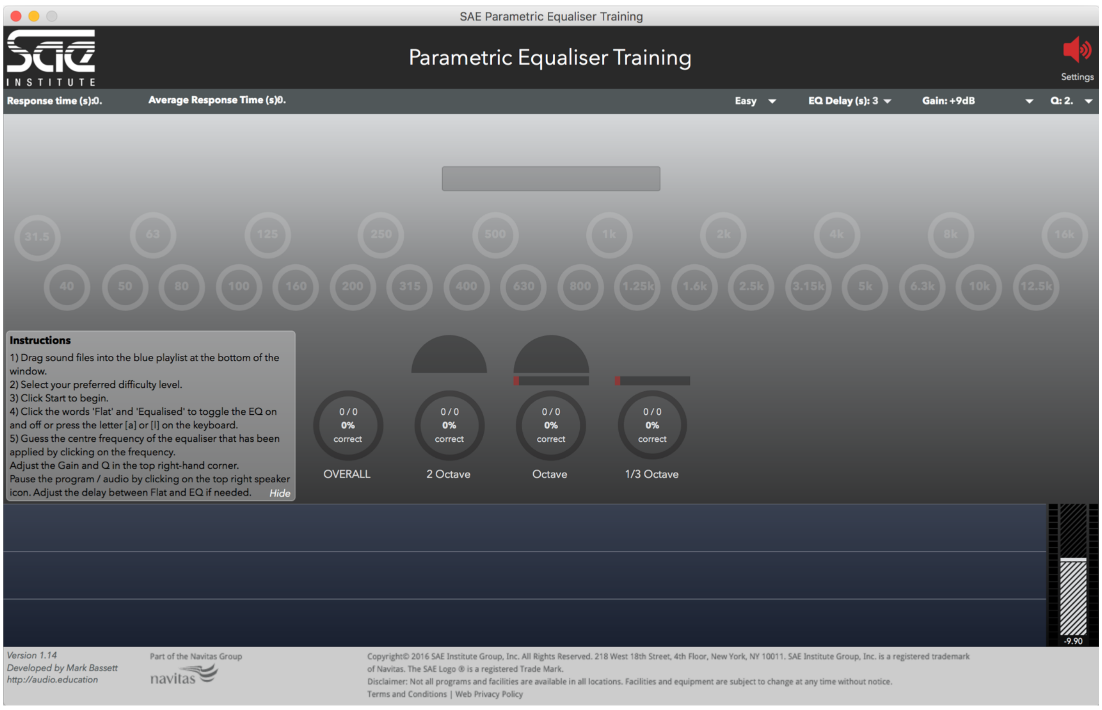

Figure 108. The first released version (1.14) of the application.

7.3.7 WALKTHROUGH – VERSION 2.0

Upon loading the application, users are first presented with a splash screen[10] (Figure 109), then after a short delay, the screen switches to show the application interface (Figure 110). A flashing text bubble prompts users to import up to three audio files into the program by dragging them onto the blue playlist at the bottom of the screen. The application supports stereo or mono .wav, .aiff, .mp3, .flac and .m4a audio files.

Once at least one audio file has been imported, users are prompted to select one of three difficulty levels: Easy, Hard or Expert. The selected difficulty level does not affect the difficulty of the presented questions; this is dictated by the training stage, the gain and Q settings and the delay time setting. The selected difficulty level (Table 22) dictates the number of available hints per stage, the number of consecutive correct questions needed to trigger stage promotion (where possible), and the number of total incorrect questions to trigger stage demotion (where possible). Stage promotion refers to either the transition from a 2-octave frequency question set to an octave frequency question set, or from an octave frequency question set to a 1/3rd octave frequency question set. Stage demotion refers to the transition between question sets in the opposite direction. In this sense, the selected difficulty level controls the difficulty of move through the stages of training, as opposed to the difficulty of individual questions.

---

[10] Launch screen with the SAE logo and application name.

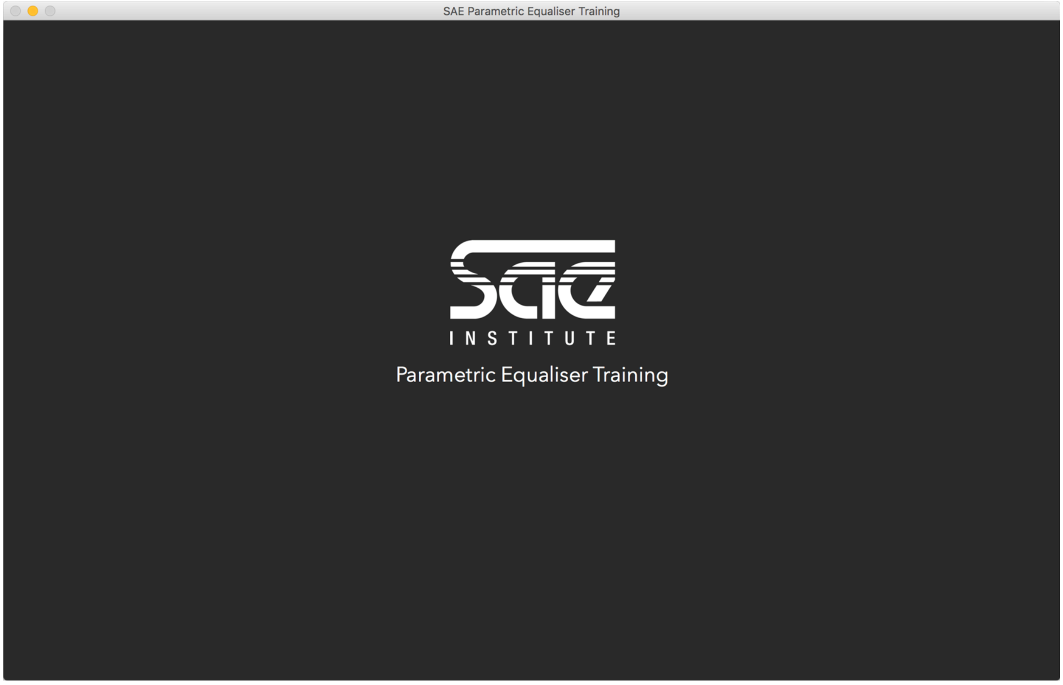

Figure 109. The splash screen.

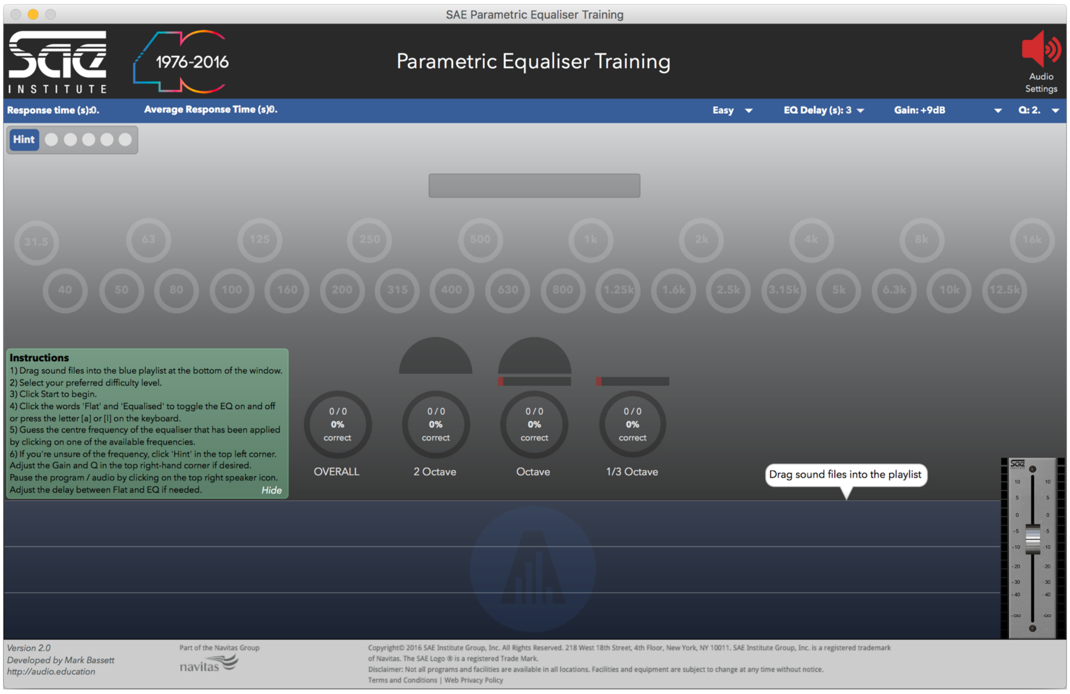

Figure 110. The initial training program interface splash screen.

