TOWARDS A NEW EQUALISATION USER INTERFACE FOR MIXING MULTI-TRACK RECORDINGS OF MUSICAL INSTRUMENTS

Christopher Paul Dewey

A thesis submitted to the University of Huddersfield
in partial fulfilment of the requirements for
the degree of Masters of Science

The University of Huddersfield

January 2014

Abstract

Mixing consoles are the primary tool used in the mixing of multi-track recordings of musical instruments and feature an equalisation (EQ) section to manipulate the spectral characteristics of each channel. Despite numerous technological advances, modern embodiments of user interfaces for mixing continue to predominantly follow the traditional linear channel strip paradigm. The aim of this research is to question and reconsider one aspect of the mixing desk user interface, namely, the EQ user interface. The research methodology involved evaluating example mixing consoles, their EQ sections and considering a range novel mixing and EQ interfaces. A literature review of Human Computer Interface (HCI) fundamentals established current design and evaluation approaches. This is supplemented by case studies of the design and evaluation of several user interfaces for music technology applications. In line with HCI theory an analysis of the EQ task was conducted and an expert user consulted throughout the design process, simulating a range of scenarios to refine and develop paper prototypes into workable designs. A corrective EQ was developed that was favoured in evaluation over a typical existing Digital Audio Workstation (DAW) EQ. The novel features of the corrective EQ user interface were spectral data visualised as a static spectral plot with five peak frequencies presented to the user allowing the user to directly manipulate the spectral plot which simultaneously provided real-time visual feedback. A multi-track EQ was developed that was favoured in evaluation over an existing DAW EQ for multi-track EQ tasks. The multi-track EQ user interface featured two tracks displayed simultaneously and added novel direct manipulation method for performing mirrored EQ in the form of ghost nodes representing peaks from the other track on a track's spectral plot. A similar visual and direct manipulation method for performing band pass and low and high shelf filtering was also included.

Table of Contents

Copyright Statement 2

Abstract 3

List of Figures 7

1 Introduction 10

1.1 Background 10

1.2 Research Focus and Methodology 11

1.3 Structure of Thesis 13

2 Mixing Interfaces 14

2.1 Introduction 14

2.2 Analogue Mixing Consoles 14

2.2.1 Allen and Heath ZED R16 14

2.2.2 Observations 16

2.3 Digital Mixing Consoles 16

2.3.1 Yamaha O2R96 17

2.3.2 Observations 18

2.4 Digital Audio Workstation 18

2.4.1 Observations 19

2.4.2 Line 6 StageScape M20d 20

2.5 Novel Interfaces 22

3 EQ Interfaces 25

3.1 Introduction 25

3.2 Filter Types 25

3.3 Analogue Console EQ Interfaces 26

3.4 Digital Console EQ Interfaces 29

3.5 Digital Audio Workstation EQ Interfaces 31

3.6 Line 6 StageScape M20d EQ Interface 32

3.7 Novel EQ Interfaces 33

3.8 Audio Analysis Interfaces 36

4 Interface Design: HCI Fundamentals 40

4.1 Introduction 40

4.2 Principles 40

4.2.1 Identification of Users and Tasks 40

4.2.2 Selection of appropriate interaction style and visualisation metaphor 41

4.2.3 Seven Stages of Action 45
4.2.4 Automation 47
4.3 Evaluation 47
4.3.1 Fitts Law 48
4.3.2 Usability 49
5 Case Studies 52
5.1 Cuebert: A New Mixing Board Concept for Musical Theatre 52
5.2 Knobs vs Faders 53
5.3 Multitouch Interface for Audio Mixing 53
5.4 Towards a more Flexible and Creative Music Mixing Interface 54
5.5 Battle of the DJs 55
5.6 Summary 56
6 Dissection of the EQ Task 57
6.1 Introduction 57
6.2 Corrective EQ 57
6.3 Creative EQ 58
7 Simple Corrective EQ Interface 59
7.1 Design 59
7.2 Implementation 63
7.2.1 Capturing the spectral information 63
7.2.2 Maximum amplitude vs average amplitude 65
7.2.3 Peak detection 65
7.2.4 The EQ algorithm 66
7.3 Test 68
7.4 Analysis of Results 68
7.4.1 Effectiveness and efficiency 68
7.4.2 Satisfaction 72
7.4.3 User comments and observations 74
7.5 Conclusions 74
7.6 Limitations of the experiment 75
8 EQ User Interface for Mixing Multi-track Recordings of Musical Instruments 76
8.1 Introduction 76
8.2 Design 76
8.2.1 Interfaces 76

8.2.2 Controls 79
8.3 Implementation 81
8.4 Test 83
8.5 Analysis of results 85
8.5.1 Effectiveness and efficiency 85
8.5.2 Satisfaction 88
8.6 Conclusions 92
9 Summary, Conclusions and Further Work 94
9.1 Summary 94
9.2 Conclusions 96
9.3 Further Work 98
10 References 99

Word count 25,267

List of Figures

Figure 1: Layout of the Allen and Heath ZED R16 Analogue Mixing Console (Copyright Allen & Heath, 2008) 14
Figure 2: Allen and Heath ZED R16 Main meter section 16
Figure 3: Yamaha O2R96 digital mixing console (Copyright Yamaha) 17
Figure 4: Layers on the Yamaha O2R96 17
Figure 5: Cubase 7 Mixing Interface (Steinberg, 2013) 18
Figure 6: Cubase pan control 19
Figure 7: a) Line 6 StageScape M20d console b) Line 6 StageScape main interface (Copyright Line 6, 2012) 20
Figure 8: Mapping volume, frequency and panning visually (Gibson, 2008) 22
Figure 9: An example of a mix (Gibson, 2008) 22
Figure 10: Multitouch interface for mixing audio (after Carrascal & Jorda 2011) 23
Figure 11: a) Personalisation of song playback b) Interface for simple multi-track mixing (Scott *et al* 2013) 24
Figure 12: High pass and low pass filter properties and response curves 25
Figure 13: High shelving filter properties and response curve 25
Figure 14: Peaking filter properties and response curves 26
Figure 15: Allen and Heath ZED R16 mono channel EQ section and corresponding response curves indicating each filter's range (Allen & Heath, 2008) 27
Figure 16: Allen and Heath ZED R16 stereo channel EQ section (Allen & Heath, 2008) 28
Figure 17: Samson S-Curve Dual 231 (copyright Samson, 2013) 28
Figure 18: Response curves for the Samson S Curve Dual 231 (copyright Samson, 2003) 29
Figure 19: Yamaha O2R96 Selected Channel Section with EQ section and core controls highlighted (copyright Yamaha, 2010) 30
Figure 20: Yamaha O2R96 display EQ screen (Copyright Yamaha, 2004) 30
Figure 21: Cubase 7 EQ (Steinberg, 2013) 31
Figure 22: Line 6 StageScape M20d EQ quick tweak interface (Copyright Sound on Sound, 2012) 32
Figure 23: Line 6 StageScape M20d EQ deep tweak interface (Copyright Sound on Sound, 2012) 33
Figure 24: The subjEQt interface (Mecklenburg & Loviscach, 2006) 33
Figure 25: Frequency ranges of subjective terms (Mecklenburg & Loviscach, 2006) 34
Figure 26: The 2DEQ interface (Sabin & Pardo, 2009) 35
Figure 27: The visual mixer EQ interface (Gibson, 2008) 35
Figure 28: Adobe Audition CS5 spectrogram view 36
Figure 29: Ozone 5 Waterfall Plot (Izotope, 2013) 37
Figure 30: Novel interface for displaying the spectral content and phase differences of a stereo channel (Eastty, 2000) 37
Figure 31: Colour mapping of phase (Eastty, 2000) 38
Figure 32: Figure 32: Novel interface for displaying volume envelope and pitch (Schnell, Cifuentes & Lambert, 2010) 38
Figure 33: The three pillars of successful user-interface development (after Shneiderman & Plaisant, 2005) 40
Figure 34: The transformation from numbers to insight 41
Figure 35: Image invariants and components for a frequency response curve 42
Figure 36: Planar dimensions and retinal variables (after Bertin, 1983) 42

Figure 37: Comparing differences in dimensions used to visualize data in Gibson's (2008) and Carrascal & Jorda's (2011) mixing interfaces 43
Figure 38: The different types of organisation (after Bertin, 1983) 43
Figure 39: Level of organisation (Bertin, 1983).. 44
Figure 40: Three types of time series data visualisations (after Heer et al, 2010) 45
Figure 41: Seven stages of action (Norman, 2002) 45
Figure 42: Design questions derived from the seven stages of action model (Norman, 2002 p. 53) .. 46
Figure 43: Fitts Law 48
Figure 44: The range of interfaces evaluated by Gelineck & Serafin (2009) 53
Figure 45: Range of methods employed to enable user interaction with the prototype mixing interface. Screenshots taken from a video of the prototype interface in use 55
Figure 46: Interface 1 59
Figure 47: Interface 2 60
Figure 48: Interface 3 60
Figure 49: Interface 4 61
Figure 50: Interface 5 61
Figure 51: Interface 6 62
Figure 52: Interface 7 62
Figure 53: FFT Time to frequency domain 63
Figure 54: Capturing the spectral information flow diagram 64
Figure 55: Identifying the bin maximums 65
Figure 56: Flow diagram of the EQ algorithm employed in interfaces 1 to 6..................................... 66
Figure 57: Flow diagram of the EQ algorithm employed in interface 7... 67
Figure 58: Normalised task completion time (NTCT) for each user and interface 69
Figure 59: Average normalised task completion time (NTCT) for each interface with 95% confidence intervals.. 69
Figure 60: Timelines of test subject thinking and activity 71
Figure 61: Number of periods of thinking followed by activity 72
Figure 62: Plot of average number of periods of thinking followed by activity 72
Figure 63: User preference ranking of interfaces 73
Figure 64: Average user preference ranking of interfaces 73
Figure 65: Interface A.. 77
Figure 66: Interface B.. 77
Figure 67: Interface C.. 78
Figure 68: Interface C logarithmic implementation with two identical channels displayed 78
Figure 69: Interface D 79
Figure 70: Multi-track EQ controls... 80
Figure 71: Flow diagram of the algorithm used in interfaces B, C and D 82
Figure 72: NTCT for all test subjects and scenarios 85
Figure 73: Average normalised task completion time (NTCT) with 95% confidence intervals for all interfaces and scenarios considered 86
Figure 74: Plot of average normalised task completion time (NTCT) for each interface and scenario considered.. 86
Figure 75: Pairwise comparison between interfaces from a post hoc analysis with statistically significant relationships highlighted. 88

Figure 76: Average preference scores for all subjects and scenarios 88
Figure 77: Percentage of positive and negative keywords selected for each interface 89
Figure 78: Wordclouds generated for each interface based on the keywords selected by the test subjects 90
Figure 79: Top five 5 keywords for each interface based on preference 91

1 Introduction

1.1 Background

Music technologists use a variety of specialist tools to produce professional recordings. After the initial recording session the user will commonly use a mixing console to blend and sculpt the multi-track recordings of musical instruments into a cohesive mix. This makes the mixing user interface the key interface in any studio environment.

Mixing consoles and their software equivalent offer a suite of tools that enable the user to achieve their goal. Generally, these tools comprise controls to set the volume of an individual track, controls that pan the individual tracks in the stereo field, equalisers to control the sonic characteristics of each track, audio buses and routing options to send audio to and from other pieces of equipment and metering to provide the user with visual feedback.

The technology behind the mixing console was developed after the Second World War and was first developed for the broadcasting industry (Schmidt Horning, 2013). It became widely used in music production in the late 1950s with the advent of equipment such as the Studer Dynavox that could record multiple individual audio tracks on to magnetic tape. Early mixing consoles were custom built for recording studios by electrical engineers using valves and electrical circuitry to meet the individual design specifications (Studer, 2012). The invention of discrete transistors in the early 1960s enabled engineers to create more compact analogue solutions with rotary knobs and linear faders grouped into vertical strips for each channel of audio. Examples of mixing consoles from this era include the Studer 069 designed in 1968 (Studer, 2012) and Soundcraft's Series 1 released in 1973 (Soundcraft, 2001). Each vertical strip contained rotary knob controls arranged in visually recognisable sections mirroring the physical path of the signal. Characteristically they also featured a volume fader at the bottom of each channel strip. Additionally buttons are included to provide functions such as attenuation, muting and soloing of channels and to determine routing options for each channel to audio buses. This layout has remained largely consistent since the 1970s and can be found in the vast majority of modern analogue mixing consoles available today e.g. Neve's flagship analogue console, the 88RS (AMS-Neve, 2013).

Digital mixing consoles were developed in the early 1980s and offered a greater number of channels and extended the range of mixing tools provided by including gates, compressors and effects. The layout of the digital console commonly mimics their analogue forerunners with channels presented as vertical strips. One notable addition to the digital console is the inclusion of an LCD screen that enables the user to access and view parameter values and provides additional graphical visualisations of the mixing tools.

As increased computer processing power and storage has become more affordable in the last two decades many studios now use a digital audio workstation (DAW) with specialist software that enables the user to record and mix multi-track recordings of musical instruments on a standard desktop computer (Huber & Runstein, 2014). Channels commonly continue to be visualised as vertical strips with controls accessed through windows that can be open and closed as required (Nash & Blackwell, 2011). Often a physical control surface is included in more sophisticated setups to enable the user to control the mixing tools featured in the DAW with the familiar vertical strips of rotary encoders, faders and buttons.

In the early days of modern music production scientists and electrical engineers were employed to operate the custom built mixing consoles under the guidance of the producer (Schmidt Horning, 2013). As mixing consoles became used more widely specially trained sound engineers and producers found themselves in charge of the mixing console using their knowledge of music technology fundamentals and creativity to realise artistic intentions. Today, an incarnation of this equipment is as likely to be found in a "bedroom-producer's" home as it is in a large professional recording studio. This transferral of power from the record company to the individual is referred to as the democratisation of music technology and is a part of a wider societal and economic trend brought about by the development and widespread adoption of software (Leyshon, 2008).

1.2 Research Focus and Methodology

Mixing multi-track recordings of musical instruments to a professional standard is not trivial. This is in part due to the technical and abstract nature of audio fundamentals and much like any other skilled practical discipline, has traditionally been learned through experience (Mixerman, 2010). Each mixing scenario encountered by the user will be different and if moving between musical genres may vary significantly. Although experience can help guide a user different mixes often cannot be created in exactly the same way simply by following a preset formula. For these reasons creating the perfect mix is often seen as a dark art by many novice users.

Following observations made by Reiss in a lecture in 2011, we can contrast the evolution of the camera with that of the mixing console and we see that cameras have evolved much further in a similar space of time. Early cameras were analogue in nature requiring an in-depth knowledge to create professional photographs. The emergence of digital technology saw the addition of extra functionality to enhance the user performance. Assistive features such as auto-focus, exposure control and red-eye reduction have been commonplace for the past two decades, with modern cameras on mobile phones now enabling the user to take high quality photographs with the click of a button on a device that barely resembles the early camera. In contrast today's common embodiment of the mixing console is very similar in appearance to those first introduced in the 1960s with product developers seemingly focussing on increasing the number of channels and extending the palate of tools provided. There continues to belittle in the way of assistive features to help the user create better multi-track mixes. This arguably limited evolution leads one to question whether the established mixing console paradigm really offers the best interface, controls and information to the user for mixing multi-track recordings of musical instruments and provides the motivation behind this research.

The related discipline of human computer interaction (HCI) first emerged in the 1980s and is concerned with the interaction between humans and computers. The discipline not only critically evaluates good design examples, extracting key factors for consideration, but also provides a framework for developing and testing the usability of an interface. Designs that exhibit good HCI are preferred by users and assist users in successfully performing computer tasks to meet desired results (Norman, 2005).

An evaluation and redesign of the whole mixing console is beyond the scope of a project of this size. Consequently, the EQ section of the console was selected as a focus for reconsideration because it is the most widely used signal processing tool in audio production (Sabin & Pardo, 2009) and is a key component in the mixing of recorded instruments. Again due to time constraints, this initial investigation was restricted to mixing multiple mono multi-track recordings of musical instruments to a mono mix.

The redesign was approached in two stages. Firstly, following HCI best practice, several novel user interfaces were developed for a simple corrective EQ task. These interfaces were compared against a range of traditional user interfaces to determine whether they improved user performance for the task under consideration. The findings from the first stage of this investigation were then used to steer and inform the development of a more sophisticated EQ interface. This interface featured controls that enable the user to employ corrective and creative EQ techniques to mix two tracks of recorded instruments together. This was intended to provide a sufficient proof of concept for a new mixing interface.

This research is of significant importance because a reconsideration of the traditional mixing console paradigm is, arguably, long overdue. Much insight can be gained from advances in the field of HCI with modern computer graphical, processing and interaction capabilities offering possibilities that currently remain largely under-utilised in recording studio equipment. In addition this research aims to offer a more informed perspective on audio user interface design and evaluation.

1.3 Structure of Thesis

Chapter Two presents and evaluates the strengths and weaknesses of traditional and novel mixing user interfaces through the study of selected example implementations. Chapter Three presents and evaluates a range of traditional and novel EQ user interfaces. An overview of HCI fundamentals for the design and evaluation of assistive interfaces is provided in Chapter Four. Chapter Five presents a number of relevant user interface design and evaluation case studies for audio/music applications. A dissection of the EQ task is detailed in Chapter Six. Chapter Seven outlines the development, implementation and evaluation of a range of novel interfaces for the specific task of corrective EQ. Chapter Eight details the development, implementation and evaluation of several novel EQ user interfaces for mixing multi-track recordings of musical instruments. Chapter Nine presents a summary, overall conclusions and outlines further work.

2 Mixing Interfaces

2.1 Introduction

Professional mixing consoles are available in three formats, analogue, digital and computer based. All three incarnations offer similar core controls for setting the volume or level, pan position and equalisation (EQ) of each channel. Controls for muting, soloing and routing channels are also provided as well as a means of metering the level of individual channels and the master output.

The following section will begin by reviewing an existing commercial example from each of the three formats of mixing user interface. This will also provide context for when we move onto considering the EQ user interface in detail. One example of each type of user interface has been chosen to illustrate their typical user interface characteristics and critique their suitability as a user interface for performing their intended task.

2.2 Analogue Mixing Consoles

Analogue mixing consoles commonly present individual channels as vertical strips with controls divided into visually identifiable sections dependant on purpose (Izhaki, 2008). The layout of these controls is governed by the arrangement of the electronic components used in the manufacture of the console because the buttons, sockets, knobs and faders presented to the user are commonly soldered directly on to the underlying printed circuit board (PCB). This provides the user with an indication of how the signal flows through the console. Analogue mixing consoles commonly use buses to route signals to different parts of the mixing console. A bus is "a common electrical signal path along which signals may travel" (White, 1997, p. 288).

2.2.1 Allen and Heath ZED R16

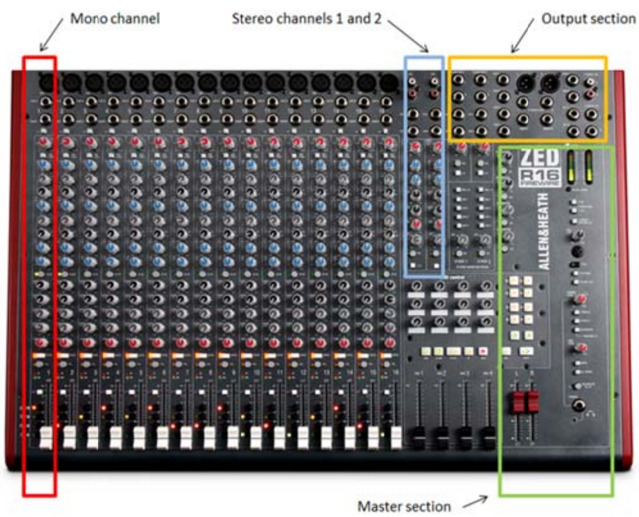

Figure 1: Layout of the Allen and Heath ZED R16 Analogue Mixing Console (Copyright Allen & Heath, 2008)

The Allen and Heath ZED R16 shown in Figure 1 is an example of a modern analogue mixing console (Allen & Heath, 2008). The ZED R16 features sixteen mono input channels, four stereo input channels, an output section and a master section.

Each mono input channel features several distinct sections. The section at the top of the channel features inputs for microphones and line level sources and an insert socket. A button is also provided to turn phantom power on and off. The next section features a knob to control the gain of the input amplifier and is used to control the signal level of the input. The section beneath these controls is concerned with EQ and will be covered in detail in Chapter Three. Below the EQ section is an auxiliary section used to send to the channel's signal to the auxiliary output section of the console. Four knobs are provided to route the signal to four different auxiliary buses. The next section features a knob to position the mono channel from left to right in the stereo field and is called the pan control. A rectangular button and red light emitting diode (LED) is placed below the pan control to mute the channel's signal. When the button is depressed the signal is muted and the LED is illuminated. A circular button and yellow LED is provided under the mute control to route the channel to the pre-fade listen (PFL) bus. Again, the LED is illuminated when the button is depressed. Two smaller LEDs are positioned beneath the PFL switch. The first LED is green and illuminates dimly at a threshold of –14dB getting brighter with higher signal levels. The second LED is red and illuminates when the signal is within 5dB of clipping. At the bottom of the channel strip is a fader to control the level of the channel sent to the main stereo bus. To the right of the channel fader are five buttons and coloured LEDs. These buttons are used to route the channel's signal to the main stereo bus and four group buses.

The four stereo input channels on the ZED R16 are positioned to the right of the mono channels and in comparison feature fewer controls. Stereo channels 1 and 2 feature more controls than channels 3 and 4 and are divided into the following sections:

- Inputs for unbalanced stereo inputs
- A knob to adjust the input level
- A simplified stereo EQ consisting of two knobs. These controls will be discussed in more detail in Chapter Three
- Two knobs to route the signal to auxiliary buses 1 and 2
- A stereo balance knob to adjust the relative level of the left and right signal
- A knob to control the overall stereo channel level
- A circular switch and yellow LED to PFL the channel in stereo
- A rectangular button to route the channel to the main stereo bus

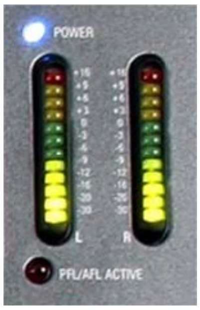

Figure 2: Allen and Heath ZED R16 Main meter section

The master section is positioned to the right of the console as indicated in Figure 1. Most importantly this includes two master faders to control the overall output level and a visual meter shown in Figure 2. An LED is provided to indicate when the PFL bus is active. Several, sockets, buttons and knobs are provided at the top of the section to facilitate a variety of routing options.

2.2.2 Observations

The main advantage of the traditional analogue mixing console is that the system state for all input and output channels is visible to the user in a single interface. Since every control performs an individual function the user can readily see all control settings simultaneously. Izhaki (2008) terms this advantage "what-you-see-is-what-you-get" (Izhaki, 2008, p. 137).

Intrinsically, the relationships between input channels can easily be deduced because each channel strip is presented identically and functionality is aligned horizontally across the console. For example, the user can view all pan settings for each channel by looking across the same horizontal row of knobs. Furthermore, as each audio channel is always located in the same place, it makes it easy for the user to quickly find and control settings (Gelineck, Büchert & Andersen, 2013).

The main disadvantage of analogue mixing consoles is that the size of the console is determined by the number of channels and controls provided. In order to make a smaller console either the number of channels featured or the range of controls provided must be compromised.

2.3 Digital Mixing Consoles

Digital mixing consoles fundamentally perform the same function as their analogue counterparts however their implementation and user interface are different. Analogue input signals are converted to digital as they enter the console and signals are converted back to analogue before they are output. Audio processing is performed by digital signal processors. Central to the console design is an in-built computer that controls the user interface, directs the processing performed by the digital signal processors and controls the audio routing. Importantly, Izhaki (2008) observes that "the individual controls on the console surface are not part of any physical signal path – they only interact with the internal computer reporting their state" (Izhaki, 2008, pp. 135-136).

2.3.1 Yamaha O2R96

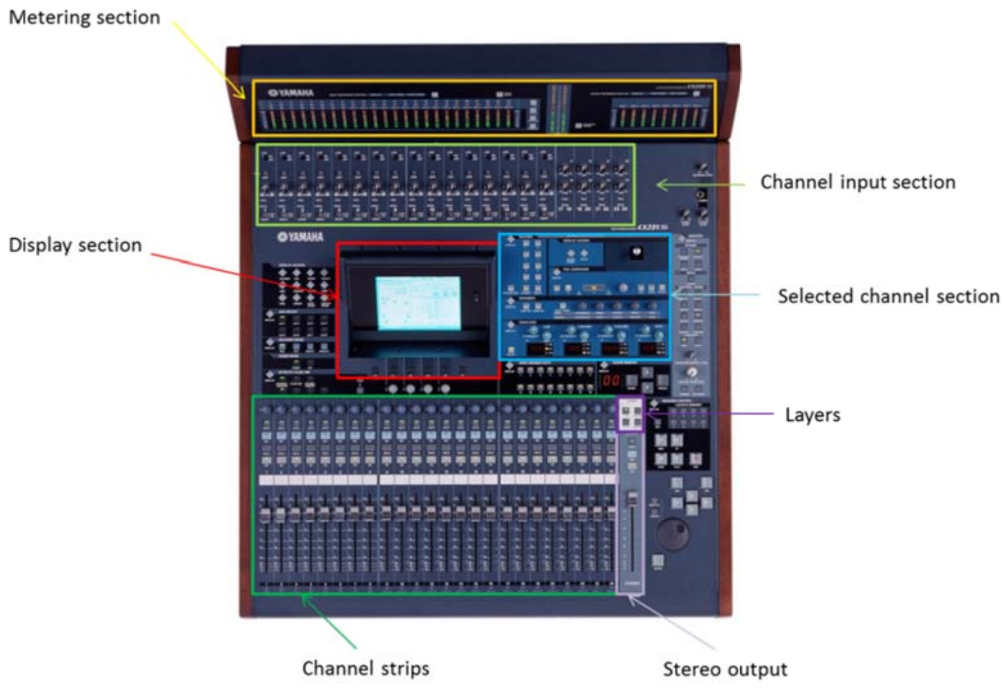

Figure 3: Yamaha O2R96 digital mixing console (Copyright Yamaha)

The Yamaha O2R96, displayed in Figure 3 is one example of a digital mixing console. This embodiment features fifty-six input channels, eight auxiliary outputs and eight bus outputs. Augmenting the core mixing tools the O2R96 also features gates and compressors for each input channel and a suite of effects (Yamaha, 2010).