Table 22. The number of hints, consecutive correct questions for stage promotion, and total incorrect questions for stage demotion for each difficulty level.

Difficulty level	Hints per stage	Consecutive correct questions required for stage promotion	Total incorrect questions required for stage demotion
Easy	5	10	10
Hard	3	20	5
Expert	1	30	2

Once the difficulty level is selected, the user is prompted to click 'Start'. Playback begins and the Stage One (two-octave) screen is shown (Figure 111). The first audio file in the playlist starts, initially with no filter applied. After the delay time shown in the 'EQ Delay' field (the default is three seconds), a parametric equaliser is applied to the signal. In Stage One, the filter is chosen from the five 2-octave band frequencies shown in white in Figure 111. The parameters of the filter are selectable by the user and default to +9 dB, Q=2.0.

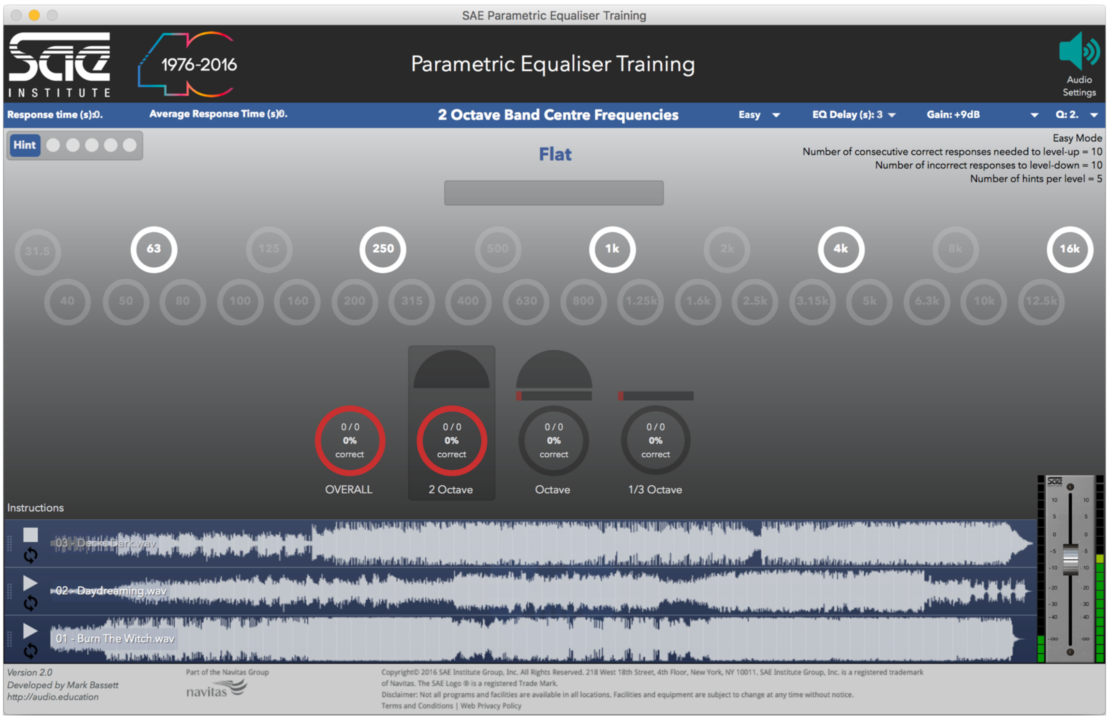

Figure 111. The Stage One (two-octave) GUI with the instructions hidden.

The task requires the user to identify the centre frequency of the parametric filter (absolute identification) applied to the music track from the five two-octave frequencies displayed in white on the screen in Figure 111 (unless the hint function is used, in which case the number of frequencies on screen reduces to three). The user can toggle between the flat and equalised versions of the signal at any time by either clicking on 'Flat' / 'Equalised' or by pressing the letter 'a' or 'l' on the keyboard. The user has an unlimited amount of time to select the centre frequency. If the user's response is correct, the word 'Correct' is displayed with a green background, the white circle surrounding the chosen frequency changes temporarily to green, and the performance indicator for the stage is temporarily highlighted in green (Figure 112). If the user is incorrect, the word 'Incorrect' is displayed with a red background, followed by 'the correct answer was X Hz'. In addition, the white circle surrounding the chosen frequency changes temporarily to red, the circle surrounding the correct frequency changes temporarily to green, and the performance indicator for the stage is temporarily highlighted in red (Figure 113). Once users answer 10 (on Easy difficulty) consecutive questions correctly, they are promoted to the octave stage of training (Figure 114).

The octave stage again requires the user to perform an absolute frequency identification task, but the number of possible centre frequencies expands to 10. If the total number of incorrect questions on the octave stage of training reaches 10 (Easy difficulty), the user is demoted to the two-octave stage. If the user answers 10 consecutive questions correctly (Easy difficulty) on the octave stage, they are promoted to the 1/3rd-octave stage (Figure 115).

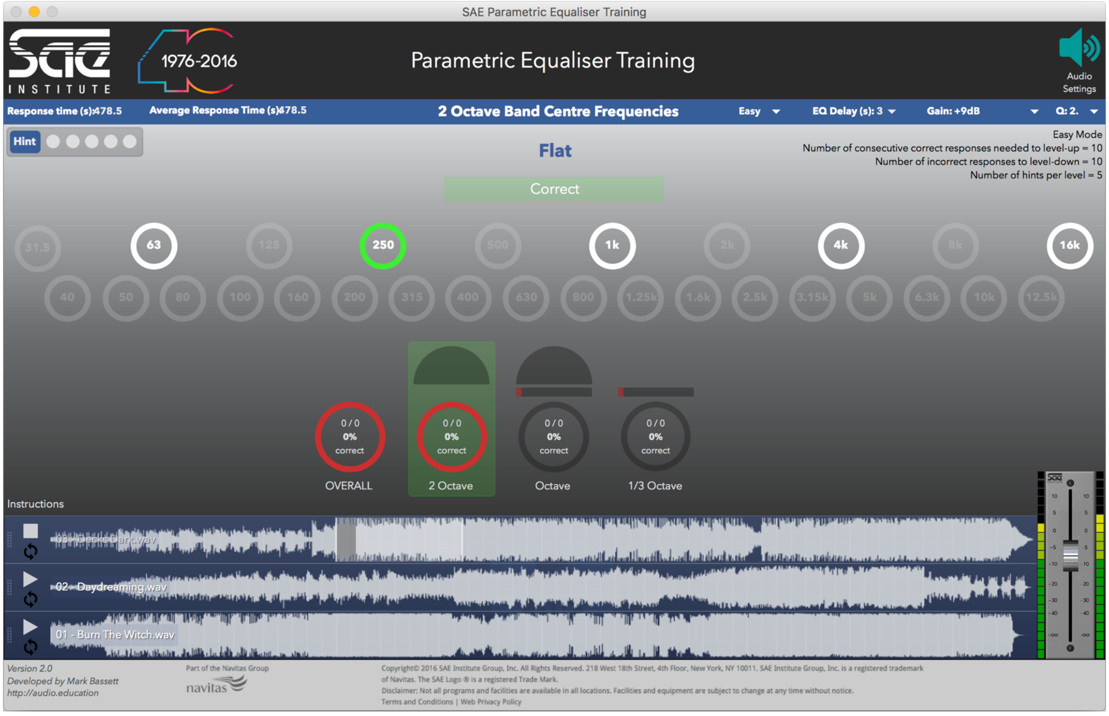

Figure 112. The GUI response when the user selects a correct answer.

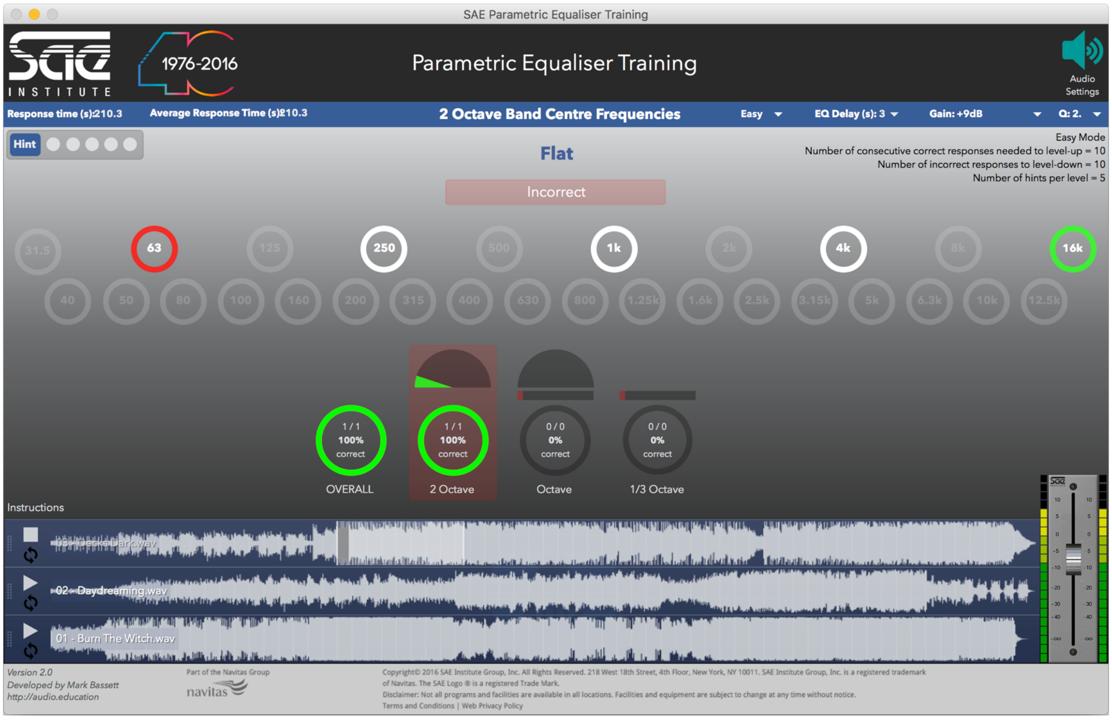

Figure 113. The GUI response when the user selects an incorrect answer.

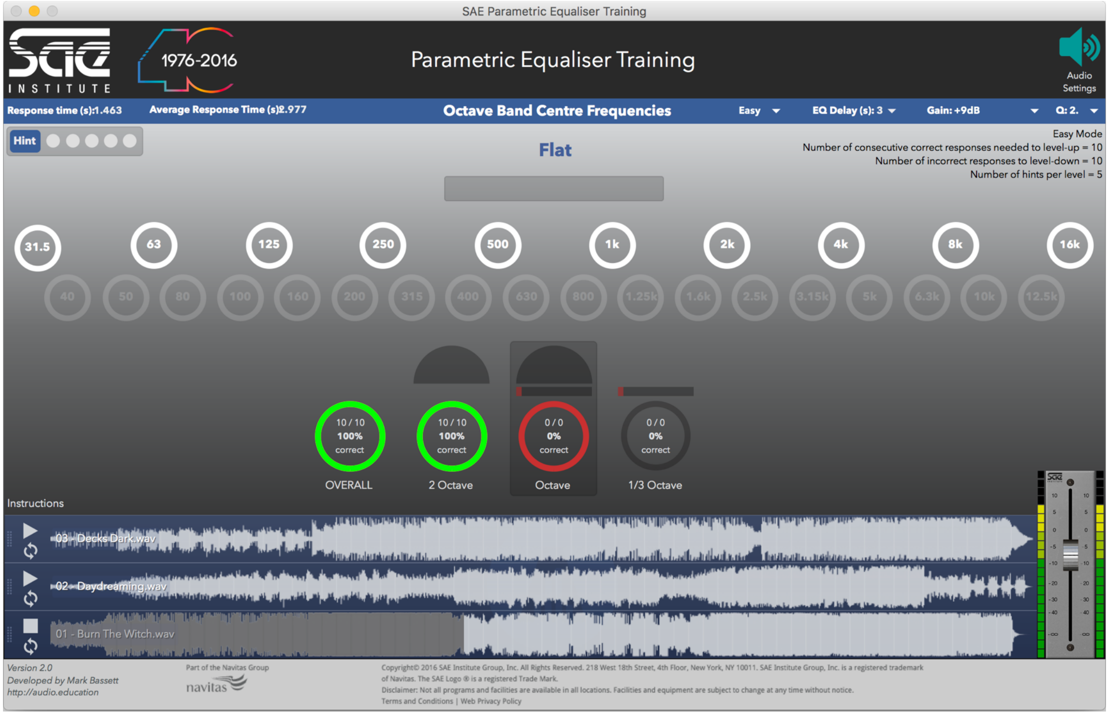

Figure 114. The Stage Two (octave) GUI with the instructions hidden.

Once the user reaches the 1/3rd-octave stage of training, the number of possible centre frequencies increases to 28. The user can be demoted to the octave stage by answering 10 questions incorrectly (Easy difficulty), but cannot be promoted to any higher stages. As such, the training program does not have an end point – the training session will continue indefinitely if desired.

The layout of frequencies was designed to encourage the user to use the octave frequencies (the top row of frequencies in Figure 115) as reference frequency anchor points; once the user estimates which octave-band centre frequency the applied filter is closest to, they then estimate if the centre frequency is higher, lower or equal to this estimated octave-band frequency.

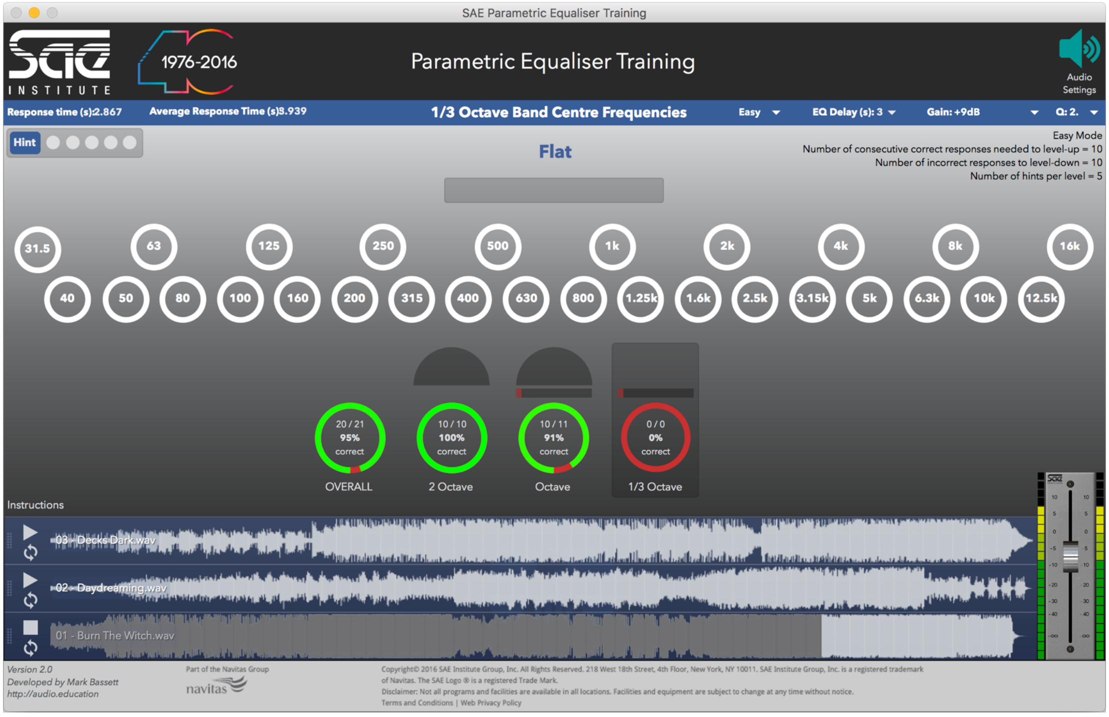

Figure 115. The Stage Three (1/3rd-octave) GUI with the instructions hidden.

The current version of the application features a 'Hint' button as detailed above. When clicked, approximately[11] half of the incorrect answers are removed from the screen (Figure 116). The hint button can only be used once per question and the number of hints per stage of training is dictated by the chosen difficulty level (Table 22).