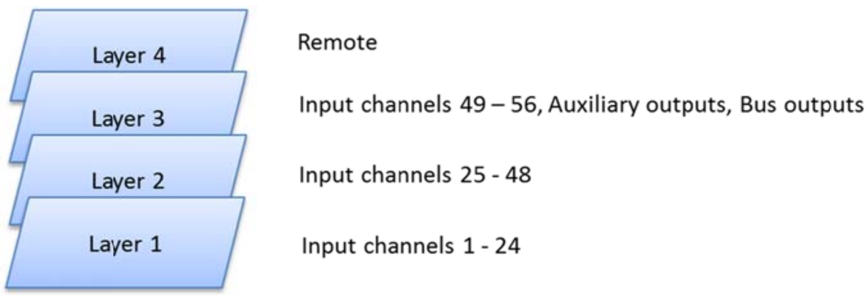

Figure 4: Layers on the Yamaha O2R96

Central to the design and operation of the O2R96 is the use of layers as illustrated in Figure 5. These layers effectively divide the input and output channels into four sets to allow twenty-four channels to be controlled via the channel strip section at any one time, dependant on the layer selected.

The channel strips are arranged as vertical rows situated along the bottom of the console interface. Each channel strip contains one rotary encoder, four rectangular buttons and a channel fader. The rotary encoder replaces the columns of knobs commonly featured on analogue mixing desks and are used to control pan position, auxiliary sends and two further assignable modes. The function controlled by the encoder is determined by buttons to the left of the display.. Two buttons have the set purpose of controlling PFL and muting the channel output. To provide the user with feedback each button is illuminated with a colour dependant on the mode selected.

The selected channel section controls a range of functions via a suite of rotary encoders, buttons and a joystick. A 320 x 240 pixel monochrome illuminated LCD panel is featured in the display section to provide the user with visual feedback through a series of pages. Display pages are grouped by function and can be selected using a range of buttons beneath the LCD panel. These sections are covered in greater detail in Chapter Three with specific regard to EQ.

2.3.2 Observations

Digital mixing consoles offer several improvements over analogue consoles. Significantly more channels can be featured on a smaller footprint through the use of layers. Digital mixing consoles can offer an all-in-one mixing solution as dynamic processing, effects and channel automation are included in many products. The inclusion of display screens is another distinct advantage providing the user with greater visual feedback.

The main disadvantage of digital consoles is that the controls, parameter values and visual graphical displays are effectively hidden in various screens and only one screen can be seen at once. Additionally there is often no longer a one to one mapping of physical controls for pan, auxillary busses and the EQ section. Interaction with the console is required before being able to set and review these controls. For this reason manufacturers such as NEVE build digital consoles with as many controls as their analogue counterparts (Izahki, 2008). Furthermore, it is not possible for the user to easily compare the control settings of multiple channels with the exception of fader level.

2.4 Digital Audio Workstation

Figure 5: Cubase 7 Mixing Interface (Steinberg, 2013)

DAWs combine the tools found in multi-track recorders and mixing consoles in one software application (Izhaki, 2008). One example of a modern DAW is Cubase 7 (Steinberg, 2013) which features two key interfaces; the multi-track sequencer interface that displays tracks horizontally and the mixing interface that displays mixer strips associated with the tracks vertically (see Figure 5)

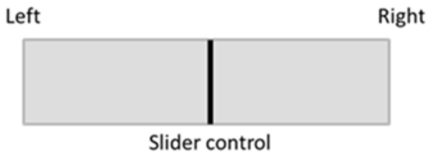

Figure 6: Cubase pan control

The mixing interface provides manifestations of the controls commonly found in hardware mixing consoles. The pan section is visually very different the controls commonly found in hardware consoles and features a small horizontal fader-type control above the level fader to place recorded instruments in the stereo field. (see Figure 6). Furthermore, each strip features a graphical visualisation of a meter placed next to the volume fader. Each meter can be configured to change the display type. A suite of four buttons is presented above and below the fader to control a variety of tools. Once activated, each button changes colour to provide feedback. The section above the fader features buttons to select mute, solo and listen mode and a button that opens the EQ section window, discussed in Chapter Three. The button section beneath the fader is visually identical to the section above and features controls for automation and routing. A graphical icon of an instrument can be used to help label and further identify each channel. The upper section of the mixing channel features a “channel rack” section. Visually this is presented as a series of horizontal bars/faders which are similar in appearance to the pan control. This section can be used to control a wide variety of mixing tools including gain, auxillary bus routing, inserts and EQ (discussed in Chapter Three). These tools are selected from a list contained within a popup menu window.

2.4.1 Observations

The DAW mixer interface largely adheres to the established analogue vertical strip paradigm and importantly provides a visual indication of the key control settings for each channel in a single interface. Benefits provided by this user interface include, a degree of channel configuration to make only the most relevant details explicit whilst still providing a large array of extended mixing tools. Additionally channels can be coloured, named and placed next to related channels to assist user interaction. The user is also often able to change the channel’s width to enable more channels to be displayed at one time although this is ultimately limited by the size of the display. For these reasons the DAW mixer interface overcomes the main drawback of digital consoles which effectively hide channel controls from the user.

Nash & Blackwell (2011) conducted an investigation into how users interact with DAWs. They discovered that users spend a large proportion of their time navigating the interface and opening, closing and resizing windows as opposed to actively using the tools provided. This is not good for efficient user interaction. Furthermore, users were observed spending more time fine-tuning individual channel settings than considering the relationships between these channels. This indicates that a balance needs to be found between providing too much information and enhancing the interface's visualisations and considering the users needs.

2.4.2 Line 6 StageScape M20d

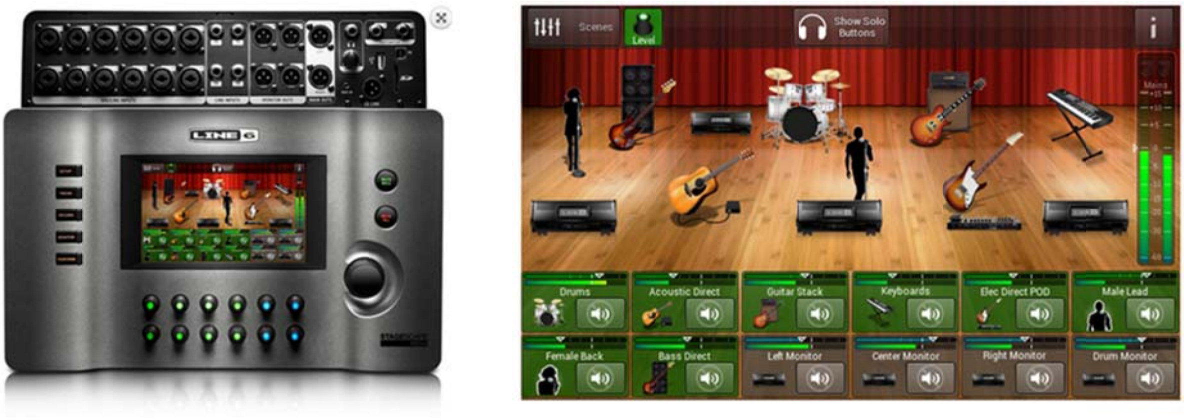

Figure 7: a) Line 6 StageScape M20d console b) Line 6 StageScape main interface (Copyright Line 6, 2012)

The Line 6 StageScape M20d (see Figure 7a) represents a commercial product that attempts to simplify the digital mixing console interface and is heralded by the manufacturers as the "world's first smart mixing system for live sound" (Line 6, 2012, para. 1). During an interview, Marcus Ryle from Line 6 explains the motivation behind the design by posing the questions "why is it that every mixer channel is identical, regardless of what you plug into it? Why can't you communicate with a console in the same way you would a sound engineer?" (White, 2012, para. 16).

The Stagescape interface is built around a 7-inch colour LED touch screen and displays sound sources as psuedo-realistic graphical icons on top of a visual representation of a stage as shown in Figure 7b. Users can interact with the graphical icons to access channel controls (Line 6, 2012).

Crucial to the design of the StageScape is the implementation of five modes which direct the workflow of the user. These modes are accessed via five rectangular buttons aligned vertically to the left of the screen. Twelve illuminated multifunction encoders are provided along the bottom of the display. These encoders work in conjunction with the LED touch screen to control various tools depending on the mode selected. The encoders change colour and brightness depending on their function.

The first mode is setup. On plugging in an instrument or microphone the user is prompted to select the graphical icon that represents the musical instrument. This selection is important as it identifies and sets the dynamics, EQ and effects commonly applied to this type of musical instrument in a live environment. The second mode is tweak which enables the user to adjust and fine tune the settings applied to the channel in setup mode. The tweak mode is covered in greater detail with regards to EQ in Chapter Three. The Third mode is monitor mode and used to control the monitor mix on stage. The Fourth mode is record which enables the user to record up to sixteen individual channels and the main mix to an internal hard drive. The fifth mode is perform which disables various controls to prevent accidental changes during performance.

Three further controls are provided to the right of the console including a large master volume rotary encoder and two buttons to mute all microphones or channels.

Line 6's Stagescape has received some praise since first launched in 2012 (White, 2012), however reviewers commonly criticise the lack of faders believing the implementation of rotary encoders and absence of faders a limitation (Price, 2013). Line 6 have clearly tried to address this criticism by adding a traditional fader screen to the suite of controls in its most recent software update.

This console is a significant development in seeking to simplify the layout of the digital mixing console and providing a more intuitive visual interface. Furthermore, the designers have considered the needs of the user with regards to workflow seeking to display controls based on the context of use. Undoubtedly, the Stagescape is more accessible to novice users in comparison with traditional interfaces because a workflow is designated and presets are applied to channels by default. With regards to more experienced users however, they may feel constrained and limited by such features. In comparison to the analogue mixing console it is not possible to easily see the settings of all channels with the Stagescape M20d as the user must navigate through the range of modes and displays to make adjustments. White's (2012) review for Sound on Sound magazine possibly offers a poignant summation.

> Whether this format of mixer takes off depends very much on whether musicians and engineers are prepared to accept a departure from the traditional 'row of faders' format, but only a few minutes spent with the M20D should be enough to convince people that its operation is, indeed, intuitive. (White, 2012, para. 36)

2.5 Novel Interfaces

Several interfaces that depart from the established mixing console paradigm have been proposed in academic literature. Gelineck *et al* (2013) argue that the first attempt to really present an alternative to the traditional mixing console paradigm is the Virtual Mixer proposed by Gibson (1997).

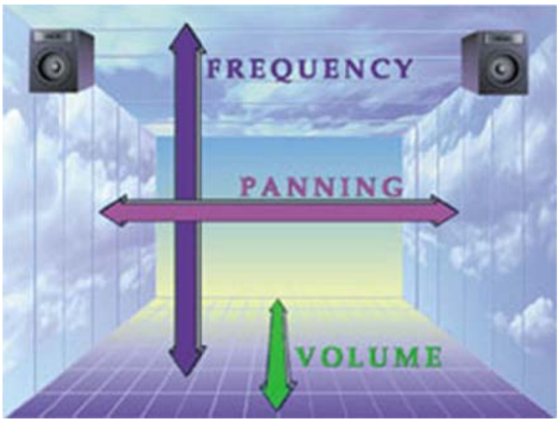

Figure 8: Mapping volume, frequency and panning visually (Gibson, 2008)

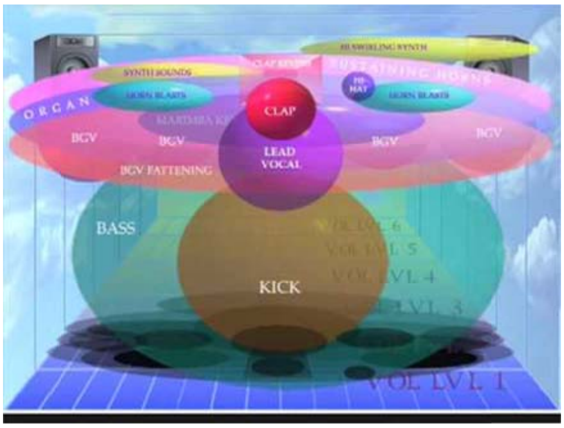

Figure 9: An example of a mix (Gibson, 2008)

At its core, Gibson's interface mapped volume, frequency and panning to create a virtual mix space shown in Figure 8. Audio channels are represented as coloured spheres placed within the virtual mix space as shown in Figure 9. The sphere's position left to right represented the channel's pan and position from front to back represents the channel's level. The sphere's height position was determined by the spectral content of the instrument's signal with higher pitched instruments positioned nearer to top of the interface. Gibson (2008) provided an esoteric explanation of why frequency content is mapped from low to high declaring "there are energy centres in the body called chakras that respond to different frequencies. These frequencies are specifically mapped out from low to high, from the base of the spine to the top of the head" (Gibson, 2008, p. 25).

The main advantage of Gibson's visual mixer is that all channels are displayed simultaneously in a visually impressive display. However the interface appears to have greater value as a visualisation of audio rather than a practical mixing tool for two main reasons. Firstly, the interface is very “busy” with channels of similar spectral content and pan position overlapping one another. Secondly and most crucially, the display changes in real time with spheres fading in and out and changing position during playback. This is very confusing and it is hard to see how the user could accurately select an instrument and modify a channel’s settings given such movement. The sphere size could be misleading to users who may interpret larger spheres being louder than smaller spheres.

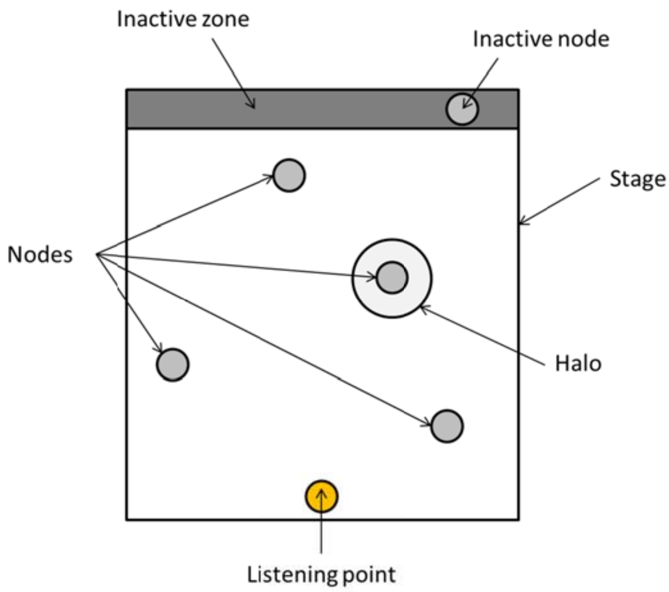

Figure 10: Multitouch interface for mixing audio (after Carrascal & Jorda 2011)

In an attempt to address the shortfalls of Gibson’s visual mixer Carrascal & Jorda (2011), building on the work of Diamante (2007), proposed the stage metaphor in their development of a multitouch interface for mixing audio (see Figure 10). They represented channels as movable nodes on a two-dimensional stage. The nodes could be moved freely from the inactive zone and placed on the stage. A listening point was also provided to make the interface suitable for surround sound mixing as well as stereo mixing. Movement of the node from left to right, relative to the listening point was used to control the channel’s pan position, with the distance from the listening point controlling the channel’s level. On selecting a node the channel gain could be controlled by changing the size of a surrounding circle called a halo.

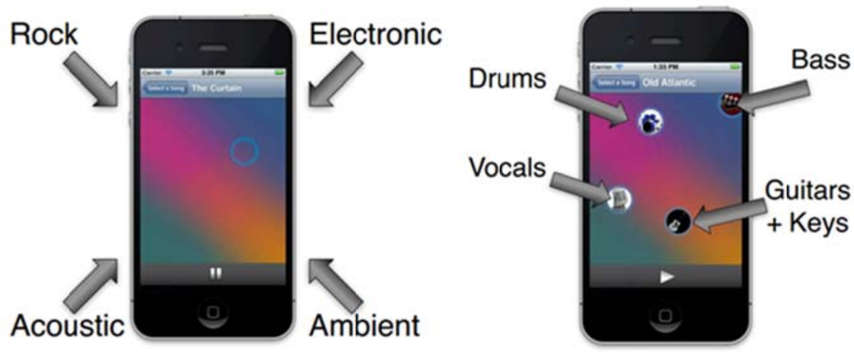

Figure 11: a) Personalisation of song playback b) Interface for simple multi-track mixing (Scott *et al* 2013)

Similarly, Scott, Chapman, Moorhead, Schwabe & Kim (2013) adopted the stage metaphor in the creation of a tool for the personalization of song playback on a smart phone as shown in Figure 11a. In this interface four different mixes of the same song are mapped to each corner of the rectangular stage. By moving the node through the two-dimensional space the user can blend the various versions of the song. Scott *et al* (2013) also detailed an interface for simple multi-track mixing. Users were provided with multiple nodes each representing an individual channel as a movable graphical icon. Again, users were able to move the nodes freely around the two-dimensional stage to mix the four instrument subgroups as shown in Figure 11b.

Carrascal & Jorda (2011) discovered that users were able to create a satisfactory mix in less time with their interface when compared with a traditional digital mixing console. Furthermore, novice users found the novel interface much easier to learn. In the development of a mixing interface Gelineck *et al* (2013) considered the stage metaphor in a preliminary investigation and provided the following explanation of why this may be the case. “Early results suggest that the conceptual model of the stage metaphor is much closer to the mental model of the user in how they conceive the mix” (Gelineck, Büchert & Andersen, 2013, p. 737)

The disadvantages of the stage metaphor were also summarised in this study with Gelineck *et al* (2013) noting.

> “The most crucial of these is the organization and overview of the different channels. With the channel strip metaphor each audio channel is always located in the same place, which makes it easy to find and control. With the stage metaphor channel representations are scattered around the virtual stage area and can be difficult to overview—especially since professional systems demand at least 24 channels (often more).” (Gelineck, Büchert & Andersen, 2013, p. 735)

3 EQ Interfaces

3.1 Introduction

The equaliser or EQ section of a mixing console is used to modify the tone and is defined as a "device for selectively cutting or boosting selected parts of the audio spectrum" (White, 1997, p. 298). EQ sections consist of a set of filters that can be characterised by the way they alter the frequency balance of a sound. Filters are often referred to as "bands" with each band containing a single filter (Izhaki, 2008). Commonly, equaliser sections offer a finite number of bands that can be harnessed by the user at any one time.

3.2 Filter Types

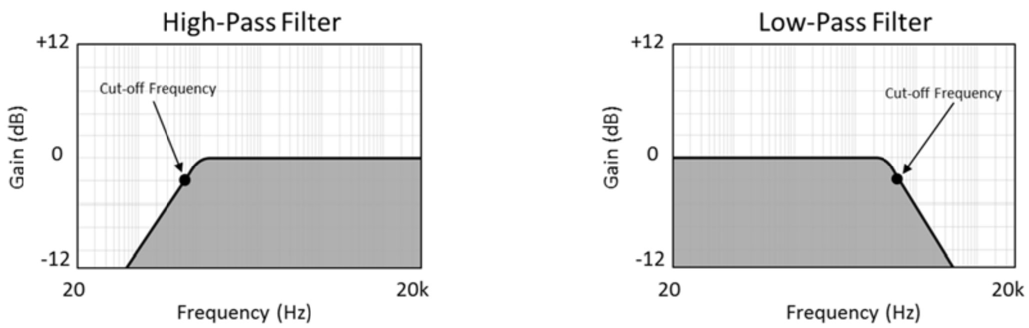

Figure 12: High pass and low pass filter properties and response curves

Pass filters allow frequencies to one side of the cut-off frequency to pass and cut frequencies to the other side as illustrated in Figure 12. High pass filters allow frequencies above the cut off frequency to pass and attentuate frequencies below the cut off frequency . Low pass filters achieve the opposite effect, allowing frequencies below the cut-off frequency to pass.

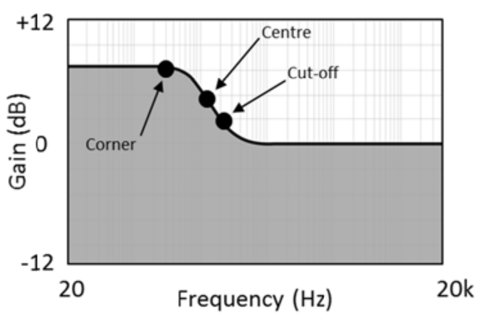

Figure 13: High shelving filter properties and response curve

Shelving filters differ from pass filters as they can be used to boost as well as cut frequencies. Figure 13 assists in explaining the characteristics of this type of filter. The centre frequency divides the frequency spectrum into two bands. On one side of the centre the frequencies are left untreated with frequencies cut or boosted by a constant amount to the other side. This is regulated by adjusting a gain control. The corner frequency refers to where the gain setting is reached. Izhaki (2008) states that there is often confusion around the term cut off, explaining that it is "the point at which 3dB of gain is reached" (Izhaki, 2008 p. 220). Two types of shelving filter exist. High shelving filters treat frequencies below the centre frequency and low shelving filters treat frequencies above.

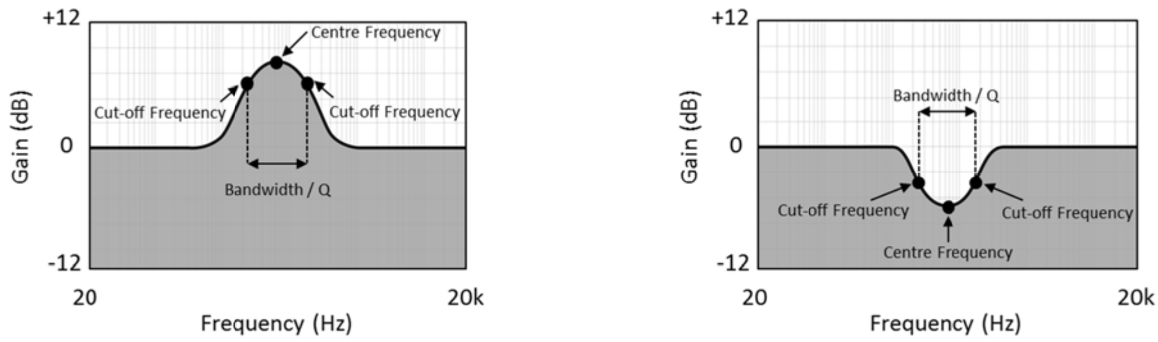

Figure 14: Peaking filter properties and response curves

Peaking filters are used to target specific areas of the frequency spectrum. Senior (2011) argues that peaking filters offer greater flexibility to the user. Izhaki (2008) describes the response curve for this type of filter as "reminiscent of the shape of a bell" (Izhaki, 2008, p. 224) and it is illustrated in Figure 14. Parametric filters enable users to control the centre frequency affected by the filter with the addition of a gain control to cut or boost the selected frequency range. The spectral bandwidth the filter affects can be altered by the user by means of the bandwidth or Q control.

The following sections will consider typical user interfaces provided for the EQ section on analogue and digital mixing desks and DAW software. Obviously the reader should be aware that whilst these are typical interfaces do vary within the categories. It will then move on to considering novel EQ user interfaces and audio analysis software displays.

3.3 Analogue Console EQ Interfaces

The mono input channel EQ section of the Allen and Heath ZED R16 features five EQ bands (Allen & Heath, 2008). The input section features a button for turning on a high pass filter. This has a fixed corner frequency of 100Hz and is positioned away from the channel's EQ section as it is commonly used to remove low frequency rumble from microphone input sources.

The main EQ section of the ZED R16 features four EQ bands as illustrated in Figure 15. The HF (meaning high frequency) knob is used to control the gain (+/- 15 dB) of a low shelving filter which has a fixed centre frequency of 12kHz. The HM (high-mid) and LM (low-mid) controls are peaking parametric filters and each feature three knobs. The knobs labelled Q are used to set the shape of the bell response with 0.8 referring to the widest setting and 6 referring to the narrowest. The knobs labelled F refer to the centre frequency of the filter. In the HM section the centre frequency can be set between 400Hz and 18kHz. In the LM section the centre frequency can be set between 18Hz and 1kHz. A gain knob is provided to provide +/- 15dB cut or boost. The LF (meaning low frequency) knob is used to control the gain (+/- 15 dB) of a high shelving filter which has a fixed centre frequency of 60Hz. A circular switch and LED is also provided at the bottom of the EQ section to bypass the the EQ controls.