7.3.8 User Performance Indicators

The GUI features several performance indicators (Figure 117). The meter on the far left shows the user's overall performance, displaying the total number of correct responses, the total number of questions attempted, and the percentage of correct responses. The coloured circle surrounding these performance indicators also changes

[11] If the number of incorrect answers is odd, one is added in order to remove half the incorrect answers. For example, at the octave stage, there are nine incorrect answers, but five incorrect are removed with the hint function as (9+1)/2=5

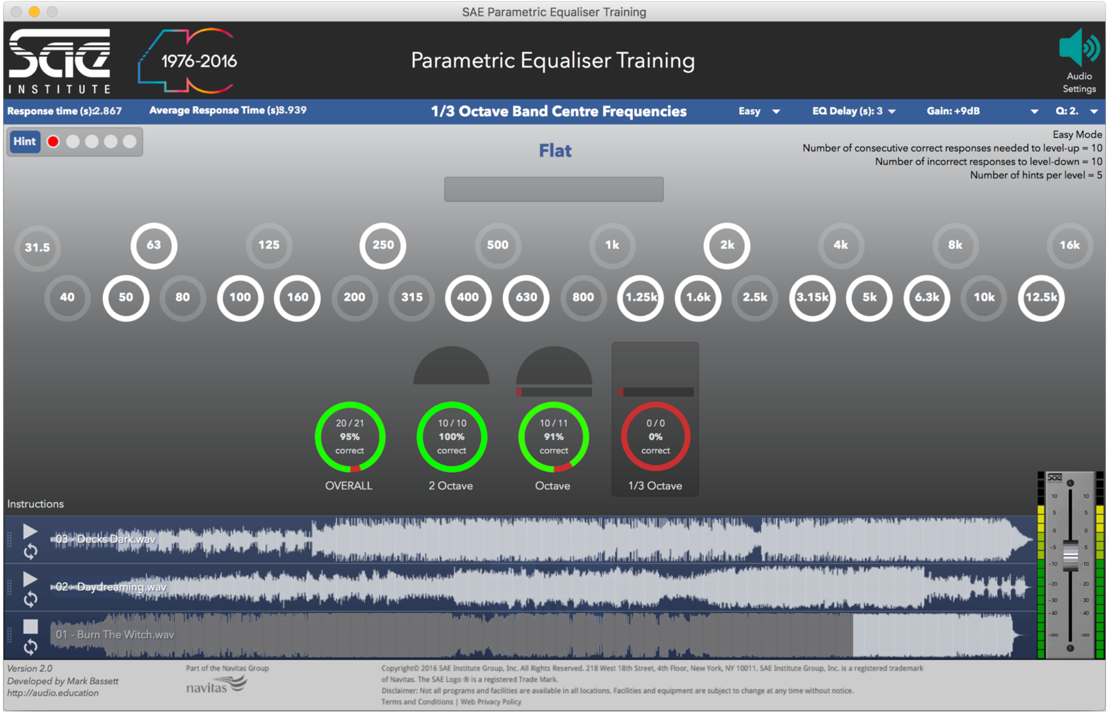

Figure 116. The Stage Three (1/3rd-octave) GUI with approximately half the incorrect answers removed after the Hint button has been clicked.

colour and circumference depending on the percentage of correct responses. Zero per cent correct changes the colour to red, while 100% correct changes the colour to bright green. Percentages between 0% and 100% are represented by a transition through the range of colours (red, orange, yellow, green). Each stage of training (2-octave, octave, and 1/3rd octave) has separate performance indicators. A performance meter is also included for each training stage.

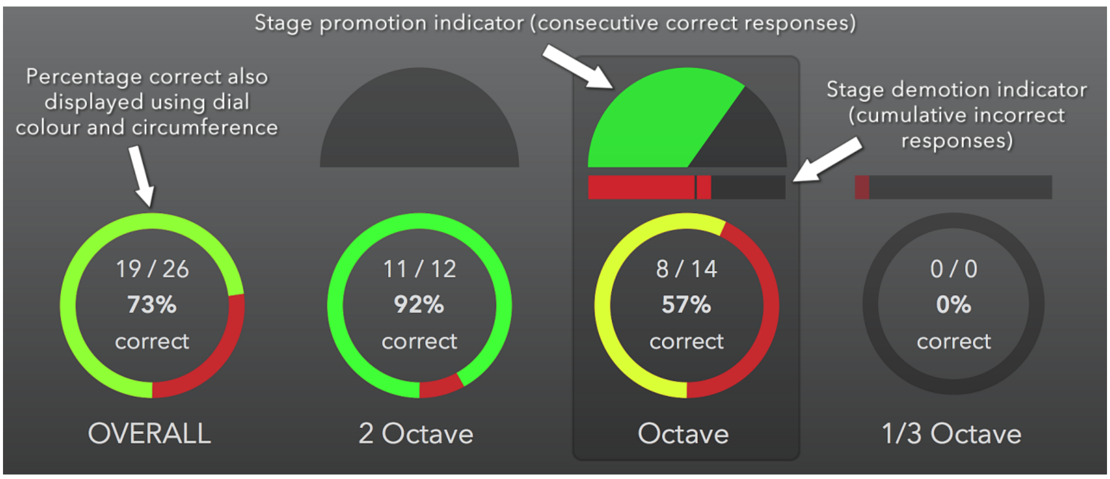

Figure 117. The performance indicators shown on the GUI.

The two-octave and octave performance indicators feature a stage promotion meter. This semi-circle is initially grey, but its area begins to fill with green as the user answers more and more consecutive questions correct. Once this meter fills up, the user is promoted to the next stage. If the user answers a question incorrectly, or if the user is promoted or demoted, the stage promotion meter resets. The 1/3rd octave stage does not have a stage promotion meter as there is no stage above it. The number of consecutive correct questions needed for stage promotion is dependent on the selected difficulty level (Table 22).

The 2-octave and 1/3rd-octave performance indicators also feature a stage demotion meter. This meter is initially grey, but its area begins to fill up with red as the user answers more and more questions incorrectly. The meter shows the cumulative total of incorrect answer for the stage, as opposed to consecutive incorrect answers. If the user get a question correct, the stage demotion meter does not reset, however it does reset if the user is promoted or demoted. Once the meter fills up, the user is demoted down a stage. The two-octave stage does not have a stage demotion meter as there is no stage below it. The number of cumulative incorrect questions needed for stage demotion is dependent on the selected difficulty level (Table 22).

The GUI displays the user's response time (measured from the onset of the applied filter) for each question and the mean response time for all questions. The current version of the application also features an 'Easter Egg'[12]; if the user answers the first 10 questions correctly (promoted to Stage 2 without answering a question

[12] An undocumented feature in a computer program, included in this case as a bonus.

incorrectly), VU meters are displayed in the header of the interface for the remainder of the training session.

7.3.9 Engagement

The application was downloaded more than 1,000 times in the first seven weeks of release (Figure 118) and 3,530 times in the first 12 months of release (Figure 119) (Apple Pty Ltd, 2017). In addition, at the time of writing (October 2017), the application had been downloaded 4,440 times in total across 91 countries (Apple Pty Ltd, 2017).

Figure 118. iTunes Connect Sales and Trends page showing total downloads in the first seven weeks of release (Apple Pty Ltd, 2017).

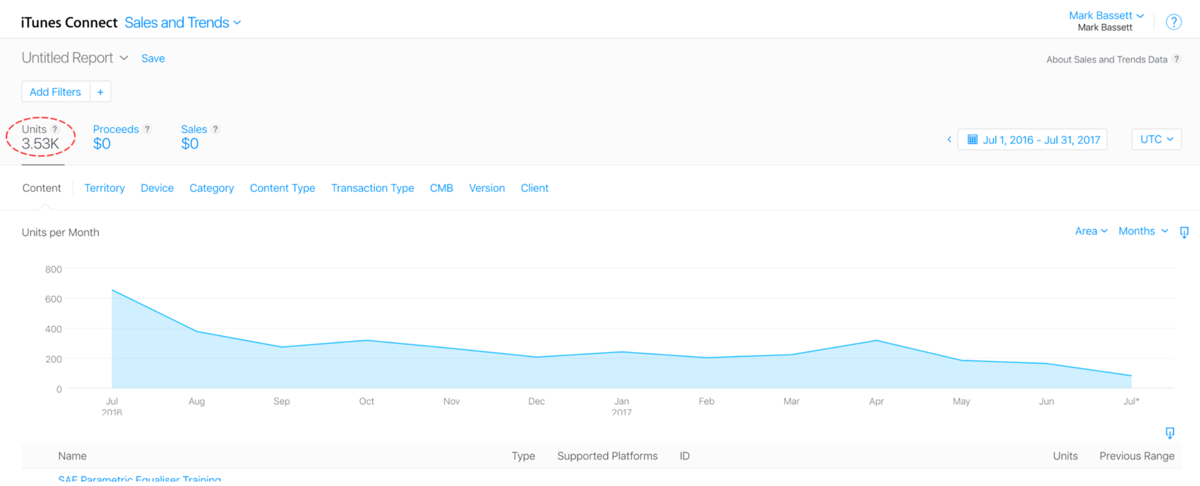

Figure 119. iTunes Connect Sales and Trends page showing total downloads in the first 12 months of release (Apple Pty Ltd, 2017).

7.3.10 Demonstration

To download the SAE Parametric Equaliser Training program, please visit http://technicalear.training/. To view a short video demonstration of the SAE Parametric Equaliser Training program, please visit http://audio.education/phd.html and enter the following credentials:

username: phd

password: MarkBassett

The password gives access to all videos on the website.

7.4 Summary

Two TETPs were newly developed as part of the research – the SAE Frequency Identification Training program used in the pilot study and the Tonal Ear Training program used in the full-scale study. These programs were developed to provide TETPs to be used in the pre/post-training test scenarios. A third standalone TETP, the SAE Parametric Equaliser Training application, was developed as a free, publicly available program accessible by anyone with access to an Apple Mac computer.

The following chapter provides a summary of the research, details areas of future work in the field, presents the limitations of the research, and finally outlines the contribution the research has made to the field.

8 Conclusion

This research was designed to test the relative importance of using the ICA and ISA methods on students' ability to discriminate tone colour when training using a TETP. As a result of the research, an pre/post-training test independent to the TETP was developed and found to be an effective measure of the influence of training method on students.

8.1 Summary

Chapter 2 presents a descriptive overview of frequency spectrum-based student and consumer-targeted TETPs for which published papers or texts were available. Chapter 3 contains a review of the relevant TETPs in the literature against three criteria developed by the author to establish the minimum requirements to assess a TETP for effectiveness. Chapter 4 details the three tests of tone colour discrimination ability that were developed and deployed in isolation to determine their suitability for use in a pre/post-training test scenario. Chapter 5 details the results of deploying the three tests – the Tic-Tac® test, a vowel categorisation test, and a conga tone dissimilarity test – in a pilot study utilising a pre/post-training test scenario. Results of the pilot study demonstrated a difference in tone colour discrimination ability between students who had trained using the ICA method, and those who trained using the ISA method, on the conga tone dissimilarity test. A full-scale study was subsequently implemented (Chapter 6), again deploying the conga tone dissimilarity test in the context of a TETP (detailed in Chapter 7). Analysis of the results of the full-scale study found small but insignificant differences between the tone colour discrimination abilities of the ICA and ISA groups of students. Despite this, the preliminary results suggest that a pre/post-training test is an effective measure of the influence of training method on students, if the test features a task that is significantly different from the tasks on which students train in the TETP.

8.2 Future Work

Many TETP designers claim their training programs are effective, despite the absence of empirical evidence. Considerable research needs to be conducted in order to establish a robust method for ascertaining the effectiveness of TETPs, to ensure that developed critical listening skills are in fact transferable from the training program to the real world, and to ensure that the TETP's goals are met.

A researcher interested in designing a frequency spectrum-based TETP should look to best practice underpinned by research. The researcher must consider several elements of training program design, including: the ideal number of questions per training set; question type – passive versus active, matching, removing, or absolute identification; the user interface – whether to include a plot showing the applied filter response, or a spectrum analyser; and the regularity and duration of training sessions. The researcher will be primarily interested in how the above program design elements affect students' learning and skill development. Perhaps most importantly, the researcher will want to use a well-established, robust method for ascertaining the effectiveness of the TETP. However, despite the growing number of TETPs available and the claims made by their developers, research on methods for ascertaining effectiveness is essentially non-existent, so no recommendation about best practice for these program design elements is possible.

An example of a TETP methodology that has earned an unwarranted reputation of 'best practice' is the use of reference frequency anchor points. Numerous TETP developers, including the author, have designed TETPs so that students must first train using octave frequency bands; as the training program progresses, the number of frequency bands is increased to 1/3rd octave or higher. The rationale for this process is that students should be presented with a small number of frequencies (octave bands) that serve as reference anchor points, then once they have committed the sound of these reference frequencies to memory, they will be better equipped to undertake training with 1/3rd octave frequency bands, as opposed to training with 1/3rd octaves in the first instance. This rationale seems logical, but there are no studies in the field that examine whether this approach to training provides any benefit to students. Hence, investigating the effectiveness of TETPs that progressively increase the number of presented frequencies, is an obvious target for future research.

8.3 RESEARCH LIMITATIONS

A significant limitation of the research is the sample size in the pre/post-training tests deployed in the pilot study. The Tic Tac® and conga tone dissimilarity tests contained less than 10 participants in each of the ICA and ISA groups, and the sample size in the vowel categorisation test prevented the analysis of results. The significance of the observed differences in results between the ICA and ISA groups on the conga tone dissimilarity test is limited by this small sample size, and it is difficult to be confident that observed differences between groups apply to the larger population. The small sample size of the vowel categorisation test precluded analysis of results and therefore the test was not considered for the large scale study. A larger sample size in the pilot study may have shown a difference between training groups, which in turn may have resulted in the test being deployed in the full-scale study, where it may have served as a suitable test for tone colour discrimination ability.

The stimuli used in the pre/post-training tests and the newly developed TETPs were all mono. The results of this research study do not necessary apply to stereo stimuli.

Participants in the full-scale study undertook either ICA or ISA-specific warm-up sessions prior to undertaking the pre-training test. As such, the pre-training test may have been delivered prior to students undertaking the TETP, but the students had already completed 105 warm-up questions before encountering the test. However, as the warm-up questions were delivered once only and students were able to complete the warm-up questions in a relatively short period of time (5-10 mins), the duration of the exposure to the training group-specific questions was limited. The pre-training task may not have been strictly 'pre-training', but the impact of this on the pre-training results between groups (which was near identical) is assumed to be limited.

Variation between the type of listening environment and the quality of the playback device may have impacted the results of the full-scale study. Participants were permitted to undertake the TETP and tone colour discrimination tests using their own headphones. Although students were instructed to use headphones, it is assumed that some students may have completed the training and test using loudspeakers. The TETP and tone colour discrimination tests are ideally completed in an environment with appropriately low levels of background noise. It is not uncommon for audio engineering students in Australia to own laptop computers and students at SAE Dubai were offered the opportunity to purchase a discounted laptop computer at the time of the study. As such, it is assumed that a proportion of the students may have completed the TETP and/or tone colour discrimination tests using their laptop, which affords portability on one hand, but may inadvertently facilitate higher levels of background noise on the other. The varying frequency response of the playback device and the varying level of background noise may affected the results of the TETP and tone colour discrimination tests in unpredictable and inconsistent ways, resulting in unproductive training on TETP and unreliable results on the tone colour discrimination tests.

A cultural bias may have been present in the full-scale study based on the type of music stimuli presented, particularly considering that many students enrolled at the SAE Dubai campus participated in the study. Although multiple genres of music were used as stimuli, the music tracks (songs) were selected entirely from Western culture. Familiarly with the style and culture of the music and the language (English) of the vocals, may have impacted the students' focus and concentration on the task at hand.

8.4 Contribution to the Field

This is the first research study to investigate the transfer of tone colour discrimination ability developed within a TETP to an external task that is sufficiently different from those encountered in the TETP, and the first research study to investigate the influence of the ICA and ISA training methods on this ability. For researchers seeking to establish the effectiveness of a TETP, this research detailed the significance of collecting data that directly relates to the goals of the TETP, and crucially, on a task that is sufficiently different from those encountered in the TETP.

The research is a first step towards establishing the influence of the ICA and ISA training methods on students' ability to discriminate tone colour. Establishing whether either method is more effective has important implications for audio engineering curricula and the way students learn to operate equalisers.

REFERENCES

Apple Pty Ltd. (2017). iTunes Connect. Retrieved from https://itunesconnect.apple.com/login

Arai, T., Satoh, F., Nishimura, A., Ueno, K., & Yoshihisa, K. (2006). Demonstrations for education in acoustics in Japan. *Acoustical Science and Technology, 27*(6), 344-348.

Audio Engineering Society. (2017a). AES education directory by geographic area. Retrieved from http://www.aes.org/education/directory/region.cfm

Audio Engineering Society. (2017b). Focal Press: AES presents. Retrieved from http://www.aes.org/publications/focal/

Australian Broadcasting Corporation. (2006). Loudness Meter Comparison Utility: Australian Broadcasting Corporation.