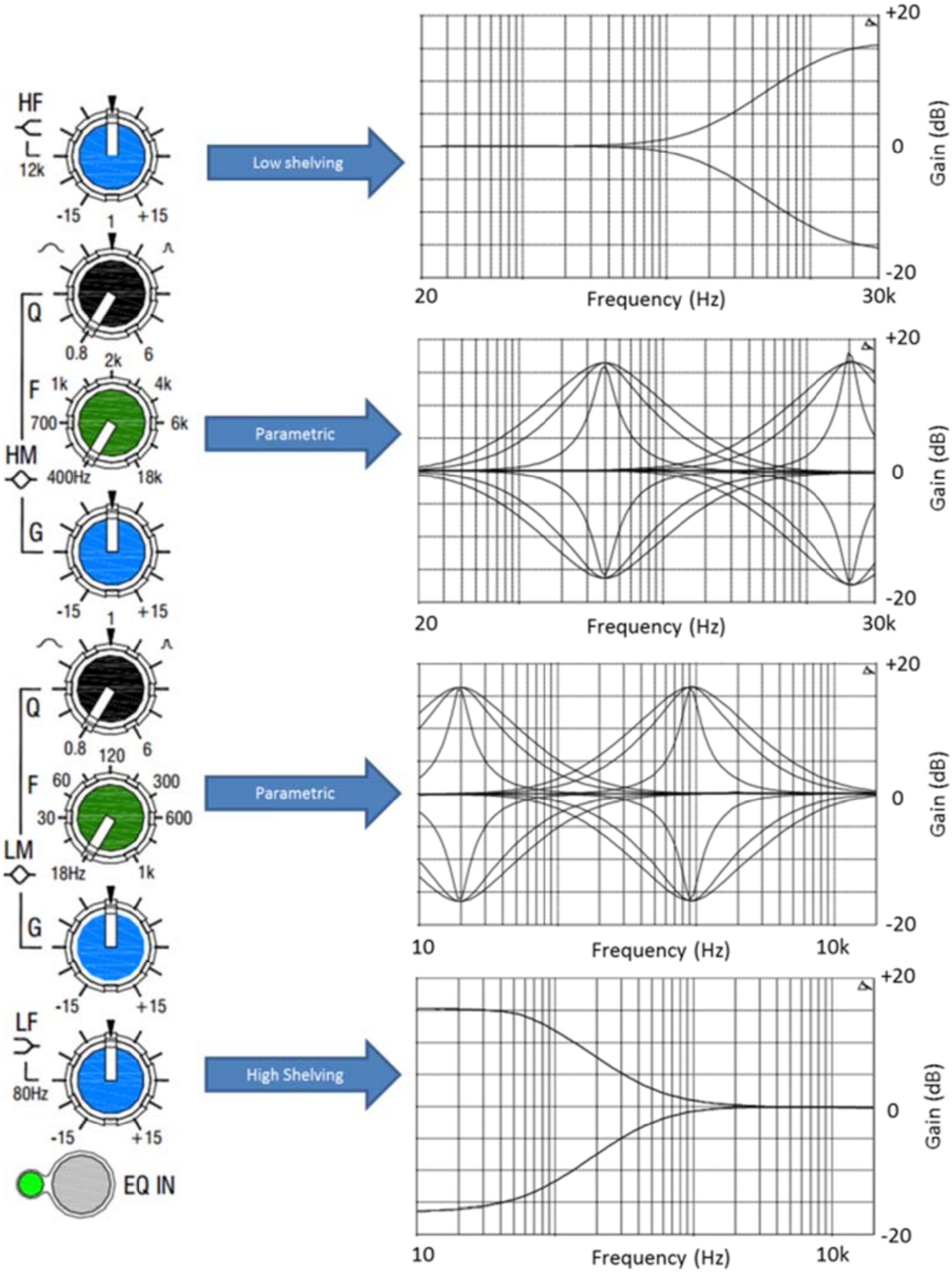

Figure 15: Allen and Heath ZED R16 mono channel EQ section and corresponding response curves indicating each filter's range (Allen & Heath, 2008)

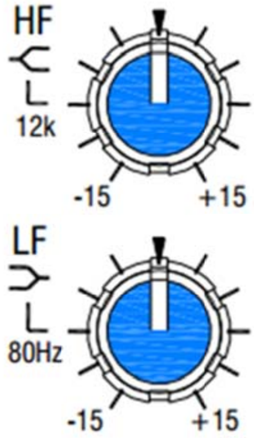

Figure 16: Allen and Heath ZED R16 stereo channel EQ section (Allen & Heath, 2008)

The EQ controls featured on the ZED R16's two stereo channels are more rudimentary in comparison as indicated in Figure 16. The HF knob is used to control the gain of a low shelving filter with the centre frequency fixed at 12kHz. The LF knob is used to control the gain of a high shelving filter with the centre frequency fixed at 80Hz.

The main drawback of traditional analogue mixing console EQ user interfaces is that the user has no real visual indication of what each filter is doing in the frequency domain. If the frequency response of filter bands found in all commerical interfaces were standardised this potentially wouldn't be such a problem however, each console embodiment employs slightly different controls and filter characterisitics. This means the user must have a sound technical and auditory understanding of EQ (Mecklenburg & Loviscach, 2006) to achieve consistent results when using the EQ controls of different analogue consoles. Furthermore, in comparison to other analogue channel strip controls which feature one control (i.e knob, fader or button) the EQ section features a suite of largely identical knobs which need to be adjusted in combination to apply settings as intended. This highlights another barrier to the user when using such an interface.

Figure 17: Samson S-Curve Dual 231 (copyright Samson, 2013)

Although not commonly featured on a mixing console, graphic EQs can also be used to modify the spectral characteristics of audio. One example of a graphic EQ is the Samson S Curve Dual 231 shown in Figure 17. Instead of providing the user with gain, frequency and Q controls, the audio spectrum is divided into thirty-one bands spaced a third of an octave apart, each controlled by a single fader (Samson, 2003).

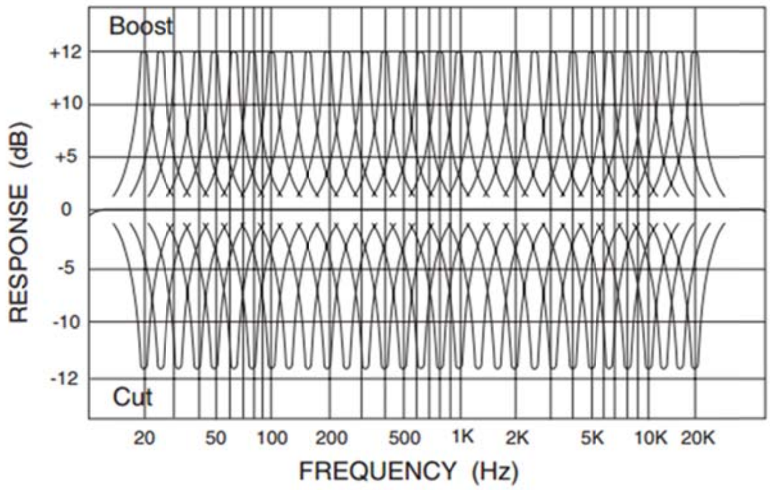

Figure 18: Response curves for the Samson S Curve Dual 231 (copyright Samson, 2003)

Senior (2011) argues that this sort of interface provides the user with a "visually intuitive" representation of how the audio spectrum has been modified. It should be noted however that this is visually misleading as adjacent frequencies are also affected when a fader is moved as shown in Figure 18. Graphic EQs are more commonly found in home hifi systems, live PA systems, mastering and if used in a recording studio used upon the main mix channels.

3.4 Digital Console EQ Interfaces

The Yamaha O2R96 console features a four-band parametric EQ for all input channels controlled via the selected channel section as shown in Figure 19. EQ settings can be stored in the O2R96's EQ library. This library also contains forty preset EQ settings for treating a range of audio sources (Yamaha, 2010).

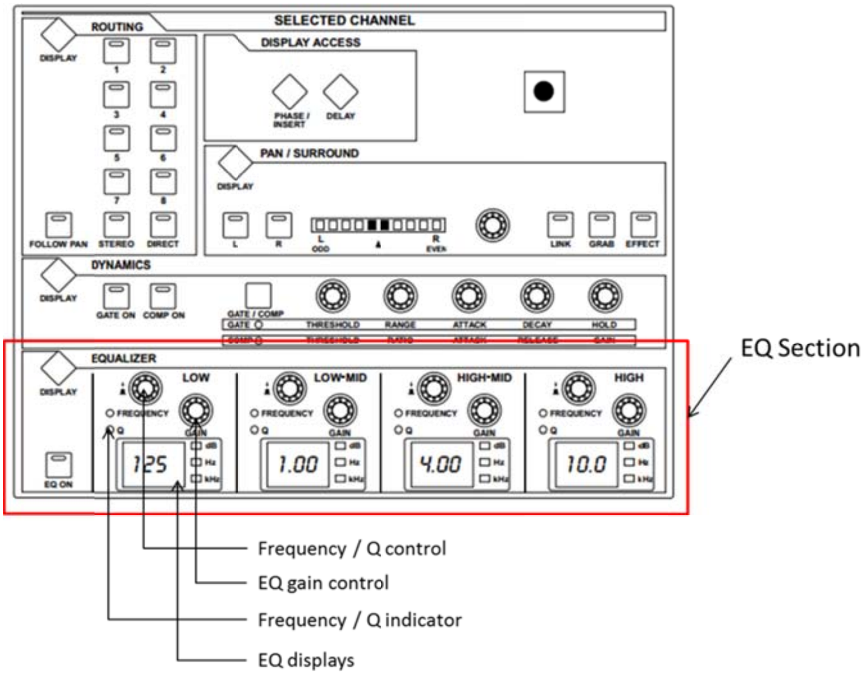

Figure 19: Yamaha O2R96 Selected Channel Section with EQ section and core controls highlighted (copyright Yamaha, 2010)

The LOW-MID and HIGH-MID bands feature peaking filters with the LOW and HIGH bands set as shelving, peaking or pass filters. Each band features two rotary encoders and a small numerical LCD display as shown above. The first encoder is a rotary push-switch encoder. The switch is used to change the mode from frequency to Q and vice versa with feedback provided via the frequency / Q indicator LEDs. Rotating the encoder sets the frequency or Q with feedback provided numerically via the small EQ display screen. The EQ gain control is also a rotary encoder and is used to set the gain (+/- 18 dB) of each filter band with feedback provided numerically via the EQ display. If the gain or Q remains unaltered for two seconds the EQ display returns to its default view showing frequency numerically.

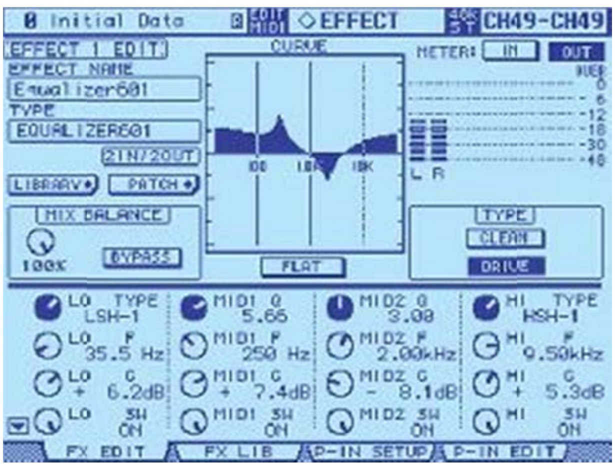

Figure 20: Yamaha O2R96 display EQ screen (Copyright Yamaha, 2004)

Further visual feedback is provided via the large display screen in the centre of the console. The user can alternatively interact via the display screen by moving from up, down and left to right and adjusting the selected visual knobs value. Most importantly this displays a graphical representation of the EQ curve applied as shown in Figure 20. Sabin & Pardo (2009) offer the following definition of the EQ curve.

> The effect of an equalizer can be described by the equalization curve, which describes the extent of energy boost or cut at each frequency. In general, frequency is on the x-axis and gain (the amount of boost or cut) is on the y- axis. (Sabin & Pardo, 2009, p. 1)

EQ curves offer a significant improvement over traditional analogue EQ controls because they provide visually intuitive feedback on what the applied EQ settings are actually doing in the frequency domain.

3.5 Digital Audio Workstation EQ Interfaces

In the last two years digital audio workstations (DAWs) have started to include a real time spectral plot of the audio currently being modified behind the EQ curve. Real time spectral plots have been included in audio analysis software plug-ins for some time and were originally implemented in audio mastering tools, for example the Oxygen 5 (Izotope, 2013). The software performs a Fast Fourier Transform (FFT) enabling the user to view audio in the frequency domain as opposed to the time domain (Pohlmann, 2011).

Figure 21: Cubase 7 EQ (Steinberg, 2013)

In Cubase 7's EQ section the EQ curve and associated controls are displayed in a graphically rendered window that can be opened and closed as required as shown in Figure 21 (Steinberg, 2013). In this example the user can interact with nodes placed directly on top of the EQ curve which represent the filter band’s centre frequency. A suite of knobs beneath the display enables the user to fine tune their EQ settings. The real time spectral plot can be configured to display both pre-EQ and post-EQ plots. This key inclusion enables users to see which frequencies are dominant and where EQ correction may be necessary in the pre-EQ plot and also to evaluate the success of their EQing in the post-EQ plot.

All major DAW solutions now offer this as a standard feature in their latest releases. This includes Studio One (Presonus, 2013), Live 9 (Ableton, 2013) and Logic Pro X (Apple, 2013) and possibly indicates a cultural shift towards software developers considering assistive mixing tools in DAWs.

3.6 Line 6 StageScape M20d EQ Interface

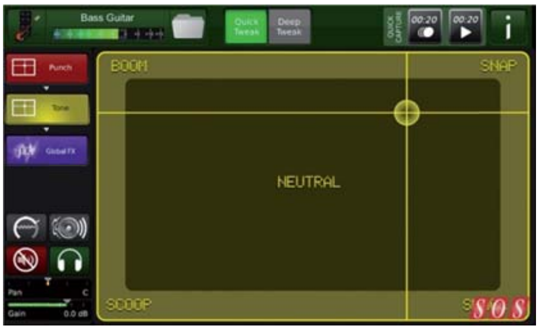

Figure 22: Line 6 StageScape M20d EQ quick tweak interface (Copyright Sound on Sound, 2012)

The EQ section of the Line 6 StageScape M20d can be accessed by selecting the graphical icon for the channel under consideration and pressing the tweak mode button to the left of the display (Line 6, 2012). By default, the user is presented with the “quick tweak” interface. This interface consists of a rectangular X Y pad with four keywords placed in each corner of the display and one keyword in the centre of the display as shown in Figure 22. The range of keywords displayed are relevant to the instrument selected in setup mode. Using the example of the lead male vocalist the words “project”, “clarity”, “air” and “full” are displayed in the corners of the interface with “neutral” displayed in the middle. The user is able to interact with the X Y pad using a single circular node. Moving the node closer to a keyword changes the EQ curve applied in real time to reflect the characteristic intended by the keyword.

Figure 23: Line 6 StageScape M20d EQ deep tweak interface (Copyright Sound on Sound, 2012)

The “deep tweak” mode is provided to further refine and configure EQ settings as shown in Figure 23 and is selected by clicking on the deep tweak button at the top of the display. In this mode the X Y pad is replaced with a traditional DAW style EQ interface featuring up to six EQ bands. These include a low shelf, high shelf and four peaking parametric bands. A coloured node is provided for each band’s centre frequency which can be moved via the touchscreen. A row of faders at the bottom of the touchscreen are provided to control the centre frequency (20Hz to 18kHz), Q (0.1 to 10) and gain (+/- 15 dB) of each EQ band. The twelve rotary encoders beneath the touchscreen augment the control and change colour providing an indication of the associated band. Following the recent trend, a real time post-EQ spectral plot is displayed behind the EQ curve.

3.7 Novel EQ Interfaces

Mecklenburg & Loviscach (2006) sought to simplify the EQ interface by developing the subjEQt. This interface adopts the subjective approach found in the StageScape M20d quick tweak interface (see Figure 24) by displaying a “visual arrangement of subjective terms such as ‘warm’, ‘present’ and ‘boomy’ instead of the standard controls” (Mecklenburg & Loviscach, 2006, p. 1).

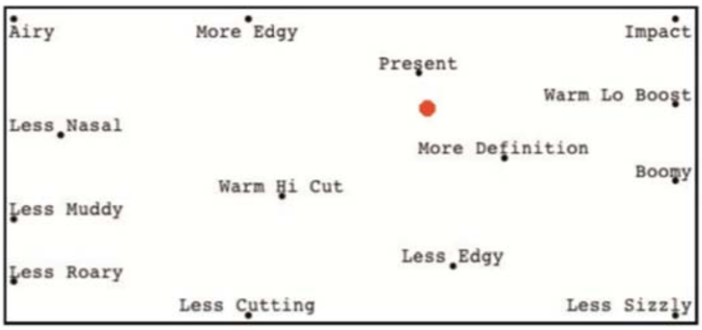

Figure 24: The subjEQt interface (Mecklenburg & Loviscach, 2006)

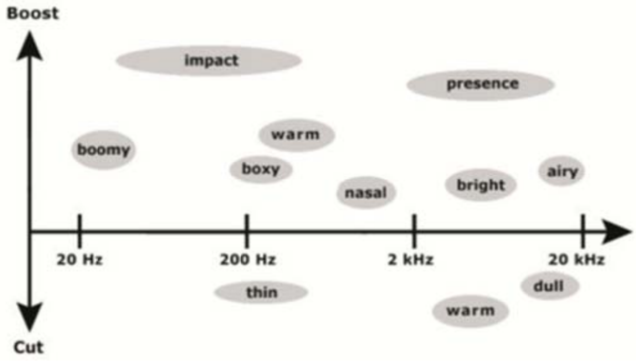

Figure 25: Frequency ranges of subjective terms (Mecklenburg & Loviscach, 2006)

In designing the subjEQt Mecklenburg & Loviscach (2006) defined a range of EQ curves for each keyword (subjective term) featuring a single peaking filter band as illustrated in Figure 25. The position of each keyword on the interface was determined by the similarity of these EQ curves. Put simply, keywords that performed a similar actions were placed close together, keywords that had opposing EQ curves (such as high cut and high boost) were placed far apart and keywords that when combined had a meaningful result (such as low cut and high boost) had a medium distance apart.

A freely movable node was provided for user interaction with the applied EQ curve being a product of where the node was positioned relative to the range of keywords displayed. A traditional graphic EQ was also provided although this was hidden by default and users were only able to use this to see the EQ curve implemented by the subjective interface.

The main problem with using subjective terms over traditional controls is that keywords infer different meanings to different people. Mecklenburg & Loviscach (2006) confirmed this limitation with regard to user skill level.

> The terms were very common to professionals; they get what they expect. Furthermore, the terms are novel to non-musicians; they do not object to what they get. Amateur musicians, however, may have heard some of the terms already, but have formed their own ideas of what they mean. So they may get different results from what they expect, which is a frustrating experience, in particular for this kind of interface which relies on agreed-upon meanings for subjective terms. (Mecklenburg & Loviscach, 2006, p. 5)

Seeking to address this limitation Sabin & Pardo (2009) used a different method for generating the keyword EQ curves in the development of the 2DEQ. In this example the EQ curves were based on weighting functions generated from a listening test. During a preliminary investigation, subjects were presented with a range of EQ curves and asked to rate how well the curve captured their concept of four subjective terms, namely tinny, bright, warm and dark. The results of this test were then interpreted and used to create an EQ curve for each keyword.

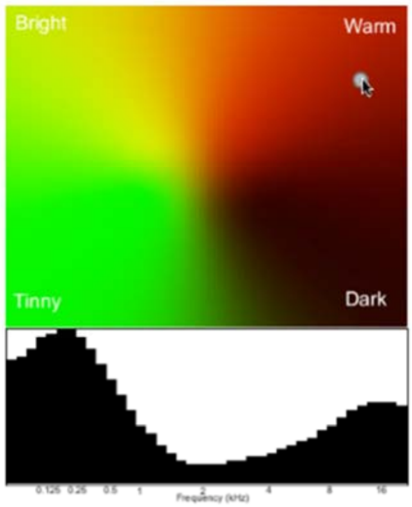

Figure 26: The 2DEQ interface (Sabin & Pardo, 2009)

The 2DEQ interface is nearly identical to the quick tweak section featured in the StageScape M20d with keywords placed in each corner of the stage and one movable node for control (see Figure 26). Interestingly, feedback is also provided to the user as a standard EQ curve displayed beneath the stage, possibly indicating a user's desire for a visual representation of what has been modified in the frequency domain.

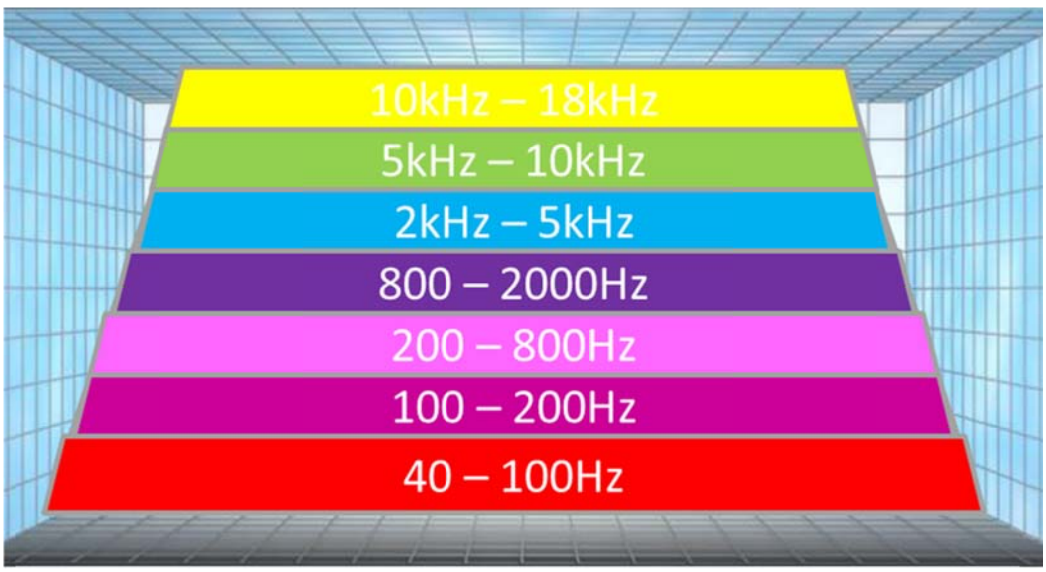

Figure 27: The visual mixer EQ interface (Gibson, 2008)

Figure 27 shows the EQ interface proposed by Gibson (2008) in his visual mixer. In this implementation a tapered four sided polygon is displayed in the centre of the stage. The polygon is divided into seven horizontal coloured bands with each band representing a frequency range.

Gibson used the brightness of each colour band to represent the frequency content of the audio channel scrutinised. Gibson also proposed using the size of each coloured band as an indication of filter bandwidth. Whilst this is in many ways a worthy alternative to the traditional EQ interface, little information is provided to explain how the user could interact with the coloured bands to apply EQ settings. For this reason the interface provides greater worth as a visual audio analysis tool than a tangible EQ interface.

3.8 Audio Analysis Interfaces

Data visualization techniques are employed to display audio analysis information to the user in a variety of applications. Software engineers have been designing interfaces that display spectral information for use in mastering for some time. More recently DAW applications have started to feature these sorts of interfaces.

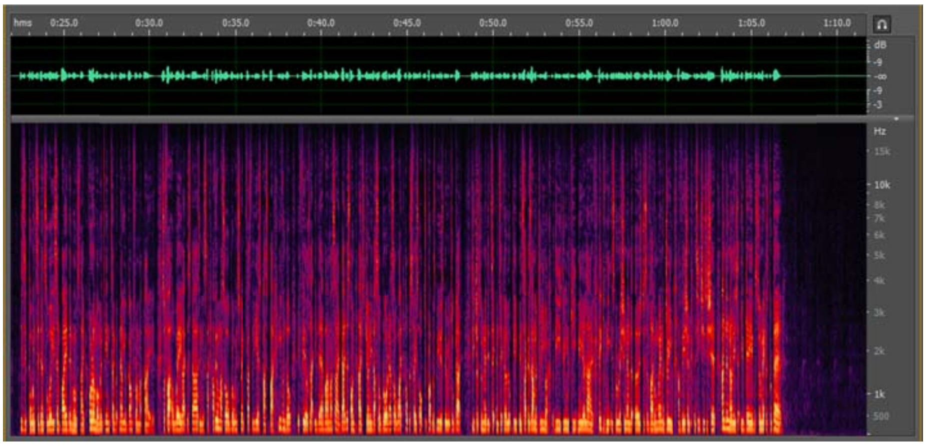

Figure 28: Adobe Audition CS5 spectrogram view

Spectrograms are commonly used to provide a visualization of changes in the frequency domain over time as shown in Figure 28. This is often displayed in a two dimensional format with time displayed on the horizontal axis and frequency on the vertical axis. Each captured FFT frame is displayed as a thin vertical time-aligned bar. The colour of the bar changes at different frequency intervals along the vertical axis to represent the relevant magnitude or perceptual “heat” of each interval sampled.