Australian Government. (2017a). Qualification details: CUA40915 - Certificate IV in Music Industry (Release 1). Retrieved from https://training.gov.au/Training/Details/CUA40915

Australian Government. (2017b). *Trade mark 496531*. Australia: Australian Government Retrieved from https://search.ipaustralia.gov.au/trademarks/search/view/496531?q=tonmeister

Australian Government. (2017c). Unit of competency details: CUSMGE12A - Maintain and expand music knowledge and critical listening skills (Release 1). Retrieved from https://training.gov.au/Training/Details/CUSMGE12A

Avid Technology Inc. (2010). MDW hi-res parametric EQ plug-in guide. Retrieved from http://akmedia.digidesign.com/support/docs/MDW_EQ_Plugin_Guide_v30_67830.pdf

Avid Technology Inc. (2015). Audio plug-Ins guide. Retrieved from http://akmedia.digidesign.com/support/compressed/AUDIO_PLUGINS_GUIDE_V12_3_E_87894.pdf?ns=1

Bartlett, B., & Bartlett, J. (2008). *Practical recording techniques: The step-by-step approach to professional audio recording.* (5th ed.). Burlington, MA: Focal Press.

Berchtold, A. (2016). Test-retest: Agreement or reliability? *Methodological Innovations, 9*(1-7).

Berman, R. (1999). *Basics of mixing: A beginner's guide to the art of mixing*: Roland Corporation.

Bielmeier, D. C. (2013). *Why didn't you learn this at recording school: Critical comments by employers*. Paper presented at the Audio Engineering Society Conference: 50th International Conference: Audio Education, Murfreesboro, TN.

Borwick, J. (1973). *The tonmeister concept*. Paper presented at the Audio Engineering Society Convention 46, New York, NY.

Brixen, E. B. (1993). *Spectral ear training*. Paper presented at the Audio Engineering Society Convention 94, Berlin, Germany.

Bureau of Labor Statistics. (2016). Occupational outlook handbook: Broadcast and sound engineering technicians. 2016-17. Retrieved from https://www.bls.gov/ooh/media-and-communication/broadcast-and-sound-engineering-technicians.htm

Case, A. U. (2007). *Sound FX: Unlocking the creative potential of recording studio effects*. Burlington, MA: Focal Press.

Commonwealth of Australia. (2011). *Report of an audit of SAE Institute PTY Limited October 2011*. (1). Retrieved from http://teqsa.gov.au/sites/default/files/auditreport_sae_2011.pdf

Conversational chaining. (1978). In D. Unwin & R. McAleese (Eds.), *The encyclopaedia of educational media communications and technology*. London, UK.: The Macmillan Press LTD.

Corey, J. (2004). *An ear training system for identifying parameters of artificial reverberation in multichannel audio*. Paper presented at the Audio Engineering Society Convention 117, San Francisco, CA.

Corey, J. (2007). *Beyond splicing: Technical ear training methods derived from digital audio editing techniques*. Paper presented at the Audio Engineering Society Convention 123, New York, NY.

Corey, J. (2010). *Audio production and critical listening: Technical ear training*. Burlington, MA: Focal Press.

Corey, J. (2013). *Technical ear training: Tools and practical methods.* Paper presented at the International Conference on Acoustics, Montreal, Canada.

Corey, J. (2017). *Audio Production and Critical Listening: Technical Ear Training* (2nd ed.). New York, NY: Routledge.

Corey, J., & Benson, D. H. (2017). Parametric Equalization. Retrieved from http://webtet.net/apcl/#/parametric

Crich, T. (2005). *Recording tips for engineers: For cleaner, brighter tracks* (2nd ed.). Burlington, MA: Focal Press.

Davis, R., & Parker, S. (2013). *Collaboration, creativity, and communities of practice: Music technology courses as a gateway to the industry*. Paper presented at the

Audio Engineering Society Conference: 50th International Conference: Audio Education, Murfreesboro, TN, USA.

Eiosis LLC. (n.d.). *AirEQ: Beyond analgoue.* Retrieved from http://downloads.eiosis.com/AirEQ/AirEQ%20User%20Manual.pdf

Everest, F. A. (1982a). *Instruction in critical listening*. Paper presented at the Audio Engineering Society Convention 72, Anaheim, CA.

Everest, F. A. (1982b). Manual for critical listening: An audio training course: Mix Books.

Everest, F. A. (1986). Manual for auditory perception: An audio training course. Santa Barbara, CA: Puclished Privately.

Everest, F. A. (1997). *Critical listening and auditory perception: The complete audio-visual training course*: Mix Bookshelf.

Everest, F. A. (2007). *Critical listening skills for audio professionals*. Boston, MA: Thomson.

Fisher, D. (2015). *Audio education - a personal view*. Paper presented at the Audio Engineering Society Conference: UK 26th Conference: Audio Education, Glasgow, Scotland.

Gibson, B. A. (1996). *The AudioPro home eecording course Vol. I*: Artistpro.

GML LLC. (Producer). (2014, 2nd July 2017). Part 4 - MDWEQ5 on acoustic guitar. [Instructional Video] Retrieved from https://vimeo.com/100724761

Gottlieb, G. (2007). *Shaping sound in the studio and beyond: Audio aesthetics and technology*. Boston, MA: Thomson Course Technology PTR.

Greenberg, M., & Huddleston, D. R. (1971). A program for developing aural discrimination of instrumental tone colors using a videosonic teaching machine. *Journal of Research in Music Education, 19*(1), 51-61.

Hess, S., & Hu, E. (2011). How to listen (Version 2.0.4): Harman International Industries.

Hirsch, S., & Heithecker, S. (2006). *Pro Tools 7 session secrets: Professional recipes for high-octane results*. Indianapolis, IA: Wiley Publishing.

Hojan, E., & Rakowski, A. (1999). Teaching psychoacoustics in Poland. *Acustica, 85*, 623-625.

Huber, D. M., & Runstein, R. E. (2014). *Modern recording techniques* (8th ed.). Burlington, MA: Focal Press.

Indelicato, M. J., Hochgraf, C., & Kim, S. (2014). *How critical listening exercises complement technical courses to effectively provide audio education for engineering technology students*. Paper presented at the Audio Engineering Society Convention 137, Los Angeles, CA.

Iwamiya, S., Nakajima, Y., Ueda, K., Kawahara, K., & Takada, M. (2003). Technical Listening Training: Improvement of sound sensitivity for acoustic engineers and sound designers. *Acoustical Science and Technology, 24*(1), 27-31.

Izhaki, R. (2008). *Mixing audio: Concepts, practices and tools*. Burlington, MA: Focal Press.

iZotope Inc. (2015). *Mastering with Ozone: Principles, tips and techniques.* Retrieved from http://downloads.izotope.com/guides/iZotopeMasteringGuide_MasteringWithOzone.pdf

Jopson, N. (2013). Skills shortage looming. *Resolution, 12*.

Kaniwa, T., Kim, S., Terasawa, H., Ikeda, M., Yamada, T., & Makino, S. (2011). *Towards a personalized technical ear training program: An investigation of the effect of adaptive feedback*. Paper presented at the SMC Conference.

Kassier, R., Brookes, T., & Rumsey, F. (2004). *A simplified scene-based paradigm for use in spatial audio listener training applications*. Paper presented at the Audio Engineering Society Convention 117, San Francisco, CA.

Kassier, R., Brookes, T., & Rumsey, F. (2006a). *A comparison between spatial audio listener training and repetitive practice*. Paper presented at the Audio Engineering Society Convention 121, San Francisco, CA.

Kassier, R., Brookes, T., & Rumsey, F. (2006b). *Designing a spatial audio attribute listener training system for optimal transfer*. Paper presented at the Audio Engineering Society Convention 120, Paris, France.

Kassier, R., Brookes, T., & Rumsey, F. (2007). *Training versus practice in spatial audio attribute evaluation tasks.* Paper presented at the Audio Engineering Society Convention 122, Vienna, Austria.

Katz, B. (2002). *Mastering audio: the art and science* (1st ed.). Burlington, MA: Focal Press.

Katz, B. (2008). *From demo to delivery: The process of production* (R. Hepworth-Sawyer Ed. 1st ed.): Focal Press.

Kawahara, K., Takada, M., & Iwamiya, S. (2013). *Effectiveness of technical listening training in Department of Acoustic Design of Kyushu University.* Paper presented at the International Conference on Acoustics, Montreal, Canada.

Kawahara, K., Takada, M., Iwamiya, S., Nakajima, Y., & Ueda, K. (2004a). *Technical listening training: Demonstration class.* Paper presented at the International Commission for Acoustics, Kyoto, Japan.

Kawahara, K., Takada, M., Iwamiya, S., Nakajima, Y., & Ueda, K. (2004b). *Technical Listening Training: Systematic training program designed to improve auditory* *sensitivity.* Paper presented at the International Conference on Acoustcs, Kyoto, Japan.

Keane, J. (2007). *The musician's guide to Pro Tools* (2nd ed.). New York, NY: McGraw-Hill Osborne Media.

Kim, S. (2015). An assessment of individualized technical ear training for audio production. *Journal of the Acoustical Society of America, 138*(1).

Kim, S., Kaniwa, T., Terasawa, H., Yamada, T., & Makino, S. (2013). Inter-subject differences in personalized technical ear training and the influence of an individually optimized training sequence. *Acoustical Science and Technology, 34*(6), 424-431.

KIQ Productions Inc. (2005). Objectives of study. Retrieved from http://www.kiqproductions.com/html/objectives_of_study.html

Klein, F., & Werner, S. (2015). *Auditory adaptation in spatial listening tasks*. Paper presented at the Audio Engineering Society Convention 138, Warsaw, Poland.

Letowski, T. (1985). Development of technical listening skills: Timbre solfeggio. *Journal of the Acoustical Society of America, 33*(4), 240-244.

Letowski, T., & Letowski, S. (2011). *Localization Error: Accuracy and Precision of Auditory Localization* (P. Strumillo Ed.): Intech.

Letowski, T., & Miśkiewicz, A. (1995). *Development of technical listening skills for sound quality assessment*. Paper presented at the Inter-noise, Newport Beach, CA.

Lightner, J. W. (1993). *A survey of the professional audio industry in an eight state region to assess employers' perceived value of formal audio education and their perceived training needs for entry-level employees.* (Master of Science in Occupational Education), Ferris State University, Unpublished.

Liu, Z., Wu, F., & Yang, Q. (2007a). *The practice and study of ear training on discrimination of sound attributes*. Paper presented at the Audio Engineering Society Convention 122, Vienna, Austria.

Liu, Z., Wu, F., & Yang, Q. (2007b). *The training and analysis on listening discrimination of pure tone frequency*. Paper presented at the Audio Engineering Society Convention 122, Vienna, Austria.

Martens, W. L., McKinnon-Bassett, M., & Cabrera, D. (2012, April). Perceptual evaluation of stochastic-event-based percussive timbres for use in statistical sonification. Paper presented at the Audio Engineering Society Convention 132, Budapest, Hungary.

Martens, W. L., & Marui, A. (2005). *Predicting timbral variation for sharpness-matched guitar tones resulting from distortion-based effects processing*. Paper presented at the Audio Engineering Society Convention 118, Barcelona, Spain.

Martin, D., & Massenburg, G. (2015). *Advanced technical ear training: Development of an innovative set of exercises for audio engineers*. Paper presented at the Audio Engineering Society Convention 139, New York, NY.

Marui, A., & Kamekawa, T. (2013). *Towards the development of objective difficulty measure in technical ear training tasks.* Paper presented at the International Conference on Acoustics, Montreal, Canada.

Massenburg DesignWorks LLC. (2009). United States Patent No. 3586395 TradeMarkia: United States Patent and Trademark Office.

McGuire, S., & Pritts, R. (2008). *Audio sampling: A practical guide*. Burlington, MA: Focal Press.

Miśkiewicz, A. (1992). Timbre Solfege: A course in technical listening for sound engineers. *Journal of the Acoustical Society of America, 40*(7/8), 621-625.

Miśkiewicz, A., & Letowski, T. (1999). Psychoacoustics in the automotive industry. *Acustica, 85*, 646-649.

Miśkiewicz, A., & Letowski, T. (2014). *Timbre Solfege training in automotive industry*. Paper presented at the Forum Acusticum, Kraków, Poland.

Moulton, D. (1993). Golden Ears. Sherman Oaks, CA: KIQ Productions Inc.

Moulton, D. (1995). *Golden Ears*. Sherman Oaks, CA: KIQ Productions Inc.

Moulton, D. (2005). *Tutorial 14: Audio ear training, or learning to hear what you can hear!* Paper presented at the Audio Engineering Society Convention 119, New York, NY.

Moulton, D. (2014, 23rd March). [Notes about the Golden Ears audio ear training course].

Moylan, W. (2007). *Understanding and crafting the mix: The art of recording* (2nd ed.). Burlington, MA: Focal Press.

Neher, T., Rumsey, F. J., & Brookes, T. (2002, May 10-13). *Training of listeners for the evaluation of spatial sound reproduction.* Paper presented at the Audio Engineering Society Convention 112, Munich, Germany.

Nishimura, A. (2006a). WWW-based multilingualized system for technical listening training. *Journal of the Acoustical Society of America, 120*(5).

Nishimura, A. (2006b). WWW-based multilingualized system for technical listening training. Chiba, Japan: Nishimura, A.

Nishimura, A. (2013). *Ear training for non-technical students using a Web-based training system*. Paper presented at the Audio Engineering Society Conference: 50th International Conference: Audio Education, Murfreesboro, TN.

Nishimura, A. (2015). *Effects of ear training on education on sound quality of digital audio for non-technical undergraduates*. Paper presented at the Audio Engineering Society Convention 138, Warsaw, Poland.

Nishimura, A. (n.d.). Trial page of auditory training system. Retrieved from http://adlib.rsch.tuis.ac.jp/~akira/atsdemo/index_en.html

Quesnel, R. (1990). *A computer-assisted program in timbral ear training - a preliminary study.* (Master of Music), McGill University, Montréal, Canada.

Quesnel, R. (1994). *Computer-assisted training of timbre perception skills.* Paper presented at the International Computer Music Conference, Copenhagen, Denmark.

Quesnel, R. (1996). *Timbral ear trainer: Adaptive, interactive training of listening skills for evaluation of timbre differences*. Paper presented at the Audio Engineering Society Convention 100, Copenhagen, Denmark.

Quesnel, R. (2001). *A computer-assisted method for training and researching timbre memory and evaluation skills.* (Doctor of Philosophy), McGill University, Ottawa, Canada.

Quesnel, R., & Woszczyk, W. R. (1994). *A computer-aided system for timbral ear training*. Paper presented at the Audio Engineering Society Convention 96, Amsterdam, The Netherlands.

Rakowski, A., & Trybuła, A. (1975). *Auditory tests of timbre sensitivity at the tonmeisters faculty*. Paper presented at the Audio Engineering Society Convention 50, London, UK.

Rościszewska, T. (2011). *Rozwój wrażliwości barwowej słuchu u studentów reżyserii dźwięku*. Paper presented at the Audio Engineering Society International Symposium on Sound Engineering and Tonmeistering 14, Wrocław, Poland.

Rościszewska, T., & Miśkiewicz, A. (2014). *Timbre Solfege: Development of memory for timbre*. Paper presented at the Forum Acusticum, Kraków, Poland.

Rościszewska, T., & Miśkiewicz, A. (2015). *Timbre Solfege: Development of auditory cues for the identification of spectral characteristics of sound*. Paper presented at the Audio Engineering Society Convention 138, Warsaw, Poland.

Rose, J. (2008). *Audio postproduction for film and video: After-the-shoot solutions, professional techniques,and cookbook recipes to make your project sound better.* Burlington, MA: Focal Press.

SAE Institute. (2017). *Postgraduate Electives*. Australia: SAE Institute.

Schaller, T., & Burleigh, I. G. (2015). *Using web audio for ear training of sound engineers*. Paper presented at the Audio Engineering Society Conference: UK 26th Conference: Audio Education, Glasgow, Scotland.

Scheirman, D. (2013). *Are audio education programs keeping pace with new developments in industry?* Paper presented at the Audio Engineering Society Conference: 50th International Conference: Audio Education, Murfreesboro, TN.

Schmidt-Horning, S. (2013). *Chasing sound: Technology, culture, and the art of studio recording from Edison to the LP*. Baltimore, MA: The John Hopkins University Press.

Schoenburg, A. (1987). *Arnold Schoenburg: Letters* (E. Wilkins & E. Kaiser, Trans. E. Stein Ed.). Los Angeles, CA: University of California Press.

Shea, M. (2005). *Studio recording procedures*. New York, NY: McGraw-Hill.

Shively, R. E., & House, W. N. (1998). *Listener training and repeatability for automobiles*. Paper presented at the Audio Engineering Society Convention 104, Amsterdam, The Netherlands.