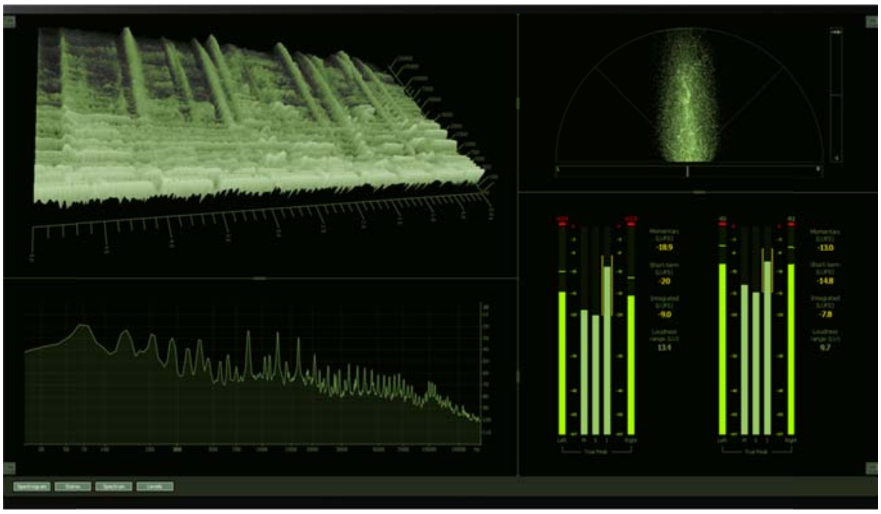

Figure 29: Ozone 5 Waterfall Plot (Izotope, 2013)

A spectrogram displayed in three-dimensions is known as a waterfall plot. One example of a waterfall plot is featured in the metering plugin Ozone 5 (Izotope, 2013) shown in Figure 29. In this example spectral plots are displayed with magnitude vertically aligned and frequency horizontally aligned. Differences in the spectral makeup over time are displayed by offsetting the spectral plots along the Z axis giving the perception of depth. These plots are used to create a topographical surface that resembles a waterfall or mountain range to assist the user in visualising the changes over time. User experience is enhanced in Ozone 5 with controls for zoom and rotation of the waterfall plot included.

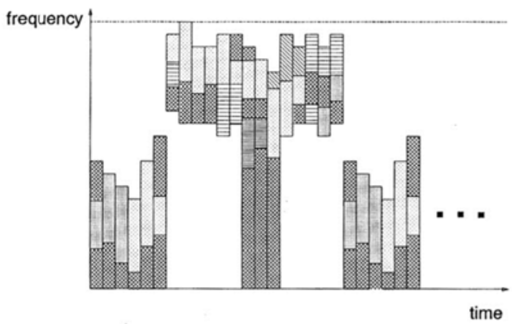

Figure 30: Novel interface for displaying the spectral content and phase differences of a stereo channel (Eastty, 2000)

Eastty (2000) attempted to display the spectral content and phase differences of a stereo audio channel in a single interface, as shown in Figure 30. Time was represented along the horizontal axis with frequency along the vertical axis split into sixty-one bands. The colour of each band was determined by the phase differences between the left and right channels.

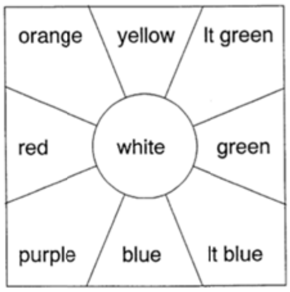

Figure 31: Colour mapping of phase (Eastty, 2000)

Figure 31 helps explain how each colour was derived from overlaying a series of coloured zones over the traditional phase scope interface. By following along the same horizontal level of the display from left to right it is possible to see the signal content and channel phase at a particular frequency band over time. By looking at the vertical direction at a particular point in time it is possible to see where in the frequency spectrum the energy is concentrated and the relative phase of the two channels at each frequency. Eastty (2000) argued that this implementation offers a more favorable visualization of spectral and phase differences over time when compared against the higher resolution spectrogram and phase scope displays.

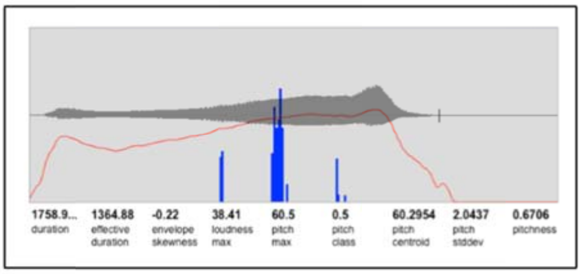

Figure 32: Figure 32: Novel interface for displaying volume envelope and pitch (Schnell, Cifuentes & Lambert, 2010)

Schnell, Cifuentes & Lambert (2010) proposed an interface to display the results of volume envelope and pitch detection analysis on for a single audio channel. In this example, the horizontal axis is used to display both time and frequency with magnitude on the vertical axis. The traditional waveform view is overlayed by a static plot indicating the volume envelope using the horizontal time measure. Vertical lines are drawn on top of these two display elements to give an indication of pitch which uses the horizontal frequency measure. Key analysis information is presented in numerical format beneath the graph. Of particular note, is the lack of scale present on the interface to indicate time, magnitude or frequency, instead relying on the established assumption that time and frequency is displayed from left to right and magnitude from low to high.

These examples introduce design ideas and concepts worthy of consideration in the design of a novel EQ interface. The inclusion of static plots is of particular interest and could be more intuitive to the user when compared with the widely adopted real time approach. With a realtime display the display is by definition always varying and can make it harder for the user to get an overview and identify frequencies requiring attention. A static display enables the user to quickly gauge the frequency content of the audio channel under scrutiny. Static displays also provide greater ease if the user is to interact directly with the display. The novel interfaces also introduce the concept of simplifying the high resolution displays featured in commercial interfaces. Furthermore, they provide examples of how a variety of data variables can be displayed as a single simple entity.

4 Interface Design: HCI Fundamentals

4.1 Introduction

The field of Human Computer Interaction (HCI) has become established over the last three decades. HCI is a multi-disciplinary field concerned with the study, planning and design of computer interfaces (Erikson & McDonald, 2008). The overall aim of HCI is to inform the creation of interfaces that enable users to successfully reach their goals (Norman, 2002). Goals can be defined as the overall objective of the user. Shneidermann & Plaisant (2005) present three metaphorical pillars that can be used to turn ideas into good user interfaces, as shown in Figure 33.

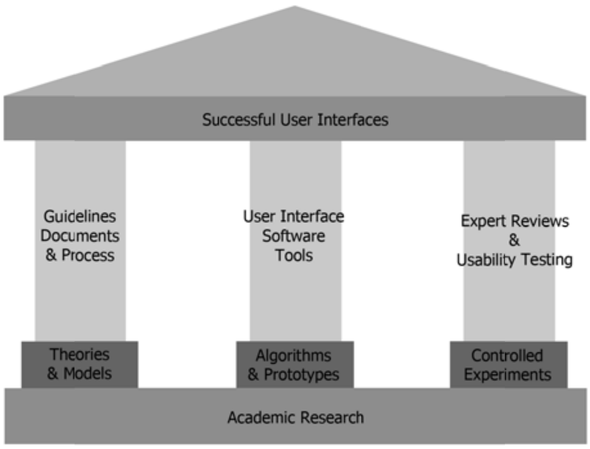

Figure 33: The three pillars of successful user-interface development (after Shneiderman & Plaisant, 2005)

4.2 Principles

A range of principles have been collated by Shneidermann & Plaisant (2005) to assist the designer in the creation of user interfaces. They argue that cluttered, inconsistent interfaces that provide little feedback directly affect a user's ability to reach their goals, placing undue cognitive stress on the user. These principles can be organised into a series of stages to help develop better interfaces.

4.2.1 Identification of Users and Tasks

Hansen's (1971) first principle of design is that one must "know thy user" and designers must be aware that users think, learn and solve problems in different ways. Users can be defined by their skill level and broadly categorised as either novice, knowledgeable intermittent or expert. Novice users require a limited number of actions and must be able to carry out simple tasks successfully in order to gain in confidence when using an interface. Knowledgeable intermittent users will have a broad understanding of the underlying concepts conveyed by the interface and processes required to perform a task successfully but need consistency and meaningful feedback to enable them to remember what processes are needed to perform a task successfully. Expert users fully understand the processes required to complete a task successfully and require tools and feedback that will enable them to work quickly. In order to design an interface that is useful to all three types of user, a multi-layered approach is often adopted.

Once the skill level of the user has been considered it is crucial to determine the range of tasks undertaken by the user to reach their goal. Tasks can be defined as the actions a user performs when interacting with an interface. These tasks can be identified by observing and interviewing users, a method called task analysis (Crystal & Ellington, 2004). Task analysis provides an indication of how frequently tasks are performed and helps identify a task order. Task analysis enables complex actions to be broken down into a series of tasks and helps the designer decide the range of controls required for the user to achieve their goals. Consequently a detailed dissection of the EQ task is included in Chapter Six.

4.2.2 Selection of appropriate interaction style and visualisation metaphor

Once the skill level and tasks are established, a designer must choose an interaction style (Shneiderman & Plaisant, 2005). Direct manipulation interfaces graphically represent the world, concept or data to be conveyed in order to allow improved user interaction. This style is favoured by designers as it provides a visual metaphor which in turn facilitates ease of learning and retention, error avoidance and user satisfaction. The Line 6 StageScape M20d is one example of a direct manipulation mixing console interface as it enables users to interact directly with a representation of musicians on a stage.

Shneiderman (1996) asserts that interface researchers and designers are increasingly using graphical means to display dynamic information, a method termed data visualisation. This is because there exists is a widespread belief amongst designers that visual displays take advantage of the users' cognitive ability to "scan, recognise, and recall images and … detect changes in colour, size, shape, movement or texture" (Shneiderman, 1996 p. 2).

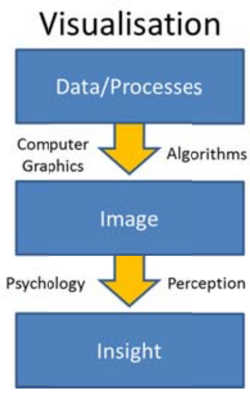

Figure 34: The transformation from numbers to insight

Bertin's image theory (1983) asserts that data visualisation is in fact a joint function of computer graphics and perception. Bertin's (1983) work identifies two stages that lead to a user gaining insight from a data visualization, as shown in Figure 34.

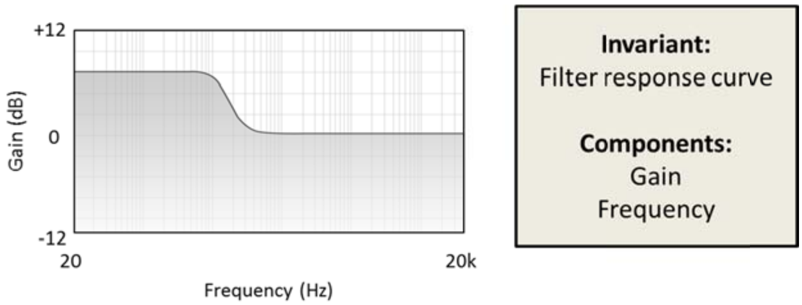

Figure 35: Image invariants and components for a frequency response curve

An image can be defined as the fundamental perceptual unit of any visualisation with each image consisting of two parts termed components and invariants. A component is the concept conveyed to the user and an invariant relates these components together. Ideally, one image should be presented to the user for simplicity. Users extract information from data visualisations by firstly externally identifying what is being represented. The user then internally identifies how the components are mapped before perceiving what is being displayed, a process termed abstraction (Myatt & Johnson, 2008). In the example provided in Figure 35 external identification is the realisation that the graph is showing something about gain and frequency. Internal identification is the perception that gain is mapped to the vertical axis of the plane and frequency to the horizontal axis to provide a frequency response curve. Perception is the realization that the filter is boosting frequencies between 20Hz and 500Hz.

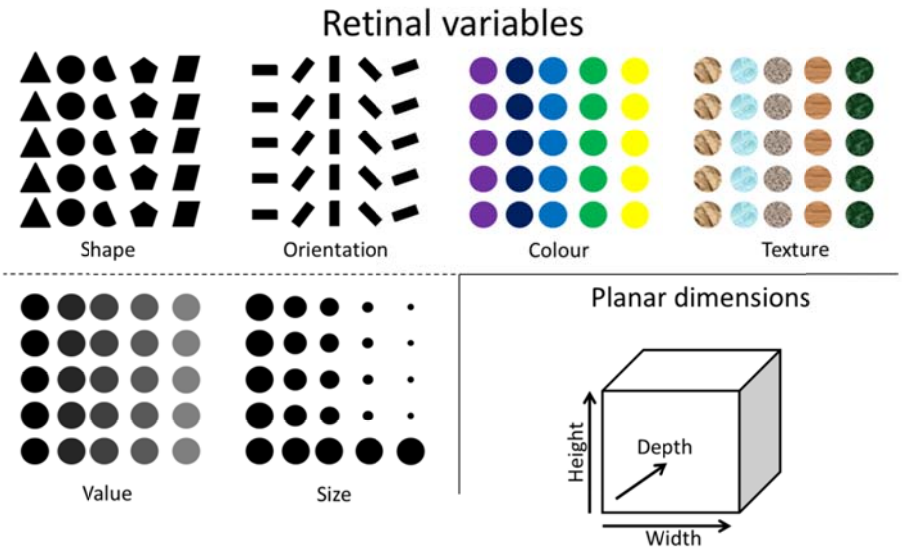

Figure 36: Planar dimensions and retinal variables (after Bertin, 1983)

Bertin (1983) advises that an optimum of three visual variables can be perceived and understood by the user in each image. These visual variables are classified as either planar or retinal, as shown in Figure 36. Planar variables exist as spatial dimensions, i.e. height, width and depth and retinal variables include size, colour, shape, orientation or texture. Importantly, each image must consist of both planar and retinal variables.

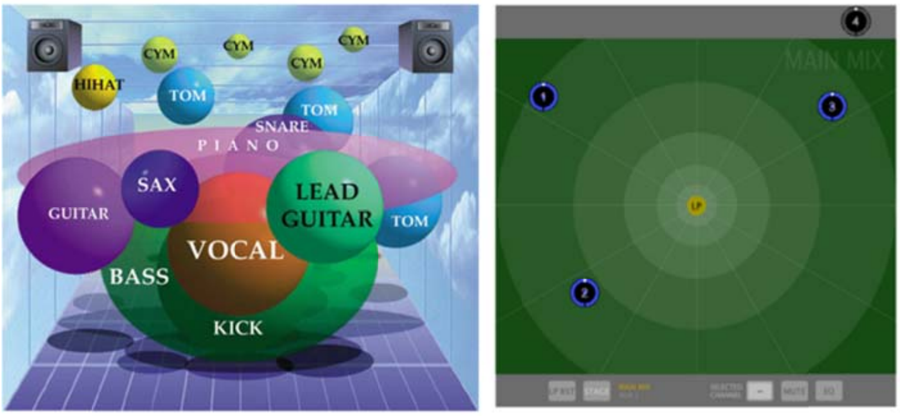

Figure 37: Comparing differences in dimensions used to visualize data in Gibson's (2008) and Carrascal & Jorda's (2011) mixing interfaces

Bertin (1983) states that if more than four dimensions are displayed in any image a user's perception will be hindered by the complexity of the display. This concept is contextualised in Figure 37. A visual analysis of the two examples reveals that Gibson's mixing interface is visually more complex than Carrascal & Jorda's. This is because Gibson's interface uses six dimensions consisting of three planar variables and three retinal variables (colour, size and transparency) whereas Carascal & Jorda's interface relies on three dimensions consisting of two planar variables and one retinal variable to display the same information. Norman (1994) cited by Myatt & Johnson (2011) offers the following guidance. "For representation to be effectively used, it must make only the relevant details explicit; in other words it must be at the right level of abstraction" (Myatt & Johnson, 2011) . This guidance is important because it promotes the design of interfaces that display only the information relevant to the overall goal. In other words less is more.

Original data

0	1	1	2	1
1	2	2	0	1

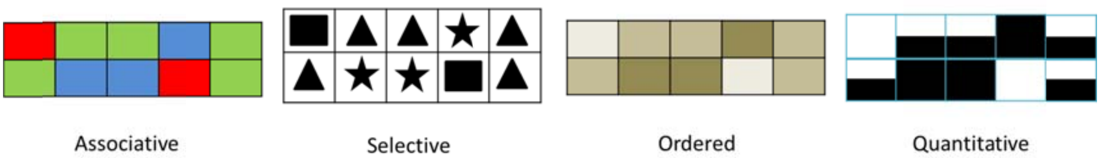

Figure 38: The different types of organisation (after Bertin, 1983)

Four different levels of organization are defined by Bertin (1983). Figure 38 provides a simple example to assist in understanding this concept. Associative organization is the lowest level defined and allows users to group data despite their values. Selective organization allows the comparison of data values. Ordered organization allows the user to visually identify that one data value is larger than another. Quantitative organization is the highest level of organization defined allowing the user to directly extract data ratios by reading the data values from the display.

	Associative	Selective	Ordered	Quantitative
Planar	X	X	X	X
Size		X	X	
Brightness		X	X	
Texture	X	X	X	
Colour	X	X		
Orientation	X	X		
Shape	X			

Figure 39: Level of organisation (Bertin, 1983)

The type of organization perceived by the user is dependent on the type of variables employed in the display with planar variables affording the greatest level of abstraction as illustrated in Figure 39. Green (2006) contextualizes Bertin's (1983) work with regard to interface design, drawing on the field of perception and psychophysics in a bid to help designers make more intuitive solutions. Green argues that designers should use multiple non-spatial dimensions such as colour, brightness and size simultaneously. Green offers further guidance stating that data is best represented on a spatial axis with a limited number of non-spatial dimensions displayed at any one time..

Shneiderman (1996) outlined a basic taxonomy of data visualisation, classifying visualisations according to the number of dimensions displayed. More recently, Behrens (2008), Lima (2011) and Lee (2012) have offered their own taxonomies to describe the various forms employed to display data visually.

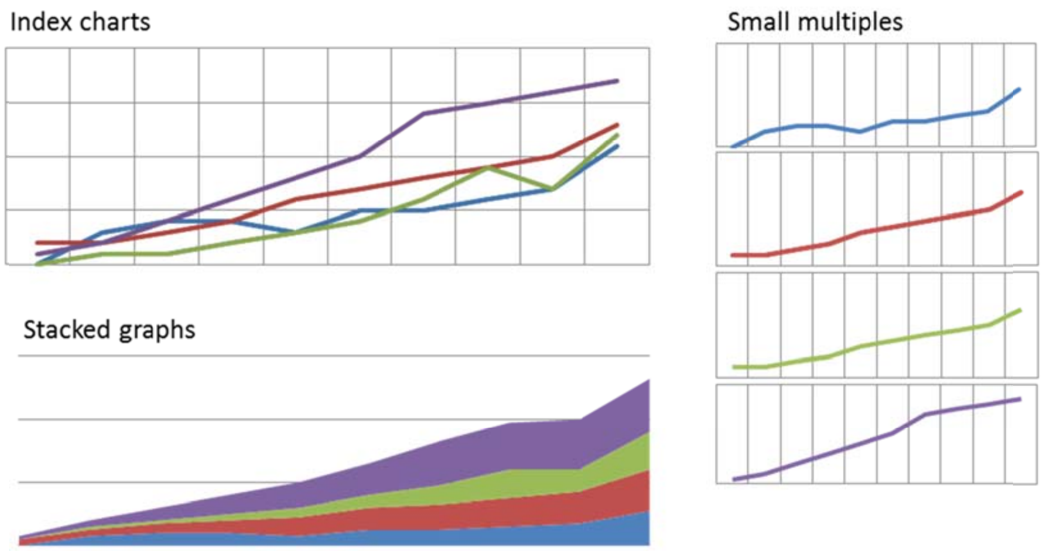

Figure 40: Three types of time series data visualisations (after Heer et al, 2010)

Heer, Bostock, & Ogievetsky (2010) presented a detailed overview of the main data visualisation forms. They identified three forms commonly used to display values that change over time, as shown in Figure 40. These forms are used to compare multiple series of data and are relevant to this project because they could also be also considered to display the frequency spectrum for multiple channels of audio.

4.2.3 Seven Stages of Action

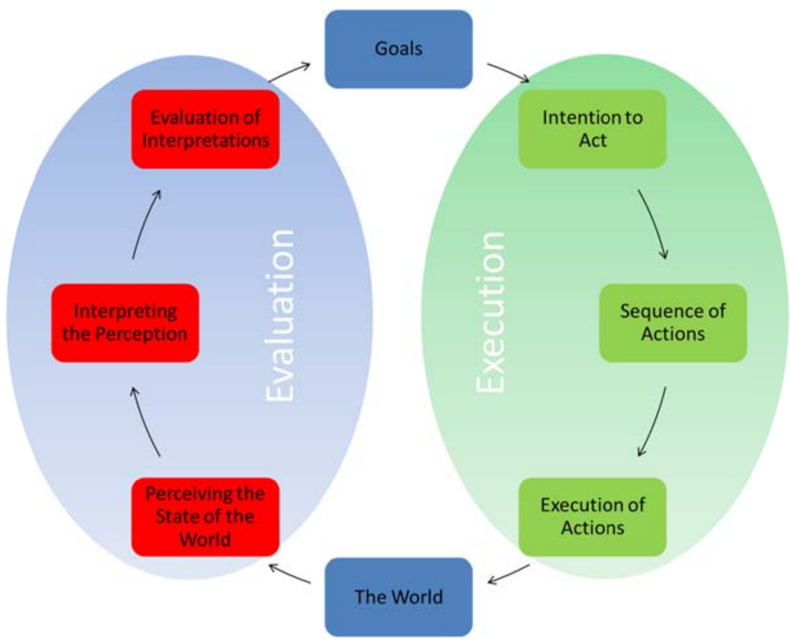

Figure 41: Seven stages of action (Norman, 2002)

Building on the image theory outlined by Bertin (1983) et al, Norman (2002) considers how humans interact with the world and identifies seven stages of action as shown in Figure 41. Actions enable us to achieve our goals within the physical constraints of the world. Actions are first identified and then executed by the user. Once the action has been executed the user then perceives and interprets the state of the world before evaluating the outcome. Norman (2002) argues that these stages should be considered as cyclical in nature, with users continually modifying their goals based on the evaluation of their actions.

How well the world is represented and how well we interpret that representation is considered by Norman (2002) as a gulf of execution and evaluation. The gulf of execution is the difference between the users intentions and the actions allowed by the interface. The gulf of evaluation represents the amount of cognitive effort a user must employ to determine the state of the system and what actions must be taken to meet the user's intentions. Good design implementations exhibit narrow gulfs.

Norman (2002) suggests using the seven stages of action model to elicit a series of questions the designer must consider when developing an interface. These are highlighted in Figure 42.

How easily can one:	
Determine the function of the device?	
Tell what actions are possible?	Tell if system is in the desired state?
Determine mapping from intention to physical mapping?	Determine mapping from system state to interpretation?
Perform the action?	Tell what state the system is in?

Figure 42: Design questions derived from the seven stages of action model (Norman, 2002 p. 53)

Norman reinforces the importance of these questions as a basis for interface design by relating the factors they address to four core design principles.

1) Visibility: by looking the user can see the state of the device and the alternatives for action.

2) A good conceptual model: the designer provides a good conceptual model for the user, with consistency in presentation of operations and results and a coherent, consistent system image.

3) Good mappings: it is possible to determine the relationships between actions and results, between the controls and their effects, and between the system state and what is visible.

4) Feedback: the user receives full and continuous feedback about the results of actions. (Norman, 2002, pp. 52-53)

4.2.4 Automation

Another facet of interface design is the inclusion of task automation. Routine tasks can be automated allowing the user to devote their attention to making creative and higher-level decisions. Reiss has been active in the field of automated mixing of audio and with others has developed a number of systems that when combined together can create a mix from a set of sound sources. This research has focused to date on automatic stereo panning (Gonzales & Reiss, 2007), a control for minimizing the masking effects between audio channels (Gonzales & Reiss, 2008) and a tool for automatically normalizing gain (Gonzales & Reiss, 2008). More recently this research has provided the foundation for the commercial development of an automated mixing tool branded MixGenius (MixGenius, 2013). With specific regard to EQ automation, MixGenius employs parametric EQ filters designed directly in the digital domain (Reiss, 2011), an intelligent EQ tool that uses an algorithm based on the Yule-Walker method (Ma, Reiss & Black, 2013) and the automatic EQ of multiple audio channels using cross-adaptive methods (Gonzales & Reiss, 2009).

4.3 Evaluation

Nielsen (1994) defines usability inspection as a general term for a set of methods used to inspect an interface. Nielsen (1994) recommends several evaluation methods to augment and steer the development cycle. Ten principles termed the heuristics are outlined:

- Visibility of system status
- Match between system and the real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design
- Help users recognize, diagnose, and recover from errors
- Help and documentation

Nielsen (1994) argues that these principles form the basis of heuristic evaluation. This is an informal evaluation method that involves experts assessing how well each design element follows the heuristics. A variation of this is heuristic estimation where experts estimate the relative usability of two or more interface designs. Cognitive walkthroughs involve a simulation of user experience for a specific task to check that the correct controls are provided to allow users to achieve their goals. In pluralistic walkthroughs a focus group of designers and experts run through scenarios with a prototype interface discussing each element. A feature inspection involves designers listing the sequences of steps required to accomplish typical tasks checking for any long, cumbersome or unnatural patterns.

4.3.1 Fitts Law

Interfaces were traditionally evaluated in terms of their mechanical characteristics. Card, English & Burr (1978) were the first to use Fitt's Law for the evaluation of HCI input devices.

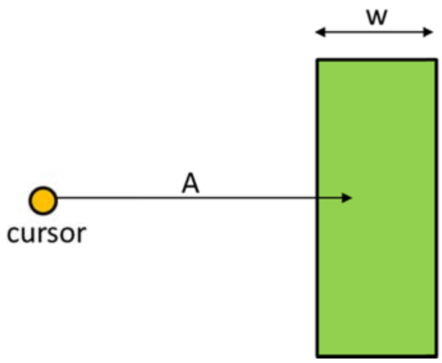

Figure 43: Fitts Law

Fitts Law states:

$$T = a + b \log(2A/W) \tag{1}$$

Where T is the time needed to point to a target, W is the target width and A is the linear distance from initial hand position to target position as illustrated in Figure 43. Constants a and b are empirically determined with the logarithmic term $\log(2A/W)$ named the index of difficulty (ID) and measured in bits. Tasks of greater difficulty have higher IDs. The reciprocal of b is the index of performance (IP) measured in bits/sec and represents the human rate of information processing for the task under consideration. The most widely used version of Fitt's Law is the Shannon formula (MacKenzie, 2003 cited by O'Modhrain, 2011).