Stavrou, M. P. (2003). *Mixing with your mind: Closely guarded secrets of sound balance engineering*. Mosman, Australia: Flux Research Pty Ltd.

Strong, J. (2011). *Home recording for musicians for dummies* (4th ed.). Hoboken, NJ: John Wiley and Sons, Inc.

Strong, J. (2012). *Pro Tools all-in-one for dummies* (3rd ed.). Hoboken, NJ: John Wiley and Sons, Inc.

Theinhaus, E. (1960). The question of an audio engineering degree [letter to the editor]. *Journal of the Audio Engineering Society, 8*(1), 68-69.

Thompson, P., Mosley, B., & Ward, M. (2013). *Critical listening and acoustics as an essential part of the audio production and sound technology curriculum*. Paper presented at the Audio Engineering Society Conference: 50th International Conference: Audio Education, Murfreesboro, TN.

Tough, D. T. (2009). *Developing a consensus-driven, core competency model to shape future audio engineering technology curriculum: A web-based modified Delphi study.* (Doctor of Education), Tennessee State University, Nashville, TN.

Tough, D. T. (2010). *Shaping audio engineering curriculum: an expert panel's view of the future*. Paper presented at the Audio Engineering Society Convention 129, San Francisco, CA.

Trademark Direct. (2017). Tonmeister. Retrieved from http://uk.trademarkdirect.co.uk/tonmeister-2119081

University of Surrey. (2016). Music and Sound Recording BMus (Tonmeister) (Hons) University of Surrey - Guildford. Retrieved from http://www.surrey.ac.uk/undergraduate/music-and-sound-recording

Walzer, D. A. (2015). Critical listening assessment in undergraduate music technology programmes. *Journal of Music, Technology and Education, 8*(1), 41-53. doi:10.1386/jmte.8.1.41_1

Webster University. (2014). Technical ear training: Mastering the art of the invisible. Retrieved from http://www.webster.edu/news/2013/news/12042013-tet-audio.html

White, P., & Houghton, M. (2008). What's the frequency?: A guide to Effective EQ. *Sound on Sound.* Retrieved from http://www.soundonsound.com/techniques/whats-frequency

APPENDIX A – HUMAN RESEARCH ETHICS COMMITTEE APPROVALS

THE UNIVERSITY OF SYDNEY

RESEARCH INTEGRITY
Human Research Ethics Committee
Web: http://sydney.edu.au/ethics/
Email: ro.humanethics@sydney.edu.au

Address for all correspondence:
Level 6, Jane Foss Russell Building - G02
The University of Sydney
NSW 2006 AUSTRALIA

Ref: [SA/KFG]

4 May 2011

Associate Professor William Martens
Audio and Acoustics
Faculty of Architecture, Design & Planning
Wilkinson Building – G04
The University of Sydney
Email: william.martens@sydney.edu.au

Dear A/Prof Martens

Thank you for your correspondence dated 2 May 2011 addressing comments made to you by the Human Research Ethics Committee (HREC). On 4 May 2011 the Executive of the HREC considered this information and approved the protocol entitled "**Testing the relative influence of training method on tone colour discrimination".**

Details of the approval are as follows:

Protocol No.:	**05-2011 / 13570**
Approval Period:	**May 2011 to May 2012**
Authorised Personnel:	**Associate Professor William Martens** **Dr Densil Cabrera** **Mr Mark McKinnon-Bassett**
Documents Approved:	**Advertisement (version 1.2, 11/04/2011)** **Participant Information Statement (version 1.2, 11/04/2011)** **Participant Consent Form (version 1.2, 11/04/2011)**

The HREC is a fully constituted Ethics Committee in accordance with the National Statement on Ethical Conduct in Research Involving Humans-March 2007 under Section 5.1.29.

The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Research Involving Humans. A report on this research must be submitted every 12 months from the date of the approval or on completion of the project, whichever occurs first. Failure to submit reports will result in withdrawal of consent for the project to proceed. Your report is due by **31 May 2012**.

Chief Investigator / Supervisor's responsibilities to ensure that:

1. All serious and unexpected adverse events should be reported to the HREC within 72 hours for clinical trials/interventional research.

2. All unforeseen events that might affect continued ethical acceptability of the project should be reported to the HREC as soon as possible.

3. Any changes to the protocol must be approved by the HREC before the research project can proceed.

Manager Human Ethics
Dr Margaret Faedo
T: +61 2 8627 8176
E: margaret.faedo @sydney.edu.au

Human Ethics Secretariat:
Ms Karen Greer **T**: +61 2 8627 8171 **E: karen.greer@sydney.edu.au**
Ms Patricia Engelmann **T**: +61 2 8627 8172 **E: patricia.engelmann@sydney.edu.au**
Ms Kala Retnam **T**: +61 2 8627 8173 **E: kala.retnam@sydney.edu.au**

ABN 15 211 513 464
CRICOS 00026A

THE UNIVERSITY OF SYDNEY

4. All research participants are to be provided with a Participant Information Statement and Consent Form, unless otherwise agreed by the Committee. The following statement must appear on the bottom of the Participant Information Statement: *Any person with concerns or complaints about the conduct of a research study can contact the Manager, Human Ethics, University of Sydney on +61 2 8627 8176 (Telephone); + 61 2 8627 8177 (Facsimile) or ro.humanethics@sydney.edu.au (Email).*

5. You must retain copies of all signed Consent Forms and provide these to the HREC on request.

6. It is your responsibility to provide a copy of this letter to any internal/external granting agencies if requested.

7. The HREC approval is valid for four (4) years from the Approval Period stated in this letter. Investigators are requested to submit a progress report annually.

8. A report and a copy of any published material should be provided at the completion of the Project.

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

S. J. Assinder

Dr Stephen Assinder
Chair
Human Research Ethics Committee

cc: Mr Mark McKinnon-Bassett
mbas4365@uni.sydney.edu.au

Research Integrity
Human Research Ethics Committee

Friday, 24 May 2013

Assoc Prof William Martens
Architectural & Design Science; Faculty of Architecture, Design & Planning
Email: william.martens@sydney.edu.au

Dear William

I am pleased to inform you that the University of Sydney Human Research Ethics Committee (HREC) has approved your project entitled "**Transfer of training on tone colour identification to a dissimilarity-rating task.**".

Details of the approval are as follows:

Project No.: **2013/349**

Approval Date: **22/05/2013**

First Annual Report Due: **22/05/2013**

Authorised Personnel: **Martens William; McKinnon-Bassett Mark; Cabrera Densil**

Documents Approved:

Date Uploaded	Type	Document Name
12/05/2013	Participant Info Statement	Participant Information Statement
01/04/2013	Questionnaires/Surveys	Questionnaire given to participants upon entry
01/04/2013	Advertisements/Flyer	Advertisement of Study

HREC approval is valid for four (4) years from the approval date stated in this letter and is granted pending the following conditions being met:

Special Condition/s of Approval

- Please update the version number on the fiurst page of the Participant Information Statement

Condition/s of Approval

- Continuing compliance with the National Statement on Ethical Conduct in Research Involving Humans.
- Provision of an annual report on this research to the Human Research Ethics Committee from the approval date and at the completion of the study. Failure to submit reports will result in withdrawal of ethics approval for the project.
- All serious and unexpected adverse events should be reported to the HREC within 72 hours.

Research Integrity
Research Portfolio
Level 2, Margaret Telfer
The University of Sydney
NSW 2006 Australia

T +61 2 8627 8111
F +61 2 8627 8177
E ro.humanethics@sydney.edu.au
sydney.edu.au

ABN 15 211 513 464
CRICOS 00026A

THE UNIVERSITY OF SYDNEY

- All unforeseen events that might affect continued ethical acceptability of the project should be reported to the HREC as soon as possible.

- Any changes to the project including changes to research personnel must be approved by the HREC before the research project can proceed.

Chief Investigator / Supervisor's responsibilities:

1. You must retain copies of all signed Consent Forms (if applicable) and provide these to the HREC on request.

2. It is your responsibility to provide a copy of this letter to any internal/external granting agencies if requested.

Please do not hesitate to contact Research Integrity (Human Ethics) should you require further information or clarification.

Yours sincerely

S. J. Assinder

Dr Stephen Assinder
Chair
Human Research Ethics Committee

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) National Statement on Ethical Conduct in Human Research (2007), NHMRC and Universities Australia Australian Code for the Responsible Conduct of Research (2007) and the CPMP/ICH Note for Guidance on Good Clinical Practice.