In their evaluation of user interfaces for discrete sound placement in film and TV post production, Hughes & Wakefield (2012) favoured the Fitt's Law based evaluation arguing that it provides a mathematical link between time to locate a target and the accuracy of that location of the corresponding sound source.

Jacob, Deligiannidis & Morrison (1994) suggested that evaluation should move away from focussing on the mechanical structure of the interface and consider the suitability of the perceptual structure for the task under consideration. Jacob *et al* (1994) believe that devices with controls matching the perceptual structure of the task afford greater user performance. This belief is mirrored by Wanderley & Orio (2002) who define the perceptual structure as a metaphor of manipulation and argue that it is the conceptual model of the interface that should be evaluated. Furthermore, building on the work of Hunt & Kirk (2000), Wanderley & Orio (2002) state that audio interface design is in fact a very specialised field of HCI involving bi-directional communication between human and computer. Users control many parameters simultaneously with feedback experienced not through on screen prompts, but by experiencing the effects of actions on the audio in real time.

4.3.2 Usability

The International Standard of usability, ISO 9241-11 (1998), defines usability as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (User Focus, n.d. para. 5).

The evaluation of audio industry interfaces in terms of usability has received little significant attention (Carrascal & Jorda, 2005) however there is some significant guidance offered in the related field of digital musical interface (DMI) evaluation.

Wanderley and Orio (2002) suggest evaluating DMIs by timing users performing a range of very simple musical tasks with the proposed interface to provide a quantitative benchmark of usability. They argue that a musical context is required to make the evaluation relevant and provide the context of use defined in the standard. Furthermore they suggest measuring efficiency as task completion time, effectiveness as accuracy and satisfaction as user preference. O'Modhrain (2011) concurs and suggests capturing a real time log of user interaction which can then analysed in terms of accuracy, task completion time or the trajectory of movements. Travis (2008) agrees with these metrics and argues that all three variables are independent and in designing tests one must consider a measure of each to form a rounded evaluation of usability.

O'Modhrain (2011) argues that evaluation of DMIs must be discussed from the perspective of each stakeholder namely audience, performer and manufacturer. Particularly pertinent to this research is the performer or in this context, the user's perspective. User experience has been captured in the past by asking the users to provide a running commentary on their experience when using a new interface (van de Haak & de Jong, 2003). O'Modhrain (2011) suggests using qualitative methods in parallel with the quantitative methods identified for measuring efficiency and effectiveness. This involves probing test subjects on a cognitive level using a post-test questionnaire.

Numerous researchers argue that greater insight can be gained from semi structured interviews (Paine, Stevenson, & Pearce, 2007; Stowell, Plumbley, & Bryan-Kinns, 2008; Chuchacz, 2009). In comparison, this approach can capture subtle differences in user experience allowing quantitative results to be extracted from an unstructured dialogue. Stowell *et al* (2008) take the analysis of user responses further by defining a discourse analysis method that can be conducted on the verbal content of post-test interviews which in essence methodically reorganises the test subject's language. User responses are firstly transcribed with any surface impressions noted by the analyst. The transcribed data is then itemised to produce a list of objects, actors and descriptions. The described world is then reconstructed to include important factors and relationships and examined in terms of the test context. Kiefer (2008) employs a similar method in the evaluation of musical controllers by combining the quantitative data analysis from user log files with key user comments from a post-test interview in a database.

A common error in measuring satisfaction is to implement only a post-test questionnaire (Travis, 2008). Furthermore, when designing a questionnaire Travis (2008) states that there are many issues to consider. Firstly, Cronbach (1946) discovered that test subjects are more likely to agree than disagree with a statement, a phenomena termed the acquiescence bias. The two most commonly used questionnaires are the Usefulness, Satisfaction and Ease of Use (USE) questionnaire (Lund, 2001) and the Computer System Usability (CSUQ) questionnaire (Lewis, 1995). These examples phrase all questions in a positive bias meaning the results gathered display a misleading positive bias. Even when this is considered Travis argues that other sources of bias are often present with very few questionnaires undergoing a test of reliability and validity.

Wiklund, Thurott & Dumas (1992) found that test subjects are reluctant to be critical when using a rating scale. This behaviour is evident even when test subjects are encouraged to be critical leading Wiklund *et al* (1992) to ponder this phenomenon. Wiklund *et al* (1992) ascribed this behaviour to test subjects wanting to be "positive" or to not "hurt the feelings of the person conducting the test".

Seeking to combat this bias, Microsoft developed the Desirability Toolkit (Benedek & Miner, 2002) consisting of 118 product reaction cards containing an equal proportion of "positive" and "negative" adjectives. Using this method, test subjects are asked to sort through the cards and select the five that most closely match their experience with the interface during the test. These selected keywords form the basis of a semi-structured interview. From his experiences using this method Travis states that it is easier to elicit negative comments as it appears to give test subjects the permission to be critical of the system. This is mirrored during the following interview, with test subjects more likely to offer criticism as they may place a negative spin on a positive adjective.

Travis (2008) has refined this technique by reducing and customising the set of adjectives used based on the context of the interface tested. Two methods of evaluation are proposed; firstly the selected keywords can be used to form a word-cloud from the responses of all test subjects. This displays graphically the range of selected keywords with the font size of the keyword proportional to the number of times selected. A second method involves logging the number of positive and negative words used to describe the interface by the test subjects in the post-test interview.

In summation, this research provides clear guidance that can be transferred to the evaluation of audio industry interfaces. In designing usability tests practitioners must first consider the context of use and select a suitable scenario within which the interfaces can be evaluated. Tests must include a measure of efficiency, effectiveness and satisfaction. By recording a log of user interaction during the test efficiency can be measured in terms of task completion time and effectiveness as accuracy. Due to its subjective nature satisfaction is harder to measure meaning a range of methods are required to accurately gauge this facet. Asking test subjects to rank candidate interfaces in terms of preference will provide some insight but greater discernment will be gained combining this with a semi-structured post-test interview. The use of keywords which are analysed both quantitatively and qualitatively in displays such as wordclouds are also considered useful for interface evaluation.

5 Case Studies

This chapter provides examples of the design and evaluation methods employed in the development of several interfaces for music technology. The aim of this chapter is to elicit a suitable range of techniques that can be used to design and evaluate the novel interfaces developed in this project.

5.1 Cuebert: A New Mixing Board Concept for Musical Theatre

In the development of a new mixing console for the live theatre environment Liebman, Nagara & Spiewla (2010) adopted a user centred approach. Initially, a group of professional theatre engineers were consulted in a series of semi-structured interviews. Notes were taken and organised by theme to create an affinity diagram. An affinity diagram is a hierarchical diagram often used in rapid contextual design to identify all issues faced by a user group. It is favoured by designers as it helps quickly identify all issues faced by the user population (Holtzblatt, Wendel & Wood, 2005).

Based on this diagram a series of paper prototypes were developed and refined to create a candidate interface design. Liebman *et al* (2010) favoured paper prototypes as they were quick to create, flexible and reconfigurable, enabling interface ideas to be mocked-up in full scale for scrutiny, comparison and task simulation.

A more detailed prototype was then developed using a redundant Tascam M-3700 mixing console with graphical representations of controls printed on paper and placed on top of the console. The core visual features and controls from the paper prototype were then created in presentation software and projected onto the console. In this proposed interface the design maintained the inclusion of faders and replaced knobs with touchscreen controls to create a digital solution that largely followed the layout of analogue consoles. The EQ section favoured a DAW style EQ curve interface featuring touchscreen nodes and a bank of rotary encoders were provided for control.

Scenarios for each phase of usage identified by the affinity diagram were simulated by creating graphical animations. These animations were then projected on to the console to create the illusion of a working interface. A user was then videoed performing mixing tasks to further illustrate how the proposed interface would work. These videos were released on the internet to the target user community via a web forum to gain feedback. This feedback was used to provide a framework for further development.

The advantage of this approach is that it enabled the interface to be prototyped relatively quickly in comparison to software and hardware prototyping. The disadvantage of this approach is the absence of quantitative evaluation. Also the users didn't actually get to use the interface personally and experience how it felt to use it – they only experienced it by proxy.

5.2 Knobs vs Faders

Gelineck & Serafin (2009) considered audio interface design "from the bottom up". They evaluated user preference for the simplest traditional controls used by the audio industry, i.e. knobs and faders.

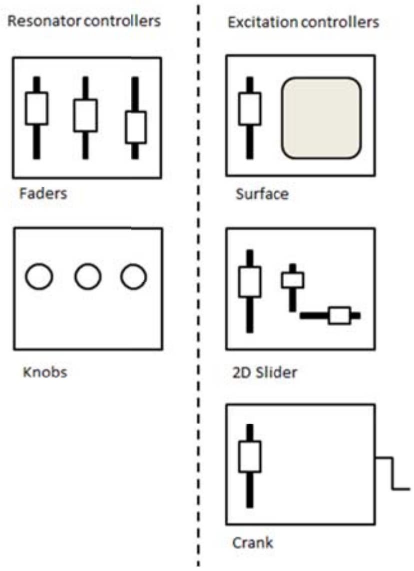

Figure 44: The range of interfaces evaluated by Gelineck & Serafin (2009)

The interface was made up of two parts comprising of a resonator controller and an excitation controller an as illustrated in Figure 44. The two parts were connected together to form one overall interface. Each combination of resonator and excitation controller was tested.

A group of test subjects were given seven minutes to freely explore the interfaces before being asked to try and recreate a reference sound. The subjects were allowed three minutes to attempt each recreation. A log of user interaction was captured for each test together with a Likert scale rating of how well the user had reproduced the reference sound. This rating was calculated as an average of Gelineck & Serafin's (2009) rating and an expert user rating. Finally subjects were asked to complete a post-test questionnaire ranking nine statements using a Likert scale.

This research is particularly relevant to this project as it used a range of methods to evaluate musical interfaces inspired by HCI research involving both qualitative and quantitative evaluation.

5.3 Multitouch Interface for Audio Mixing

Carrascal & Jorda (2011) argued that no significant usability studies have been conducted on mixing consoles and question whether they really are fit for purpose. Their work sought to address this oversight through the development and evaluation of a control interface for the mixing console. A group of six test subjects were selected for a preliminary usability test consisting of five novice users and one expert user. An even gender split of subjects was selected to help represent a cross section of the target user group.

The test consisted of a comparison between a traditional digital console (Yamaha 01V96) and the novel multitouch interface developed. The basic working principle of each interface was explained before the test and subjects were asked to mix a 4 track song consisting of percussion, guitar, piano and voice. The subjects were requested to achieve a certain spatial positioning of each instrument. In designing the test, Carascal & Jorda pondered how best to measure effectiveness as mixing audio is very subjective in nature. They chose to measure this facet of usability by recording task completion time after asking subjects to take as long as they wanted with each interface to create a satisfactory mix. At the end of the test the subjects were asked to rate each interface with a Likert scale of 1 to 5. User interaction, comments and observations were also recorded for evaluation.

5.4 Towards a more Flexible and Creative Music Mixing Interface

Gelineck et al (2013) divided evaluation into two stages to help steer the development of a novel music mixing interface. Firstly, an initial functioning prototype interface was developed that displayed audio channels as graphical circles termed "interactive widgets" on a 2D stage. Several parameters including compression, reverb and EQ could be altered by changing the length of coloured sliders.

Figure 45a: Multitouch

Figure 45b: Tangible objects on all widgets

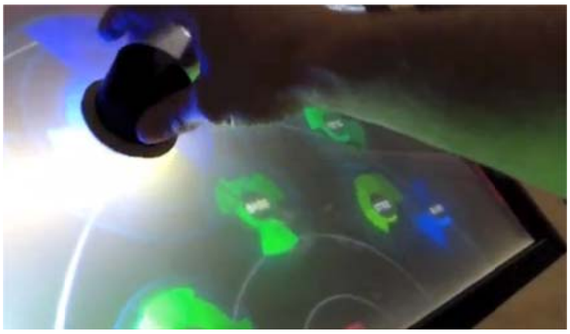

Figure 45c: Smart tangibles

Figure 45: Range of methods employed to enable user interaction with the prototype mixing interface. Screenshots taken from a video of the prototype interface in use http://www.youtube.com/watch?v=9WAI3FCzugE

Three different methods of user interaction were developed as illustrated in Figure 45. These included:

1) Multitouch: as shown in Figure 45a. Users interacted directly with the widgets via the touchscreen.
2) Tangible objects on all widgets: as shown in Figure 45b. Plastic cylinders that resemble knobs controlled the interactive widgets. In this instance each widget had its own tangible object.
3) Smart tangibles: as shown in Figure 45c. One tangible object was used to select and control all interactive widgets on the stage, a process termed "drag and drop".

An exploratory focus group of expert users were asked to use the interface with each interaction method to mix two eight track songs as a group whilst considering key design fundamentals. The panel then ranked the interaction methods in terms of preference and provided comments.

The second strand of evaluation involved conducting a preliminary test to assess the usability of each interaction method. Fourteen test subjects were given the task of changing the position and one parameter of the widgets with each test lasting 90 seconds. The goal was to keep the widgets as close to their corresponding moving ghosts as possible. To provide a benchmark, a fourth interface was included in this test which featured a mouse as the method of interaction. A tracking error score was calculated as the root-mean-squared Euclidean distance for all widgets to their corresponding ghost for each participant in each of the four conditions. This method was also employed by Hughes & Wakefield (2012) in their evaluation of user interfaces for discrete sound placement in film and TV post production. Finally the test subjects ranked the control tools in order of preference.

5.5 Battle of the DJs

In the development and evaluation of DJ interfaces, Lopes, Ferreira & Pereira (2011) elicited design guidelines by evaluating the three most commonly used types of interface; traditional, virtual and hybrid. This study directed Lopes *et al* to develop a prototype multitouch interface.

The prototype was based on the recommendations of four expert users who were regularly consulted during the design process. The prototype was inspired by Lemur (Jazzmutant, 2013) and enabled users to customise the layout of the interface to suit their own preference. The controls were displayed on the interface as graphical representations of turntables, faders, killswitches, knobs and a crossfader.

Four amateur and six professional test subjects were selected who were either radio, club or scratch DJs. A usability test was conducted comprising of three stages. Firstly test subjects completed a pre-test questionnaire to determine their relative experience. The subjects were then asked to complete four tasks with the traditional, virtual and hybrid and prototype multitouch interfaces. Each task consisted of mixing and beat-matching pairs of songs of similar style but differing tempi. Prior to the test the songs used were critiqued by an expert DJ to ensure they had the same difficulty level and were capable of being mixed with other songs. The pairs of songs were randomly selected from a pool of approved material to ensure the test subjects could not memorise pitch values during the test. One final task was included which involved the test subjects using the prototype interface with their own audio material. The time for task completion was recorded together with a video and audio recording of each subject performing the tests. Finally, users were interviewed individually and contents of the discussion transcribed. In the evaluation the average task completion time and standard deviation for each interface was calculated and compared in terms of task and user group.

5.6 Summary

The case studies provide the following considerations in the design and evaluation of a music technology user interface:

- Traditional interfaces should be evaluated prior to developing a novel interface
- An expert user should be selected and consulted to help guide the design process
- Paper prototypes should be developed and refined prior to software and hardware development
- Usability tests should include a pre-test questionnaire to ensure the test subjects selected meet the requirements of the selected user group skill level
- Measurements should be taken to evaluate effectiveness and efficiency quantitatively
- Satisfaction can be evaluated in terms of user preference and analysis of comments and observations
- A benchmark interface should also be tested to provide a comparison with the novel interface
- Some elements of the test should be randomised to ensure accurate evaluation

6 Dissection of the EQ Task

6.1 Introduction

Following Crystal & Ellington (2004) this chapter presents a task analysis of mixing audio with specific regard to EQ and is based on a study of relevant literature. As this project is restricted to developing an intuitive EQ user interface for the mixing console for mono sound source inputs to produce a mono mix output, other key mixing techniques such as panning, dynamic processing (compression and gating) and reverb are omitted.

Owsinski (2006) cited by Quarz (2013) defines the process of EQ as a broad objective (i.e. goal) with the user working in the frequency domain, to achieve a "true mix of timbre, bass, treble and midrange." EQ techniques can be broadly categorised as either corrective or creative in nature. Corrective techniques are concerned with treating specific spectral issues found in individual audio channels and creative techniques are employed when considering multiple audio channels to create a cohesive and interesting mix.

6.2 Corrective EQ

Resonant frequencies are often captured during the recording process. These are characteristically dominant frequencies that can colour the overall tone of a sound source unpleasantly. They occur for many reasons, including the following causes in isolation or in combination:

- the frequency response of the microphone used to capture the recording
- the placement of the microphone relative to the instrument
- the acoustic properties of the room e.g. a resonance in an acoustically untreated room
- the physical properties of the sound source recorded e.g. the "ring" of a snare drum

Gibson (2008) recommends using a peaking parametric filter for treating resonant frequencies and provides a step-by-step solution. This involves firstly setting a high Q and gain and then sweeping across the frequency spectrum to find the resonance. Once identified, the gain is attenuated to remove the resonance. Some expert users will not need to sweep the frequency spectrum and will be able to get close to the resonance based on their experience.

Another problem tackled through corrective techniques is sibilance. Sibilance is defined as a "high frequency whistling or lisping sound that affects vocal recordings" (White, 1997, p. 323). Izhaki (2008) recommends using a low pass filter at a high frequency to remove this effect.

Mynett, Wakefield & Till (2010) identifies high end noise or "hiss" as another corrective issue that can be treated by employing a low pass filter.

6.3 Creative EQ

The biggest issue an engineer faces with regards to EQ when mixing multiple audio channels is the effects of masking. Mynett et al (2010) defines masking as “the ability of one sound to obscure, or inhibit, (i.e. mask) the frequencies of another sound” and advises that masking especially occurs in a dense mix and is most pronounced in the lower regions of the frequency spectrum (Mynett 2012).

One technique that can be used to combat this phenomenon involves sharing out the frequency spectrum. This can be achieved by the removal of unnecessary frequencies. Mynett et al (2010) suggests using high-pass filters on all audio channels to remove anything below 60Hz. In many cases all frequencies below the audio channel’s lowest frequency can be cut freeing up the lower part of the frequency spectrum for the instruments that really belong there. Examples include the treatment of cymbals or hi-hats. Mynett et al (2010) also suggests using low pass filters to mark the channel’s upper boundary but acknowledges this technique is not as commonly used in mixing. Once the main effects of masking have been treated with the pass filters, the remaining essential areas can be selectively treated with peaking, and shelving filters to accentuate pleasant frequencies.

Additionally, Mynett et al (2010) documents a more localised technique for the sharing of the frequency spectrum termed mirrored EQ. Two example scenarios are provided to help explain the technique. The first involves a snare drum which lacks definition. Instead of boosting the snare channel with a peaking parametric filter Mynett et al (2010) suggests cutting other the channels at around 200Hz. The second scenario involves a rhythm guitar and vocal channel competing for space. Mynett et al (2010) advises the engineer to find and boost the pleasing frequency range in vocal and cut the same range in the rhythm guitar.

Gibson (2008) and Mixerman (2010) provides more general advice with regards to creative EQ which again favours cutting as opposed to boosting frequencies. Specifically, they suggest working from the low to high frequency cutting “muddiness” from 100Hz to 300Hz with a high shelving filter, cutting “irritating mids” from 800Hz to 5kHz with a wide Q peaking parametric filter and adding “brightness” with a low shelving filter from 5kHz to 8kHz.

7 Simple Corrective EQ Interface

7.1 Design

In the design of a corrective EQ interface to treat a single mono recording of a musical instrument the project supervisor acted as the expert user. Regular design meetings were conducted to help steer development and provide further guidance and insight.

The motivation behind the design of these interfaces was to represent the recorded instrument considered in terms of its spectral content. Initial discussions focussed on the merits and limitations of displaying real time and static spectral information. It was agreed to focus on a static visualisation because a rapidly changing display was considered to be confusing to the user and unsuitable for direct manipulation. An analysis of the corrective EQ task documented in Chapter Six indicates that a commonly performed task involves the targeting and attenuation of resonant frequencies. For these reasons the novel interfaces developed presented the frequency spectrum as static spectral information with peak frequencies identified and presented for modification.

Different design ideas and controls were initially sketched on paper and their relative merits and suitability for the task considered discussed. Many of these ideas were found to be unsuitable and quickly disregarded when task based scenarios were considered. This led to four potential designs being selected for realisation. Mock-ups of the most commonly used EQ interfaces currently featured in audio products were also developed to provide a benchmark against which to evaluate the new interfaces.

The interfaces developed included:

Figure 46: Interface 1

Interface 1 replicated the design of a typical SSL style channel strip, with two bands of peaking parametric filters for high and low mid frequencies. It also featured a high-pass, low-pass, high shelving and low shelving filter (see Figure 46).

Figure 47: Interface 2

Interface 2 replicated the design of a typical DAW style channel strip, again featuring high and low pass filters with the addition of four peaking parametric filters that are not band limited. The EQ curve was displayed for user feedback (see Figure 47).

Figure 48: Interface 3

Interface 3 had the same EQ controls as Interface 2 and replicated the typical recent advances in EQ plugins, with real time spectral information displayed behind the EQ curve (see Figure 48).

Figure 49: Interface 4

Interface 4 had the same EQ controls as Interfaces 2 and 3, however in this implementation the spectral information was presented as a static plot behind the EQ curve (see Figure 49). This interface was developed to enable the static spectral visualisation to be presented in an interface that featured the same controls as traditional solutions.

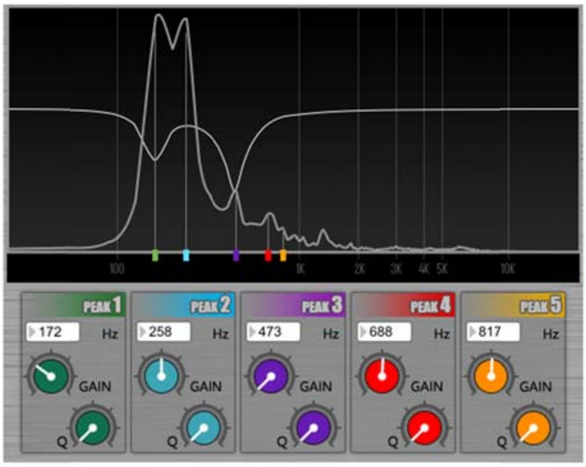

Figure 50: Interface 5

Interface 5 displayed spectral information as a static plot together with the EQ curve as in interface 4. However this interface added five vertical lines that indicated the five peak frequencies. Each vertical line had a coloured marker at its base which corresponded to one of five coloured parametric EQ sections in ascending frequency order from left to right. Because the candidate frequencies for attenuation had been identified the user was not provide with a frequency control. The peak frequencies were displayed in a numerical format in place of the frequency control of interfaces 2, 3 and 4. Q and gain/attenuation rotary encoders were provided for EQ control (see Figure 50).

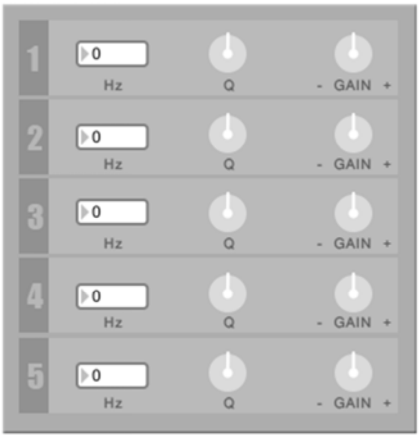

Figure 51: Interface 6

Interface 6 presented similar information to Interface 5 but did not display spectral information or the EQ curve to the user. The five peak frequencies were ordered by amplitude from top to bottom on the display. Each identified frequency was displayed numerically with corresponding Q and gain/attenuation rotary encoders next to it (see Figure 51). This interface was considered worthy of consideration because of its minimalist characteristics providing the minimum of information and controls needed to complete the task.

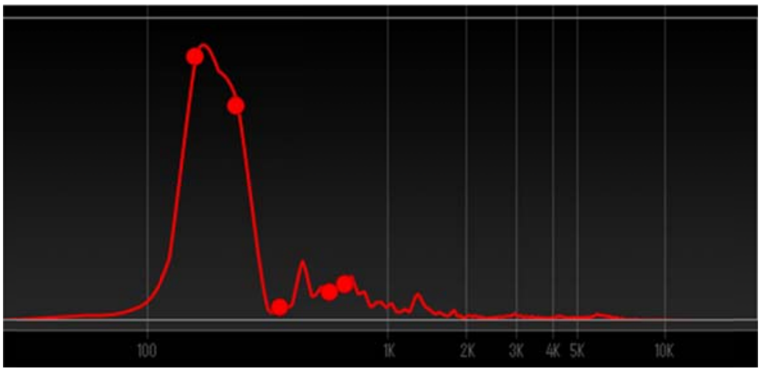

Figure 52: Interface 7

Interface 7 presented the frequency bin maximums across the audio sample under consideration as a spectral plot combined with five movable nodes. These nodes corresponded to the five peak frequencies as in Interface 5 and 6. Downward movement of the nodes provided EQ attenuation at that frequency and the nodes were restricted to movement in the vertical axis only. The spectral plot changed in real time to reflect the EQ adjustments made by the user (see Figure 52). If the user attempted to move the node above its original position the EQ was effectively bypassed. This design was considered a worthy addition as it closely adheres to the guidelines proposed by Bertin (1983), Nielsen (1994) and Norman (2002).

Interfaces 4 through to 7 represented novel ideas for the EQ user interface, with Interfaces 5 and 7 offering a more assistive interface.

7.2 Implementation

The Max/MSP Jitter environment (Cycling74 , n.d.) was used for developing the prototype interfaces as it not only provides a good platform for conducting spectral analysis but also facilitates the creation of prototype user interfaces. For these reasons together with the wealth of supporting literature and active online community it was selected as an appropriate development tool. Adobe Photoshop was selected for the creation of image assets as it is widely regarded as the industry standard tool for creating bitmap images.

7.2.1 Capturing the spectral information

The Fast Fourier Transform (FFT) is an algorithm employed to convert audio from the time domain to the frequency domain as shown in Figure 53. Brixen (2011) provides a concise explanation of how spectral analysis is performed in the FFT stating "the (audio) signal can be resolved into a set of constituents based upon the fundamental frequency. Each constituent contains information on both the amplitude and phase" (Brixen, 2011 p. 193). It is possible to plot the relative amplitude of each constituent relative to the fundamental frequency as a curve.

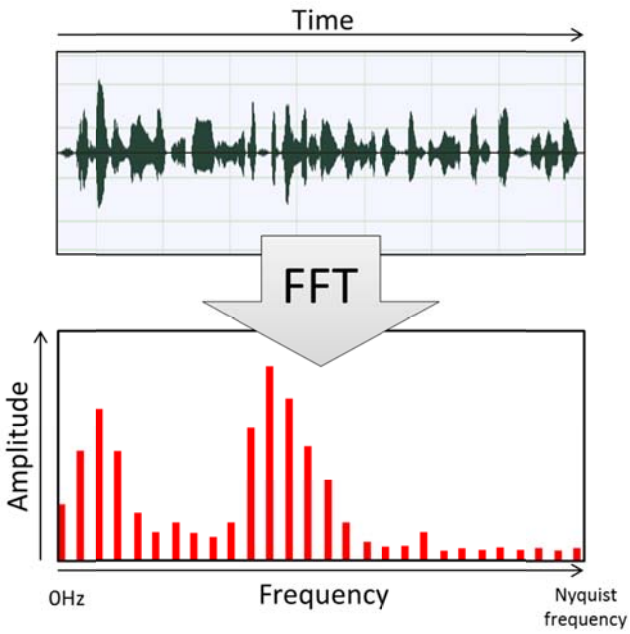

Figure 53: FFT Time to frequency domain

The number of constituents or samples taken in each frame is termed the FFT size. Samples are often visualised as being captured in frequency bins. The distance between each bin is termed the frequency resolution and is measured in Hz. Adhering to the Nyquist sampling theorem, which states that the sampling frequency must be at least two times the highest audio frequency to be reproduced (Pohlmann, 2011) it is the combination of the sampling frequency and FFT size that determines the frequency resolution of the FFT. This relationship can be explained by the following expression (Brixen, 2011):

$$R = F_s/N \tag{2}$$

where:

R = frequency resolution in Hz
F_s = sampling frequency in Hz
N = the number of samples (FFT Size)

In FFT analysis a windowing function is often utilised to prevent errors in data capture occurring when the audio sampled is longer in duration. This windowing function could be described as a sort of “fade in” and “fade out” with consecutive windows often being overlapped by a specified percentage. One commonly used function in FFT analysis is a Hanning window with 50% overlap.

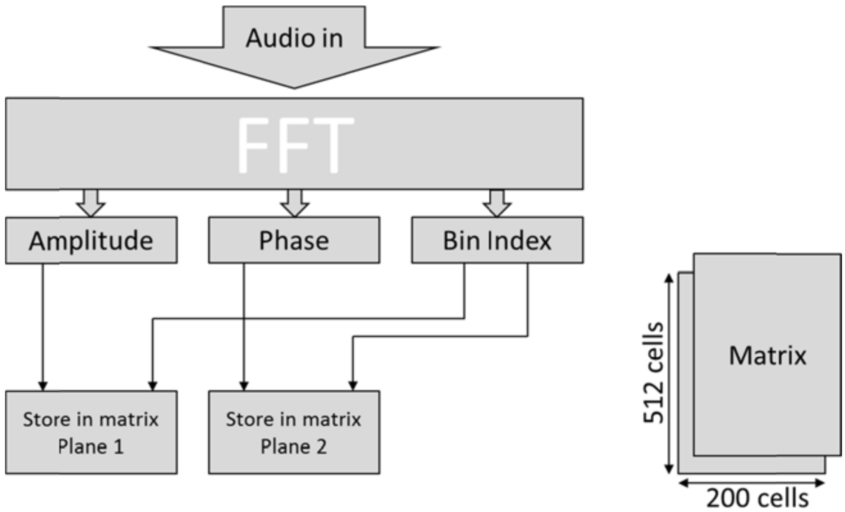

Figure 54: Capturing the spectral information flow diagram

The algorithm used to capture the spectral information is presented as a simplified flow diagram in Figure 54 and is based on a method proposed by Charles (2008) in his spectral processing tutorial. In this implementation the audio sample undergoes FFT analysis using the *pfft~* object with the results stored in a Jitter matrix. The dimensions of the matrix are set to match the FFT window size i.e. 512 cells (corresponding to 512 bins) by 200 cells (corresponding to 200 FFT frames). Various objects were required in Max/MSP to overcome its idiosyncrasies such as employing a *cartopol~* object to convert the output of the FFT from cartesian to polar coordinates. This implementation uses a 1024-point FFT with a Hanning window function with 50% overlap. Higher resolution FFTs were also trialled during implementation but were not selected partly because of the limited processing power of the computer employed and partly because the spectral information gained from the 1024-point FFT was deemed to be sufficiently detailed to be used in this context.

7.2.2 Maximum amplitude vs average amplitude

The FFT bin maximums for the audio segment analysed were selected to generate the static spectral display in preference to FFT bin averages for two reasons. Firstly, it was felt that with regard to EQ, and particularly corrective EQ, sound engineers are generally concerned with dominant frequencies. This was confirmed after consultation with the expert user and comparing both average and maximum values for several audio segments. Secondly, unless a gating algorithm was implemented, using the bin averages as the basis of a visualisation would prove misleading because audio signals generally contain periods of silence. Clearly further research could be conducted to experimentally determine whether this is the best approach and evaluate alternatives.

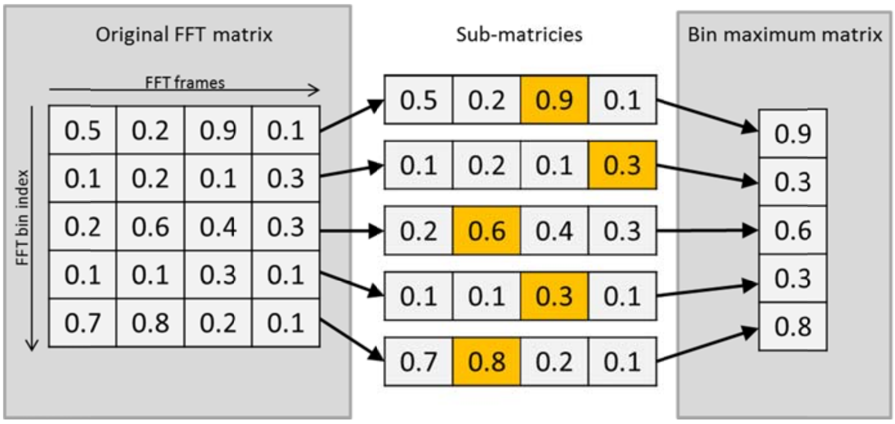

Figure 55: Identifying the bin maximums

Figure 55 outlines how the bin maximums were identified from the original stored matrix. The original FFT matrix shown in Figure 54 was split into 512 bin specific sub-matricies. Using the *jit.3m* object it was possible to identify the bin maximums in each sub-matrix. A new jitter matrix was employed to store these bin maximums in 512 cells.

7.2.3 Peak detection

Sound sources usually contain several peak or dominant frequencies, spread across the frequency spectrum, that define their character. To determine these spectral peaks from the bin maximums a simple algorithm was employed to query the bin maximum matrix. A bin was considered a peak if its value was higher than the bin either side of it. This was simply achieved using an *if* statement with peak bin maximums flagged and stored as a list of values and corresponding bin indices. This list was then sorted in order of magnitude using the *coll* object and the top five values and bins stored in a jitter matrix 2 cells wide by 5 cells high.

A total of five peak frequencies were selected for two main reasons. Firstly, the traditional interfaces developed contain four or five filter bands. As the automatically detected peaks effectively replace the traditional EQ controls in the novel interfaces, it was considered good practice in terms of consistency in the evaluation. Furthermore, during implementation displays with a greater number of peak frequencies displayed were trialled; however in all of the audio sources considered the resonant frequency was identified within the first five peak frequencies. Further research would be necessary to determine if this was always the case for different audio sources.

7.2.4 The EQ algorithm

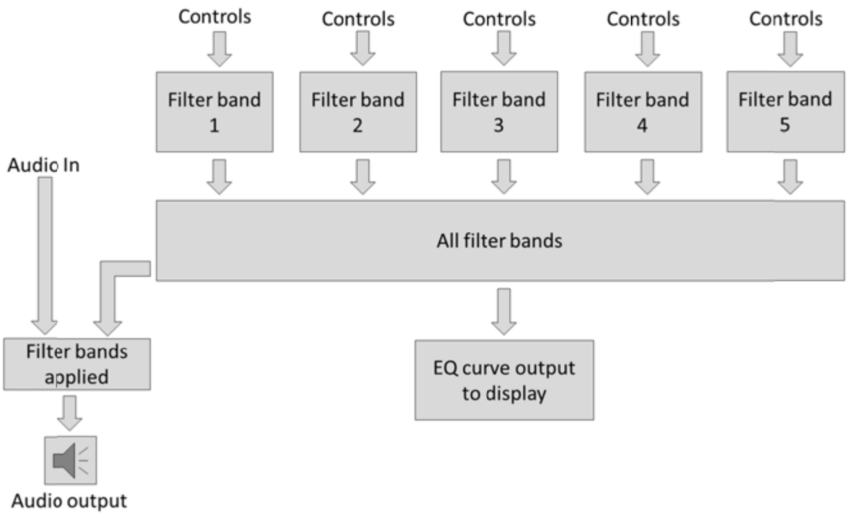

Figure 56: Flow diagram of the EQ algorithm employed in interfaces 1 to 6

Figure 56 outlines the EQ algorithm employed in interfaces 1 to 6. Each filter band was set using a *filtergraph~* object with the controls provided in the interface. These *filtergraphs~* were then combined to create one EQ curve using the *cascade~* object. The output EQ curve was used to modify the audio channel considered using the *biquad~* object and was also displayed as a plot in interfaces 2, 3 4 and 5.

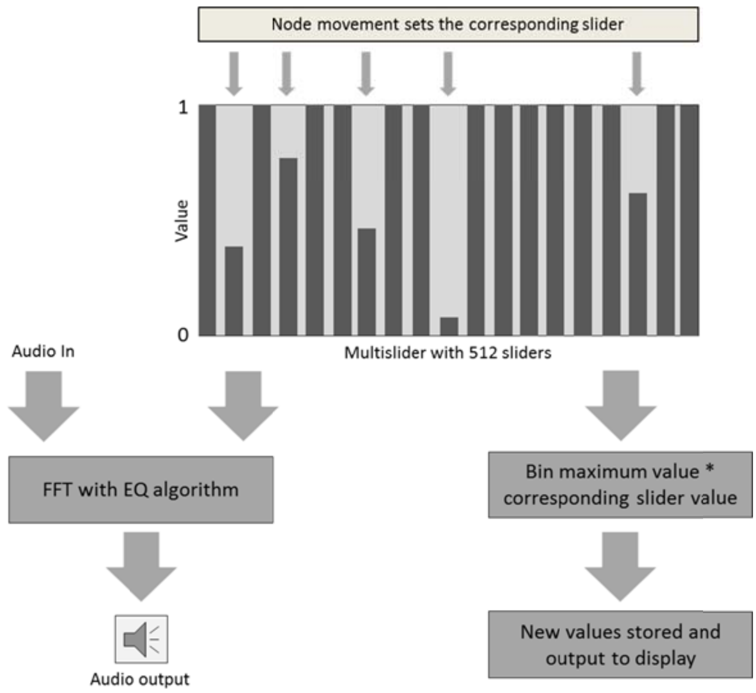

Figure 57: Flow diagram of the EQ algorithm employed in interface 7

The EQ algorithm employed in interface 7 is based on the Forbidden Planet patch, provided in the Max/MSP install (Cycling74, n.d.) and is outlined in Figure 57. This method is conceptually similar to a graphic EQ and uses a *multislider* object with 512 sliders (one slider per FFT bin). When a node is moved on the display the new position relative to the original node position is used to calculate the amount of attenuation to be applied. This attenuation value (0 to 1.0) is used to set the corresponding slider value.

The *mutislider* is then used to perform two operations:

1. The *multislider* values are used to perform filtering in the frequency domain as follows. The *multislider* values are stored in a *buffer*. The audio signal for the track considered is passed into a *pfft* object and the *buffer* used to perform the desired filtering by multiplying the amplitude values for the FFT bin considered using the *~ object. These are then converted back into the time domain by the *fftout* object and the modified sound is output.
2. The bin maximum values are multiplied by the multislider values and stored in a new Jitter matrix. These new values are then output to the static spectral plot display to reflect the changes made.

FFT filtering was performed because the emphasis of this work is on user interface design and this was the easiest way to perform the filtering and keep the user interface display of the output frequency spectrum updating in real-time. Further work could incorporate traditional EQ algorithms.

7.3 Test

Test subjects were set the task of removing a resonant frequency from an audio file consisting of a snare drum with an artificially added resonant frequency.

The snare drum audio samples were recorded with a Shure SM57 microphone and edited to create 14 six second audio files. An enveloped sine wave was mixed with each audio file to create a snare with an obvious resonant frequency ranging from 200Hz to 2 kHz.

The EQ user interfaces ran on a laptop computer with each interface controlled by a standard mouse. The tests were conducted in a small acoustically treated studio environment equipped with a professional monitoring system. Seven test subjects were selected for this preliminary investigation all of whom were familiar with, and had at least basic experience of, mixing music being either industry professionals, Music Technology lecturers or Music Technology students.

The test subjects were introduced to each interface in turn. The users were given a verbal description of each interface and any queries were clarified before and during a practice test. Users were then asked undertake the test for real and to identify and remove the resonant frequency. A different snare sample was used in each practice and test and the order in which interfaces were presented to each test subject was randomised. The effectiveness of each interface was measured by logging the user interaction with the interface using the Max/MSP external IPEM text logger developed by Moens (2011). The efficiency of each interface was measured by timing how long it took each user to remove the resonant frequency and how the user interacted with each EQ user interface. The users were videoed to provide supplementary information to the user interaction logs. A user satisfaction preference ranking was conducted at the end of the tests. The test subject's comments and interactions with each interface were video recorded for the duration of the test.

7.4 Analysis of Results

7.4.1 Effectiveness and efficiency

In all tests all users correctly found and removed the added resonant frequency with each interface. All interfaces tested are therefore deemed 100% effective for this specific task.

Normalised task completion time (NTCT)							
Interface	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Subject 1	1.00	0.65	0.65	0.48	0.17	0.29	0.12
Subject 2	1.00	0.46	0.30	0.36	0.13	0.11	0.31
Subject 3	0.53	0.76	1.00	0.36	0.22	0.49	0.35
Subject 4	0.30	1.00	0.37	0.48	0.40	0.21	0.17
Subject 5	0.86	1.00	0.81	0.76	0.38	0.90	0.14
Subject 6	0.93	0.76	1.00	0.82	0.53	0.48	0.12
Subject 7	0.84	0.78	0.83	0.36	0.36	1.00	0.45
Average	**0.78**	**0.77**	**0.71**	**0.52**	**0.31**	**0.50**	**0.24**
Std. Dev.	0.26	0.19	0.28	0.20	0.14	0.34	0.13

Figure 58: Normalised task completion time (NTCT) for each user and interface

The normalised task completion time (NTCT) was calculated for each user to enable the relative task completion times to be compared between the test subjects. Figure 58 shows the NTCT for each user to complete the task using each interface and presents the average and standard deviation of the NTCT for each interface. Figure 59 presents this data as a plot of the average NTCT to complete the task for each interface. The results suggest that users were able to perform the task faster with the new interfaces (i.e. Interfaces 4 to 7) with Interface 7 the fastest followed by Interface 5. On visual inspection there appears to be a statisitcally significant reduction in NTCT with interfaces 5 and 7 over interfaces 1, 2 and 3.

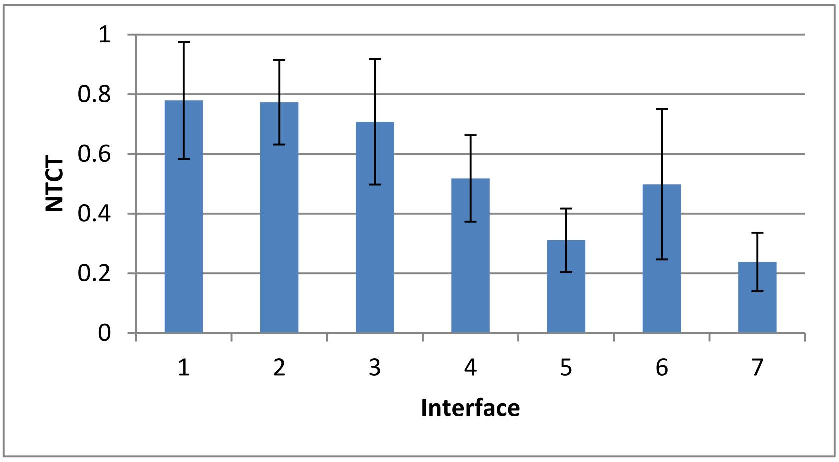

Figure 59: Average normalised task completion time (NTCT) for each interface with 95% confidence intervals

A repeated measures ANOVA was conducted to confirm this finding. There were no outliers in the data, as assessed by inspection of a boxplot. NTCT was normally distributed for each interface tested, as assessed by Shapiro-Wilk's test ($p > .05$) with the exception of interface 4 where the data violates normality as $p = .041$. This could be due to the small number of test subjects used in this investigation. As normality was only marginally violated in one of the interfaces tested, it was worthy of note but did not prevent analysis with a repeated measures ANOVA. Mauchly's test of sphericity indicated that the assumption of sphericity had not been violated, $X2(2) = 17.493$, $p = .770$. Importantly, NTCT was statistically significantly different for the different interfaces evaluated, $F(3.646, 21.876) = 7.498$, $p = .001$ partial $n2 = .556$. Post hoc analysis with a Bonferroni adjustment revealed that NTCT was significantly decreased when comparing interface 5 with interface 2 (0.46 (95% Cl, 0.74 to 0.19), $p = .003$) and comparing interface 7 with interface 2 (decrease of 0.56 (95% Cl, 1.03 to 0.04), $p = .034$).

Further analysis of the log files and video footage highlighted distinct periods of thinking and activity, with a period of cognitive processing being followed by user being observed making changes via the interface. Often there were multiple alternating periods of thinking followed by activity. To investigate this further we defined a thinking period as a period of time one second or longer in which the test subject makes no EQ changes and visually appears to be thinking and produced the plots shown in Figure 60.

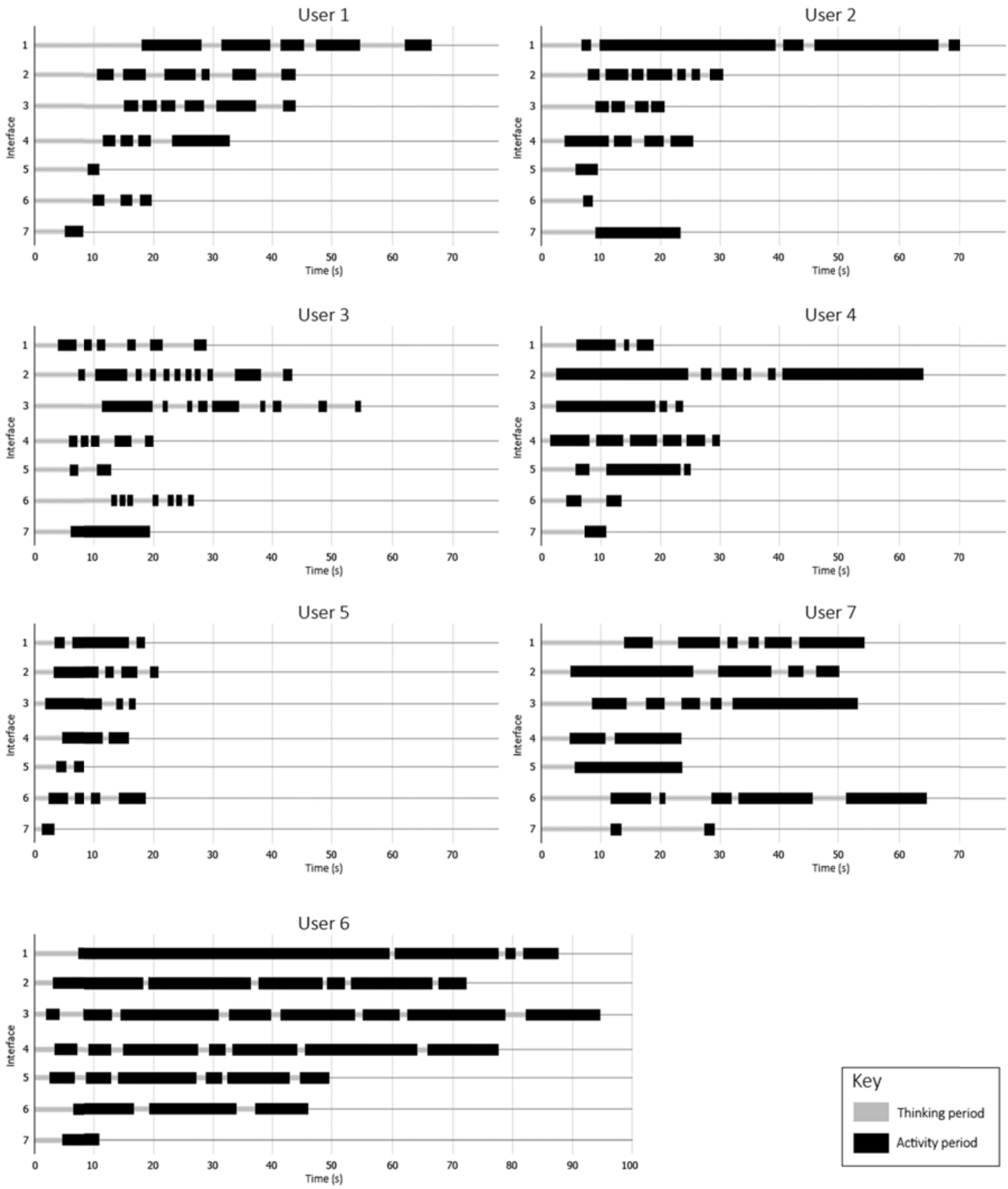

Figure 60: Timelines of test subject thinking and activity

On analysis, all subjects started the test with a period of thinking and ended the test with a period of activity. This results in an equal number of thinking and activity periods for all subjects. Figure 61 shows the number of periods of thinking followed by activity for all test subjects for each test and the average number for each test. Figure 62 presents this as a plot of average number of periods of thinking followed by activity.

Interface	Number of periods of thinking followed by activity							Average number of periods
	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	
1	5	5	6	3	3	4	6	4.57
2	6	7	11	6	4	6	4	6.29
3	6	4	9	3	3	8	5	5.43
4	4	4	5	6	2	7	2	4.29
5	1	1	2	3	2	6	1	2.29
6	3	1	7	2	4	3	5	3.57
7	1	1	1	1	1	1	2	1.14

Figure 61: Number of periods of thinking followed by activity

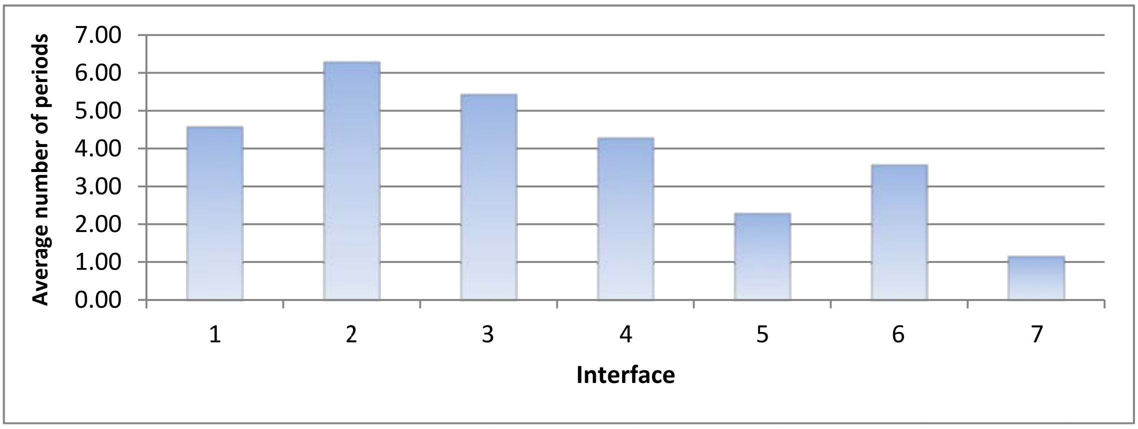

Figure 62: Plot of average number of periods of thinking followed by activity for each interface tested

Counting the number of periods of thinking followed by activity appears to show the advantages of the new interfaces over the exisiting interfaces. This count provides further clarification of the increased usability offered by interfaces 5 and 7. In addition to Interface 7 having the best average number of periods, it is worth noting that six of the subjects tested had 1 period of thinking followed by activity and one subject had 2. With Interface 5 three subjects tested had 1 period of thinking followed by activity. None of the non-assistive interfaces (i.e. Interfaces 1 to 4) had any instances of 1 period of thinking followed by activity. Arguably, if a user only has to think once and act once that is a sign of an effective interface. Additionally the identification and use of this metric in user interface evaluation represents a new addition to interface evaluation research that may be of use in evaluating other types of interface.

7.4.2 Satisfaction

Figure 63 shows the user preference ranking and average ranking for each interface. Figure 64 presents the average user preference ranking for each interface graphically. Overall, Interface 5 is the preferred implementation with 4 out of 7 test subjects putting this interface as their first choice. It is interesting to note that Interface 5 is visually similar to what users are most used to whilst providing extra assistance in performing the task.

Interface 7 was ranked 4th overall in the preference ranking. However it is interesting to note that 2 users ranked this interface as their favourite and 2 users ranked it as their least favourite. This polarization in preference could be attributed to the fact that interface 7 is a departure from the norm and also hints at the difficulty of introducing radically different interfaces and achieving acceptance from all potential users.

Interface 6 was ranked 5th equal overall and its low ranking could be due to the lack of visuals which are common in modern EQ user interfaces, again hinting at a reaction to change.

	User Preference Ranking							
Interface	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Average Ranking
1	7	7	7	6	7	6	4	6.3
2	6	6	6	3	4	5	3	4.7
3	3	5	5	2	2	4	1	3.1
4	4	4	4	4	3	3	2	3.4
5	1	3	2	1	1	1	5	2.0
6	5	2	3	5	5	7	6	4.7
7	2	1	1	7	6	2	7	3.7

Figure 63: User preference ranking of interfaces

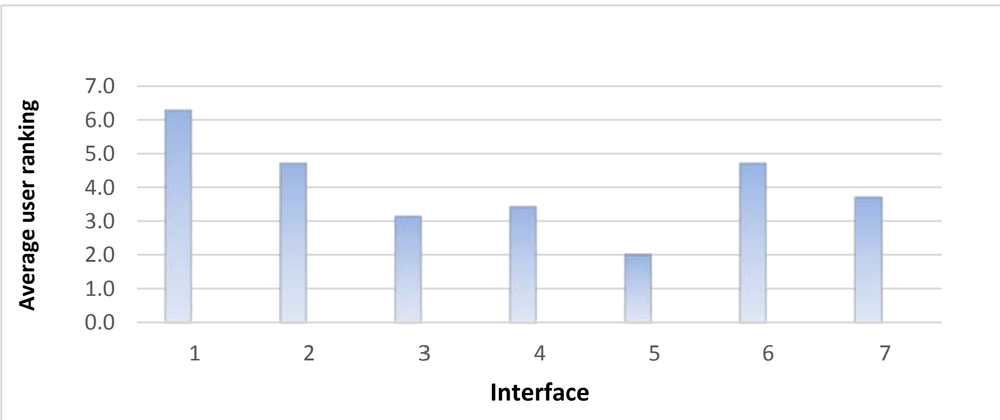

Figure 64: Average user preference ranking of interfaces

7.4.3 User comments and observations

All subjects tested were observed setting a high Q, increasing the gain and sweeping to find the resonant frequency in Interfaces 1 to 4.

Only two subjects tested used the high pass and low pass filter in Interfaces 1, 2 and 3 in the removal of the resonant frequency.

Five users commented that they were most familiar with interface 2, using this type of EQ regularly in their work. One user commented that the real time spectral plot provided in interface 3 acted as a "sanity check" as resonant frequencies can often be seen to tail off slower than the dominant frequency components.

Some users commented that they did not like being restricted to only five peak frequencies with interfaces 5 to 7, with one user passionately resisting the imposed restriction. This user was observed becoming rather stressed during the tests with unfamiliar interfaces.

One user commented that he was able to identify the resonant frequency in interface 6 by looking at the list of peak FFT bin maximums and found ordering by amplitude useful. All other subjects tested preferred the more conventional approach of ordering the peaks by frequency from low to high.

Several users recognised the benefits offered by interfaces 5 - 7 with one highly experienced user commenting that if he became more familiar with the interface and the peak frequencies identified proved reliable, he would happily switch to such a tool. One user commented that on reflection he was surprised that interfaces similar to interface 5 have not already been developed. One user commented that the assistive interfaces presented would be ideal for novice engineers and untrained bedroom producers.

7.5 Conclusions

Four novel EQ user interfaces specifically designed for the corrective EQ task have been introduced in this investigation and undergone preliminary HCI usability testing.

Of the three novel assistive interfaces, Interface 7 appears to be the best interface to complete a specific corrective EQ task based on average NTCT and average number of periods of thinking followed by activity to complete the task. Statistically, this interface has a lower average NTCT compared with interface 2. Interestingly this interface was only mid-rated on the user preference ranking due to a polarization of view amongst the test subjects.

The test subjects clearly preferred Interface 5. This interface is close to current DAW plugin and digital mixing desk EQ interfaces and the test subjects valued the addition of candidate resonant frequency identification it provided. However the average NTCT and average number of periods of thinking followed by activity to complete the task indicated that this interface was not as efficient as Interface 7.

Interface 6 was not favoured in the preference ranking, possibly due to its lack of a display showing the EQ curve and spectral information and the results from the NTCT and the number of periods of thinking followed by activity appear to suggest it is not as efficient as the other assistive interfaces.

User comments indicated that they were largely receptive of the new EQ user interfaces presented and saw their benefits apart from the polarized views on Interface 7.

7.6 Limitations of the experiment

The above experiment had the following limitations:

- Test subjects only had a short practice period with one audio example to become familiar with each interface. Different results may be observed if users were given a period of time to work with the interfaces on real projects and became familiar with each interface.
- These preliminary tests were only conducted with small number of test subjects so a wider test is required to confirm these initial findings.
- Also tests with different audio samples are required to confirm whether these results apply to different musical instruments

8 EQ User Interface for Mixing Multi-track Recordings of Musical Instruments

8.1 Introduction

The findings of the corrective EQ usability test were used to steer the development of a multi-track EQ user interface for treating mono sound sources of musical instruments. There were measured improvements in usability when the corrective task was undertaken with the novel assistive interfaces as opposed to the traditional interfaces. Furthermore interface 5 and 7 were clearly better across the metrics evaluated than interfaces 4 and 6.

Interface 7 was selected over interface 5 for further development for several reasons. From the user's perspective interface 7 is arguably better than interface 5; based on average NTCT and average number of periods of thinking and activity From the designer's and expert user's perspective this interface most noticeably adheres to Nielsen's (1994) heuristics and Norman's (2002) design guidelines outlined in Chapter Four. This is mainly because the controls are not separated from the display. Furthermore, by directly interacting with a data visualisation of spectral content as opposed to an EQ curve the interface is not only visually simpler but the feedback provides a greater indication of what the EQ controls are actually doing in the frequency domain. For these reasons, interface 7 was selected as the starting point for developing a multi-track EQ user interface.

8.2 Design

Regular design meetings were held with the expert user to steer development and provide further guidance and insight. Initial meetings focussed on how best to refine and develop the design ideas presented in interface 7 to make it suitable for the multi-track mixing of recorded musical instruments. It was agreed to focus on designing an interface for two channels as a proof of concept which could then be scaled up to include more channels.

8.2.1 Interfaces

Fundamentally, EQ is used in multi-track mixing to blend the audio channels together to create a cohesive mix, as identified in Chapter Six. A wide range of ideas were prototyped on paper and discussed in terms of their effectiveness for assisting the user in achieving this goal. Three potential interfaces were selected for development with each solution employing a different visualisation approach to display more than one channel. A tradition DAW-style interface was also selected for development to provide a benchmark in usability testing.

Figure 65: Interface A

Interface A shown in Figure 65 replicated the design of a typical DAW style channel strip, featuring five peaking parametric filters that are not band limited. Two sets of EQ display and controls are provided side-by-side with one for each channel. The first and fifth filter control is switchable between parametric mode and low or high shelving mode respectively. User feedback is provided by real time spectral information displayed behind the EQ curve.

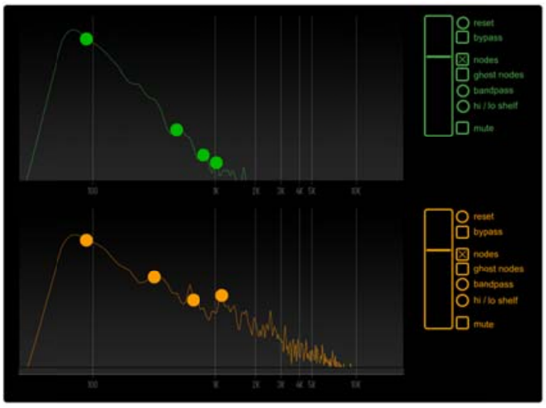

Figure 66: Interface B

Interface B shown in Figure 66 displayed the spectral plots for each audio channel separately and arranged vertically. The motivation behind this design was the multiple plot style identified by Heer *et al* (2010). The spectral plots were presented logarithmically on both axes with the vertical axis displaying magnitude as decibels (-70 to 0dB) and the horizontal axis displaying frequency (0 to 22.5kHz). Buttons and a volume control for each channel were displayed to the right of each plot.

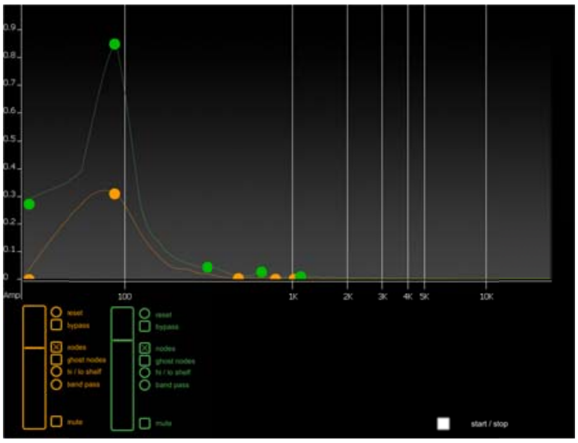

Figure 67: Interface C

Interface C shown in Figure 67 seeks to replicate the stacked graph concept identified by Heer *et al* (2010). Interface C is fundamentally different from the other novel interfaces as the spectral plots were displayed on a linear amplitude scale. Both channels were displayed together with the first channel presented as a normal linear amplitude plot. The plot for the second channel differs and was displayed as the addition of the second track's frequency content to the first track's thus representing the cumulative frequency content.

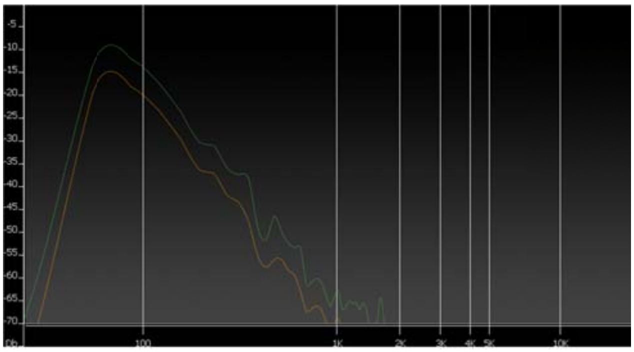

Figure 68: Interface C logarithmic implementation with two identical channels displayed

Initially during the development of interface C a decibel scale was implemented, however the logarithmic nature of this display meant that the area underneath each plot was not representative of the spectral characteristics. This is indicated in Figure 68 which shows the display for two channels of the same audio source. Clearly a logarithmic scale would provide an unrepresentative mental model of spectral characteristics displayed. Buttons and a volume control for each channel were displayed beneath the main interface.

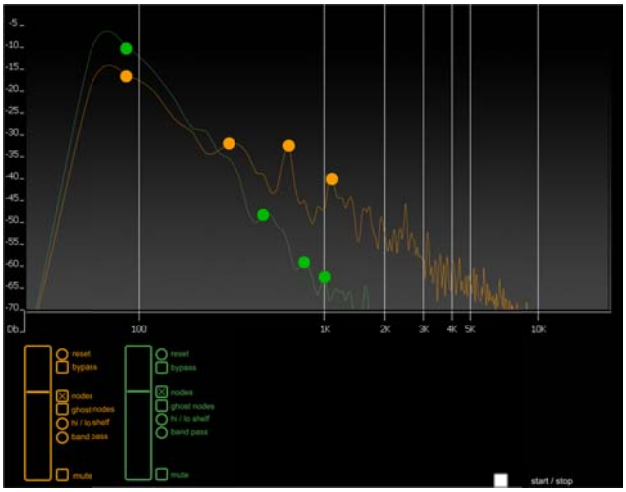

Figure 69: Interface D

Interface D shown in Figure 69 is similar in nature to interface B however the spectral plots were overlayed on one larger interface in an attempt to replicate the time-series display identified by Heer *et al* (2010). As with interface C buttons and volume fader for each channel were displayed beneath the main interface.

8.2.2 Controls

The dissection of the EQ task presented in Chapter Six provided the guidelines required to design a suite of controls that enable the user to interact directly with the spectral plots and perform multi-track EQ tasks successfully. Specifically, these tasks include:

- Sharing the frequency spectrum between channels
- Accentuating pleasant frequencies
- Performing mirrored EQ

Figure 70 highlights the controls designed to address these needs. The controls and spectral plot for each channel are given the same colour to assist user interaction.

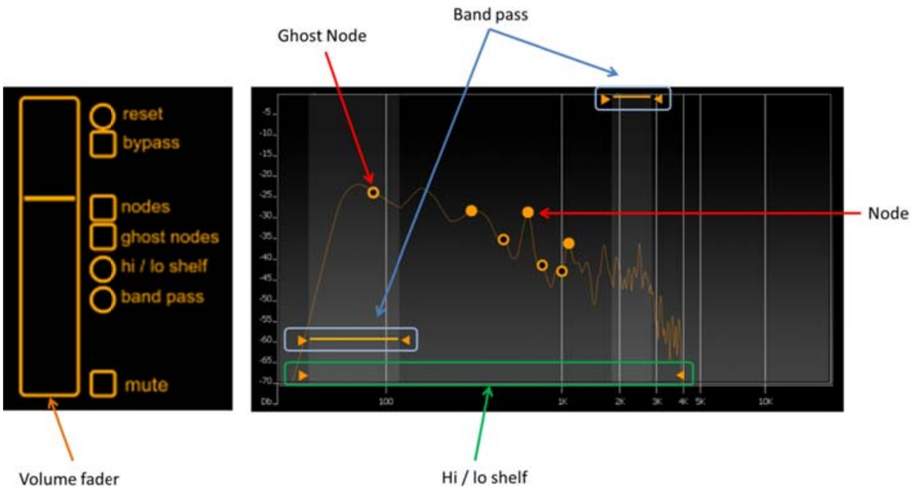

Figure 70: Multi-track EQ controls

- Nodes: these nodes correspond to the five FFT bin maximum peaks and are displayed as coloured circles. Node movement is restricted the vertical axis only with downward movement of the nodes providing EQ attenuation and upward movement providing EQ boosting; the latter being a change to the nodes featured in interface 7 in the first test.

- Ghost Nodes: these nodes correspond to the five FFT bin maximum peaks of the other track currently being mixed and are identifiable as coloured hollow circles. Ghost nodes were provided to specifically assist the user in performing mirrored EQ tasks.

- Band-pass: this control is made visible by clicking the band-pass button for the corresponding track on the display and consists of two coloured triangular markers that can be moved along the horizontal axis to select the lowest and highest frequency range for the band pass filter. A horizontal coloured line is drawn between these two markers and acts as a fader. Downward movement of the coloured line provides attenuation within the selected frequency range with upward movement providing gain. If no changes are made within a 10 second window, the band-pass control returns to the default invisible state. This was an attempt to not clutter the user interface and keep it visually simple after Norman (2002). Two band-pass sections are provided, one operating in the lower half of the frequency spectrum and the other in the upper half. These controls were considered a worthy addition as they provide a similar function to the bandwidth/Q control featured in parametric filters. Importantly static spectral plot updates in real-time reflecting any changes. This representation of band-pass filtering represents an improved visual display and importantly allows a form of direct manipulation with the visual interface.

- Hi / lo shelf: this control operates in similar way to the band-pass control and is also hidden by default. The control runs vertically along the bottom of each interface. Movement of the left triangle attenuates all frequencies below the marker and movement of the right triangle attenuates all frequencies above the marker. This control was developed to help share the frequency spectrum between channels. This approach has similar user interface benefits to those described for the band-pass filters,

- Bypass: this control temporarily removes the EQ applied to the audio, returning the spectral plot to its original state. Edited controls remain in place to provide the user with visual feedback of the changes made. This control is commonly employed in traditional EQ sections and helps the user assess the merits of EQ related decisions.

- Reset: The reset button removes all EQ changes with all controls moving to their default position.

- Volume: the volume for each audio channel is controlled by a simple vertical fader. Gain and attenuation are controlled by moving the fader upwards or downwards. Visual feedback is provided to the user by changing the height of the spectral plot. This control was developed to help set the relative level of different channels.

8.3 Implementation

The algorithm implemented not only processes the audio signal with regards to EQ but also changes the static spectral plot displayed in real time to provide visual feedback for a user’s action. There are three key sections to this algorithm as shown in Figure 71.

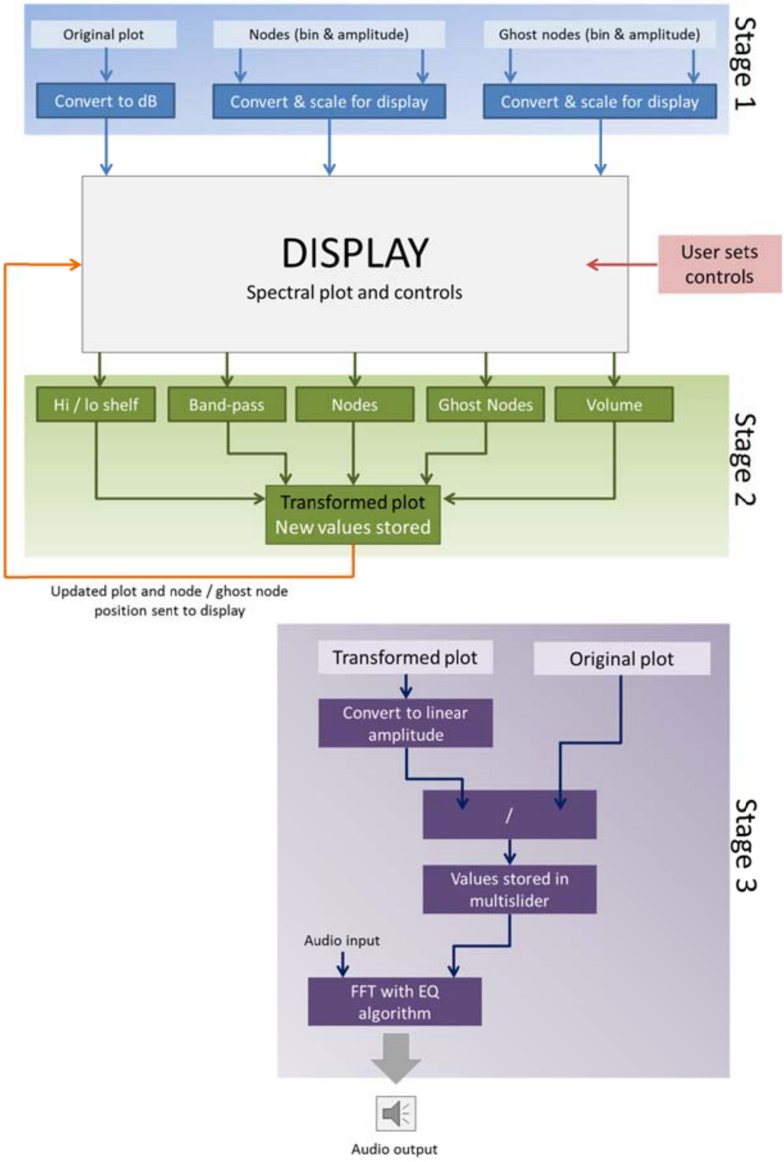

Figure 71: Flow diagram of the algorithm used in interfaces B, C and D

The first stage is concerned with the retrieval of analysis information as detailed in Chapter Seven. The bin maximum matrix for the channel considered is firstly converted from linear amplitude to logarithmic decibel values, and output for display as a spectral plot. The initial display position of the nodes is also calculated in this stage by applying a scaling formula proportional to the size of the display (in pixels) to align the individual node positions with the spectral plot.

The second part of the algorithm uses the output from the interface controls to modify the spectral plot and associated node positions in real time and is based on the scaled spectral values calculated in stage one. On moving a node, the vertical position is converted to a decibel value and stored in the transformed plot matrix. The plot is then output and displayed to reflect the changes made in real time.

The band-pass and shelving controls use arrays with a length equal to the number of bins sampled in the FFT analysis to store their corresponding multiplication factors. These multiplication values are set by moving the relevant control then output as lists from the array and stored in matrices. By multiplying the values stored in the transformed plot matrix with these matrix values the changes made by the controls are shown in real time on the plot.

The volume control operates in a simpler way, multiplying all spectral values by the volume level set. This is then reflected in real time in the channel's spectral plot. Any changes in node position made by these controls is also calculated and output.

The third stage provides the EQ algorithm and is similar to the EQ algorithm used in interface 7 in the first test. The transformed plot matrix is divided by the original plot matrix calculated in stage one. These values are then output to an array which is stored in a buffer. An FFT is performed on the audio signal of the track considered and the amplitude value of each frequency bin is multiplied by the corresponding value stored in the buffer. The signal is then converted back into the time domain by performing an inverse FFT and the modified sound output.

8.4 Test

Four mixing scenarios were created from unprocessed audio. The scenarios considered were specifically designed to target the EQ techniques commonly employed in multitrack mixing, as identified in Chapter Six.

1. The first scenario was used as practice material and featured two channels for an acoustic guitar. The first channel was captured by placing an AKG Perception 100 microphone close to the body of the guitar and the second channel, by placing an AKG C1000 microphone near to the neck. This exploratory task was considered suitable as practice material as there were many deficiencies in the recording, prompting the user to become familiar with the interfaces' EQ controls and method of feedback.
2. The second scenario featured one channel of electric rhythm guitar and one channel of electric lead guitar recorded using a DI box. This scenario was selected because of the overlapping frequency spectrums.
3. The third scenario features one bass guitar channel and one kick drum channel. This scenario was selected because it provided a good example of where the mirrored EQ technique could be employed.
4. The fourth scenario featured the recording of a female singer using a Shure SM58 and a closely miked acoustic guitar. Naturally occurring resonant frequencies and some ground

noise featured in both channels and provided a simple yet realistic scenario for mixing two instruments.

Eight test subjects were selected based on their level of experience in audio production and engineering, with an “expert frequent” level of user selected. The interfaces were configured for each scenario and exported from Max/MSP as standalone applications and loaded onto a suite of Apple Mac computers. The test subjects were provided with a standard Apple mouse and pair of headphones to enable multiple subjects to undertake the usability test simultaneously. All interaction during the test was captured as a log file using Moens’ (2011) dJogger, which also stored the total time for task completion. The EQ matrix data was also captured during the test, as well as a video file of each test using screen capture software.

Prior to the test, subjects completed a pre-test questionnaire. All users were provided with the practice scenario, the other three scenarios were divided between the group, with three subjects attempting scenarios 2, three attempting scenario 4 and two subjects attempting scenario 3.

During the practice period, each interface was explained and any concerns or queries addressed. Any important initial comments were recorded for evaluation. Once confident, test subjects were asked to use the interfaces in a specified randomised order to achieve a mix that they were happy with. Once all interfaces had been used for the task considered, test subjects were asked to rank the interfaces in order of preference and select from a list of keywords the five most pertinent adjectives in order of importance after Travis (2008). These results were collected using a Google form, a comments box was also provided for each interface to capture opinions.

A second round of testing was conducted two weeks later with the same group of test subjects. The rationale behind the decision to conduct a usability test in two stages was to gauge whether there were any changes in the test subjects opinion once they had time to reflect on their experiences. It was anticipated that a short gap between tests would support the subjects in feeling more confident using an EQ interface that was fundamentally different to traditional EQ designs and therefore provide a better assessment of usability.

The test followed the same structure established in week one with each subject attempting the two scenarios they did not try in week one. At the end of the test the subjects were given a short break before being interviewed by the designer. The selected keywords formed the basis of this discussion and the subjects were also asked to reflect on the differences in experience from week one to week two. It was intended that all subjects would be interviewed together in a focus group, however in reality the designer had to conduct the interviews in smaller groups to fit in with the test subjects prior commitments. Two of the subjects were unable to attend this second test. Due to time constraints they only attempted one more scenario three weeks after the first round of testing.

Carrascal & Jorda (2011) contemplated the best way to evaluate a subjective mixing task. They pondered, "though mixing has an inherent technical component, its aesthetic aspect is difficult to measure; there's not a precise "good" or "bad" way of doing it" (Carrascal & Jorda, 2011 p. 102). In the context of this usability test these aesthetics arguably relate to effectiveness. This is because the test subjects were instructed to mix the two recorded instruments in each scenario until they were satisfied with the results, implying the mix should be deemed 100% effective upon task completion. Therefore NTCT was selected as an appropriate measure of effectiveness. NTCT was also used to measure the efficiency of each interface in the corrective EQ experiment and was selected as an appropriate measure to use again. Satisfaction was considered in terms of preference ranking, keyword selection, post-test interview and any comments and observations recorded during the test.

8.5 Analysis of results

8.5.1 Effectiveness and efficiency

	1: Rhythm & Lead Guits				2: Vocals & Guitar				3: Kick & Bass				
	A	**B**	**C**	**D**	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**	
Subject 1	1.00	0.86	0.82	0.91					0.48	0.45	0.51	1.00	
Subject 2	0.83	0.85	1.00	0.69	1.00	0.43	0.69	0.48	1.00	0.60	0.97	0.67	
Subject 3	0.86	0.32	0.17	1.00	0.93	0.83	0.92	1.00	0.40	0.27	1.00	0.37	
Subject 4	0.51	0.83	0.68	1.00	1.00	0.04	0.28	0.26	0.22	0.25	1.00	0.49	NTCT
Subject 5	0.85	0.41	1.00	0.71	1.00	0.37	0.42	0.44					
Subject 6	0.58	0.64	1.00	0.40	1.00	0.30	0.77	0.11	1.00	0.42	0.70	0.67	
Subject 7	0.81	0.43	0.53	1.00	0.81	0.59	0.75	1.00	0.70	1.00	0.64	0.48	
Subject 8	1.00	0.43	0.78	0.43	0.76	0.65	0.65	1.00	0.82	1.00	0.84	0.59	
Average NTCT	0.81	0.60	0.75	0.77	0.93	0.46	0.64	0.61	0.66	0.57	0.81	0.61	
Std. Dev.	0.18	0.23	0.29	0.25	0.10	0.26	0.22	0.38	0.30	0.32	0.20	0.20	

Figure 72: NTCT for all test subjects and scenarios

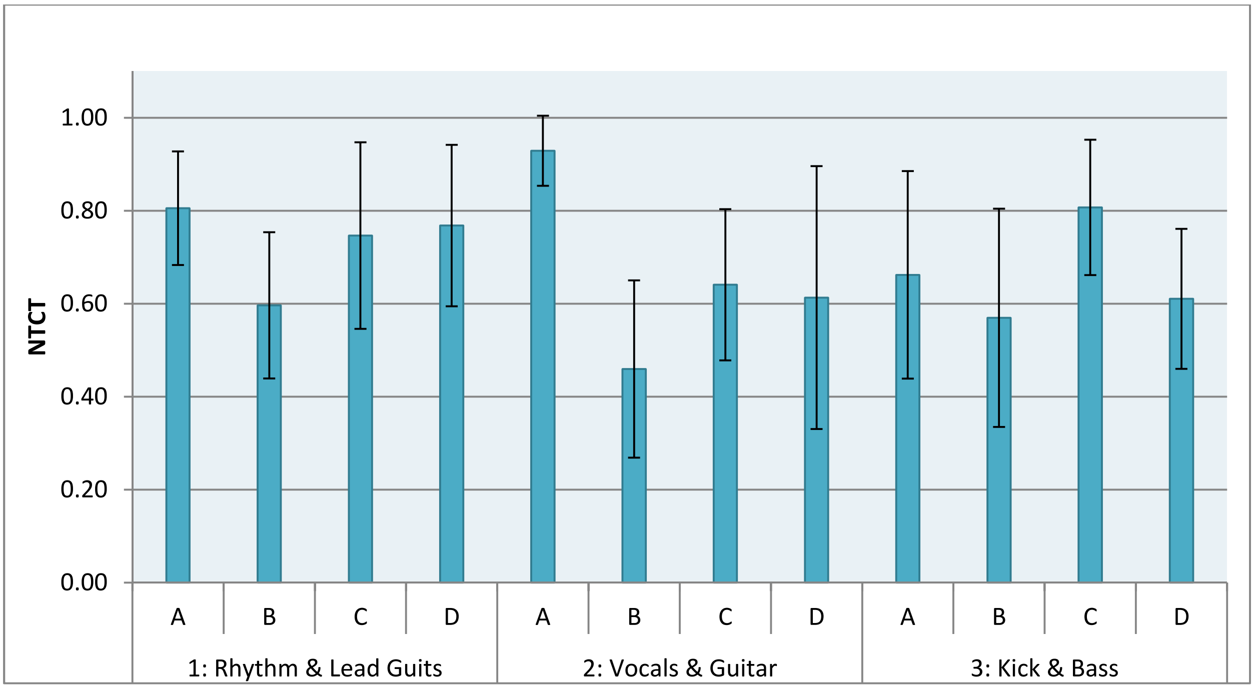

Figure 73: Average normalised task completion time (NTCT) with 95% confidence intervals for all interfaces and scenarios considered

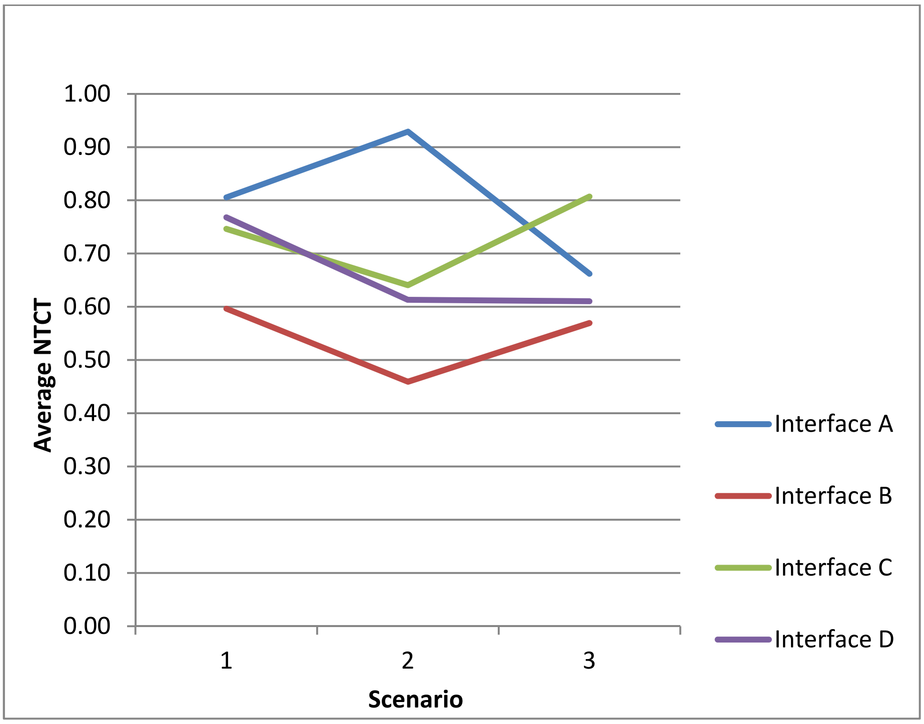

Figure 74: Plot of average normalised task completion time (NTCT) for each interface and scenario considered

Figure 72 presents the NTCT for all subjects and scenarios tested and the average and standard deviation of the NTCT for each interface. Figure 73 presents the average NTCT for all interfaces and scenarios with Figure 74 presenting this information as a plot for each interface tested. A visual inspection of the results suggests that the subjects were able to perform the tasks fastest with interface B as it has the lowest average NTCT in all scenarios tested.

In two of the three scenarios interface A had the highest average NTCT. Inspection of the histograms in Figure 73 suggests a statistically significant improvement when using interfaces B and C over interface A in scenario 2. With regards to interface B, this may be due to subject 4 having a very low NTCT (0.04) in the scenario 2 test. In scenario 2 the difference in average NTCT between interface A and the three novel interfaces tested is greatest.

A within-within subjects ANOVA was conducted to determine whether there were statistically significant differences in NTCT for the interfaces and scenarios tested. There were no outliers in the data, as assessed by inspection of a boxplot. NTCT was normally distributed for each scenario and interface tested, as assessed by Shapiro-Wilk's test ($p > .05$) with the exception of scenario 1 interface D where the data violates normality as $p = .048$. This could be due to the small number of test subjects used in this investigation. As normality was only marginally violated in one of the interfaces tested, it was worthy of note but did not prevent analysis with within-within subjects ANOVA. Mauchly's test of sphericity indicated that the assumption of sphericity had not been violated for scenario (($X2(2) = 1.64$, $p = .092$) and interface (($X2(2) = 1.64$, $p = .440$). Furthermore, there was no statistically significant interaction between scenario and interface on NTCT, $F(6,30) = 0.990$, $p = 0.45$, partial $n2 = .165$. The main effect of interface showed a statistically significant difference in NTCT, $F(3, 15) = 5.06$, $p =.013$, partial $n2 = .503$. Post hoc analysis with a Bonferroni adjustment revealed that NTCT was statistically significantly reduced using interface B as opposed to interface A (0.24 (95% Cl, 0.06 to 0.43, $p = 0.02$). No further statistical significant differences were apparent with other interfaces tested, as displayed in the pairwise analysis in Figure 75. The main effect of scenario showed no significant difference in NTCT, $F(2, 10) = 0.06$, $p = .943$, partial $n2 = .503$.

(I) interface	(J) interface	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
A	B	.242*	.044	.016	.056	.428
	C	.049	.030	.947	-.076	.175
	D	.144	.090	1.000	-.235	.523
B	A	-.242*	.044	.016	-.428	-.056
	C	-.193	.061	.152	-.451	.065
	D	-.098	.062	1.000	-.361	.165
C	A	-.049	.030	.947	-.175	.076
	B	.193	.061	.152	-.065	.451
	D	.095	.093	1.000	-.298	.488
D	A	-.144	.090	1.000	-.523	.235
	B	.098	.062	1.000	-.165	.361
	C	-.095	.093	1.000	-.488	.298

Based on estimated marginal means
*. The mean difference is significant at the .05 level.
b. Adjustment for multiple comparisons: Bonferroni.

Figure 75: Pairwise comparison between interfaces from a post hoc analysis with statistically significant relationships highlighted.

8.5.2 Satisfaction

8.5.2.1 Preference Ranking

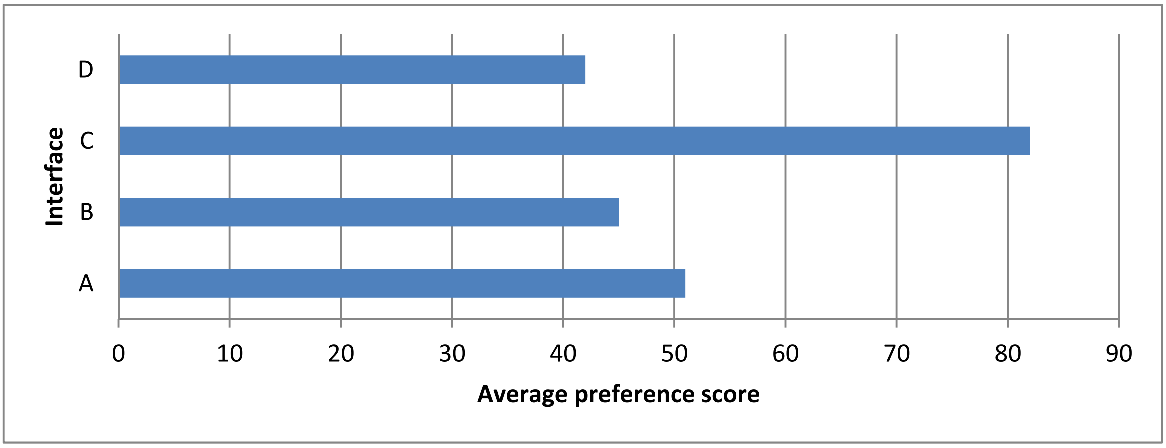

Figure 76: Average preference scores for all subjects and scenarios

The test subjects put each interface in rank order of preference. These rankings were then converted into an average score by averaging each interface's rank position across all test subjects, as shown in Figure 76. This indicates a user preference for interfaces D and B over the traditional interface A. Interface D ranks as most preferred with interface B second. Interface C ranks poorly in comparison with the other interfaces tested. It is interesting to note that during the week 1 tests, all but one subject selected the final interface they were presented with during the test as their preferred interface for the scenario considered. This was possibly caused by the test subjects taking more time than the practice period allowed to become familiar with the interfaces considered and structure of the of the usability test.

8.5.2.2 Selected keywords

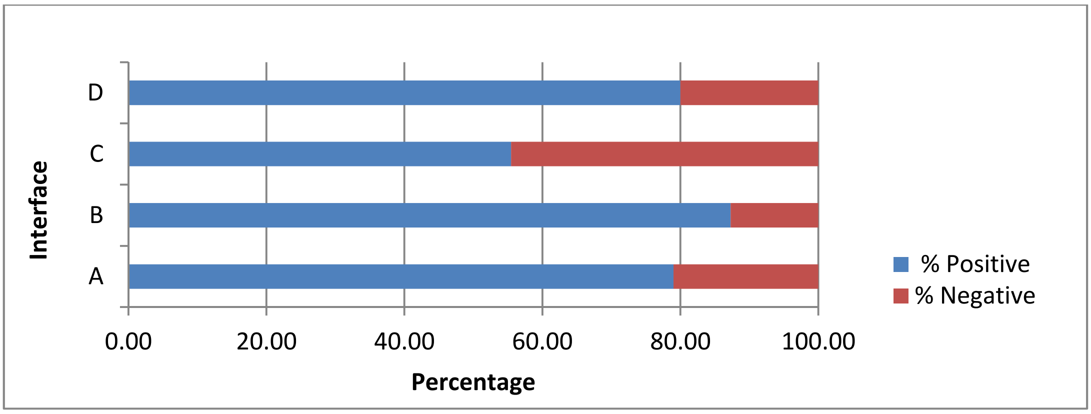

Figure 77: Percentage of positive and negative keywords selected for each interface

The percentage of positive and negative keywords selected to describe the interfaces by the test subjects (see Figure 77) suggests a preference for interface B as only 13% of the keywords selected for this interface were negative. Interface C has the highest number of negative keywords selected at 45% with interfaces A and D receiving around 20% negative keywords.

Figure 78a: Keywords selected for interface A

Figure 78b: Keywords selected for interface B

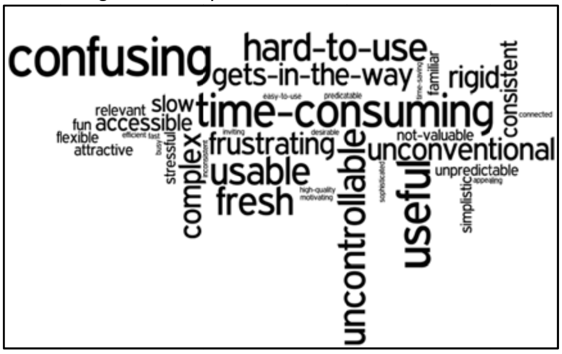

Figure 78c: Keywords selected for interface C

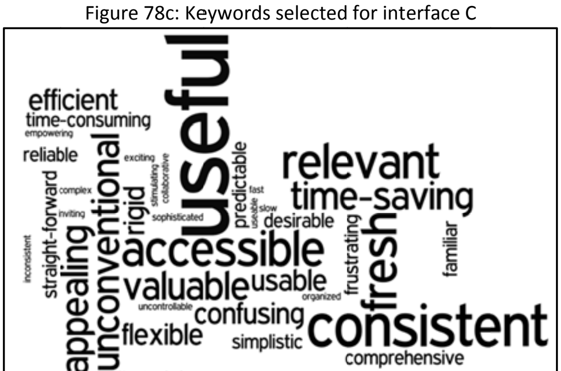

Figure 78d: Keywords selected for interface D

Figure 78: Wordclouds generated for each interface based on the keywords selected by the test subjects

Figure 78 displays all the keywords selected to describe each interface as a word-cloud with the size of the keyword proportional to the number of times it was selected. As the test subjects were asked to preference rank their keyword selection further analysis was possible.

Interface A	Interface B	Interface C	Interface D
familiar	accessible	confusing	useful
easy to use	easy to use	useful	appealing
time consuming	useful	hard to use	relevant
consistent	simplistic	fresh	consistent
accessible	fast	uncontrollable	time-saving

Figure 79: Top five 5 keywords for each interface based on preference

The number of keyword occurrences was summed based on user ranking. The top five scoring keywords for each interface are displayed in the Figure 79 in order of ranking. The keywords selected for interface A support the test subjects "familiarity" with DAW style interfaces, interestingly, the negative descriptor "time-consuming" is the third highest scoring keyword. The top keywords for interfaces B and D feature only positive words. The top keywords for interface C indicate a negative user perception, however, as with the other novel interfaces tested "useful" also scores highly.

8.5.2.3 Post-test interviews

All subjects said that they were more confident with using the novel interfaces in the second week of testing with 7 out of 8 subjects favourable to the tools and features employed with 4 subjects commenting that they began to mix with their eyes once they became familiar with the controls. Opinions were divided over whether interface B or D was the best interface overall with subjects commenting that although the overlayed spectral plots in interface D did prove useful in some instances, it felt cluttered in comparison with interface B which appeared more organised and logical.

With regards to the EQ tools provided in the novel interfaces the test subjects found the band-pass tools most useful. Several of the test subjects would have liked more than two band-pass controls. One user commented that the naming convention may be misleading for this tool. All users commented that they would prefer to manually turn on and off the band-pass and shelving controls, finding the automated display features hard to use.

All users saw the significance of the nodes in identifying and treating peak frequencies, but would like to be able to fine tune the frequency range they control. All users commented that although 5 nodes were sufficient in scenario 3 more nodes were required to treat the vocal track in scenario 2. Users were less receptive to the ghost nodes. Half of the subjects tested liked this feature although they initially struggled to understand the concept conveyed.

All test subjects commented that interface C was hardest to use. The linear display of amplitude was a particular barrier as the nodes and spectral plots were placed very close together at lower amplitudes, making it physically harder to remedy EQ issues. 6 test subjects found the layered display of spectral plots misleading, struggling to make the link between control and visualisation.

One subject found that they were able to approach the task of EQing differently given the tools and visualisations provided.

8.5.2.4 Observations

All users questioned the linear scale used in interface C, seeking clarification of what was being represented. This departure from the norm clearly affected the test subjects satisfaction with 2 subjects clearly stressed when using this interface.

When presented with the interfaces in week one all users were observed initially pushing the capabilities interfaces tested before attempting the task considered.

8.6 Conclusions

- Three novel EQ user interfaces specifically designed to mix two channels of recorded instruments have been introduced in this investigation and undergone preliminary HCI usability testing.
- Interface B has the lowest average NTCT for the three scenarios considered. Furthermore, Interface B offers a statistically significant reduction in NTCT when compared with interface A indicating Interface B is a more usable interface in comparison to a typical existing DAW EQ user interface.
- Although interface D was ranked most preferred, Interface B received the most praise from the test subjects in the post-test interview. The overlayed plots in interface D were considered by many subjects as only useful periodically and seen as a visual barrier in most scenarios considered.
- Despite perceiving the display style in interface C as useful, the test subjects clearly considered this interface the least satisfying to use and often questioned the controls, despite being fundamentally identical to those in interfaces B and D. This could be due to interface C offering a very different style of display in part due to the linear amplitude scale used.
- Using the keywords to structure the post-test interview stage helped elicit a greater understanding of the test subjects opinions regarding the interfaces and controls provided. The test subjects commented that they were largely receptive to a static display and identification of peaks. Most of the users also commented that they would like to have more control over the peaks selected and would welcome a feature that enabled them to fine-tune their selection. The band-pass control was the most favoured control overall with the

addition of more bands suggested. All the test subjects found the auto-hide features annoying.

9 Summary, Conclusions and Further Work

9.1 Summary

The overall aim of this research project was develop a new EQ user interface for mixing multi-track recordings of musical instruments. The specific research objectives included:

- an investigation of related current and novel interfaces
- establishing an understanding of HCI fundamentals with regard to interface design
- the development and evaluation of a simple EQ interface for corrective EQ tasks
- the development and evaluation of a more sophisticated EQ interface for mixing multi-track recordings of musical instruments

The investigation into current mixing console interfaces revealed that all traditional consoles divide the interface into discrete sections with input channels displayed as a series of vertical strips. The analogue console interface is favoured because it enables the user to quickly see the state of each control. Although this feature is impaired in traditional digital consoles, the inclusion of graphical control visualisations counteract this shortfall. This is particularly pertinent with regards to EQ controls, with the EQ curve display offering significant improvements over the row-of-knobs embodiment.

With regards to novel interfaces the two-dimensional stage metaphor is widely regarded as a plausible alternative to the traditional linear strip display. It enables multiple settings to be visualised simultaneously and the relationships between channels considered.

The evolution of the EQ section indicates that the audio industry is receptive to further interface development. This is alluded to in the widespread adoption of real time spectral plot displays behind the EQ curve and the development of subjective interfaces to assist the novice user.

A review of audio analysis displays offered some design considerations, as they present the results of analysis as a static visualisation. The novel interfaces considered utilised a single visual entity to display a variety of information to the user in a simplified manner.

A series of user centred guidelines can be adhered to help develop a new interface. Bertin (1983) *et al* reveal that relevant data can be used as the basis of a visual interface with a maximum of four visual variables offering the optimum number of variables to the user. The taxonomies of data visualisation provide a wide range of examples that can be considered to help develop a meaningful visual display. Developing an understanding of the cognitive processes undertaken by the user can also assist the designer in developing more intuitive interfaces.

Evaluation of interface designs should be carried out throughout the design process by consulting an expert user. Paper-prototypes can be used initially to check the suitability of the design in a variety of scenarios and rapidly refine these design ideas. More formal methods of evaluation were identified by studying the methods used to evaluate DMIs. An interface's usability can be evaluated by considering suitable measures of efficiency, effectiveness and satisfaction.

An analysis of the EQ task was undertaken through the study of relevant literature. It revealed that EQ tasks can be broadly categorised as single channel corrective tasks or multi-channel creative tasks. Several discrete methods were identified to help mix multi-track recordings of musical instruments.

Four novel interfaces were developed for the specific task of corrective EQ. Central to each design was the inclusion of an algorithm that automatically detected the five peak frequencies. These were presented in a variety of ways. A preliminary investigation revealed that the test subjects displayed fewer periods of thinking followed by activity when using interfaces 5 and 7. The subjects tested had the lowest average NTCT when using Interface 7. Furthermore, Interface 7 offered a statistically significant reduction in NTCT when compared with the benchmark Interface 2. The test subjects were largely receptive to the assistive controls featured.

Interface 7 was selected for further development because it most closely adhered to the HCI related guidance and was deemed the best interface in the preliminary usability test. A second phase of development was undertaken to create an interface and range of controls for mixing multi-track recordings of musical instruments. Three novel interfaces were developed based on the three most common forms of time-series data visualisations identified in the study of HCI literature.

Following a user centred design process; a second usability test was conducted. Three scenarios were constructed based on the findings of the EQ task analysis with the test conducted over a two week period. The test revealed that the subjects were most efficient and effective when using Interface B which featured multiple individual plots. Interface B also offered a statistically significant reduction in NTCT when compared with the benchmark Interface A, indicating it was a more efficient and effective interface. A study of user satisfaction clearly revealed that the test subjects preferred interface C the least and were most favourable to Interfaces B and D. With regards to the novel controls offered, the test subjects saw the importance of the nodes and ghost nodes but found the band-pass control most useful .

This research project has shed light on the range of methods required to develop a new mixing console interface in considering the redesign of a single control. Adopting the approach of conducting usability tests to inform and steer the design process has helped structure this project significantly. It has provided a key insight into a user's perception of audio fundamentals and has helped develop new visual metaphors from which further redesigns can embark.

9.2 Conclusions

This work has shown that:

- HCI research is useful in guiding the development of new interfaces for the corrective EQ and multi-track EQ of recorded musical instruments. Specific design techniques employed in this research included:
    - Determination of user skill level – "expert frequent" user being chosen
    - Task analysis of the EQ process.
    - Creating paper prototypes for exploration and task analysis before developing prototypes in software.
    - Use of analysis data as the core of the user interfaces graphical visualisations – specifically a static spectral plot
    - Design interface to allow user to directly manipulate the visualisation presented by the user interface – for example interacting with the peak nodes on the spectral plot.
- HCI research can be used to guide the evaluation of new interfaces for the corrective EQ and multi-track EQ of recorded musical instruments. Specific evaluation techniques employed in this research include:
    - Informal evaluation by consulting an expert user throughout the design process.
    - Simulating a range of scenarios to refine and develop the paper prototypes into workable designs.
    - Evaluating candidate interfaces in terms of their usability by identifying suitable measures of efficiency, effectiveness and satisfaction and suitable test scenarios.
    - Efficiency was measured in terms of task completion time.
    - Effectiveness was measured in terms of accuracy in simple tasks objective tasks. This was more difficult to measure in the multi-track EQ tests where we adopted the approach of allowing the test subject to complete the task to their own satisfaction and used task completion time for this measure.
    - Satisfaction was measured in terms of user preference ranking, observing user interaction and comments and selection of keywords to describe candidate user interfaces which then formed the basis of a semi-structured post-test interviews.

- By recording a log of user interaction during the corrective EQ interface tests it was possible to identify periods of thinking followed by activity and use this as an additional measure of efficiency.
  - It was observed in these evaluations that using many measures can provide contradictory results and this makes conclusive evaluation of a user interface difficult. Furthermore, when test subjects are presented with user interfaces that differ radically from the established norms they can exhibit a resistance to change. This results in radical redesigns potentially being quite likely to suffer poor preference ratings in comparison to more traditional looking designs.
- A corrective EQ was developed that was favoured in evaluation over a typical existing DAW EQ.
- The novel features of the favoured corrective EQ user interface were:
  - Spectral data visualised as a static plot.
  - Peak frequencies calculated and presented to the user for consideration and attenuation.
  - Use of nodes upon the spectral plot to enable the user to interact directly with the visual information
  - Visual feedback provided by the spectral plot changing in real-time to reflect the settings made.
- A multi-track EQ was developed that was favoured in evaluation over using an existing DAW EQ for a multi-track EQ task
- The novel features of the multi-track EQ user interface were:
  - Tracks were simultaneously displayed as multiple spectral plots.
  - Each individual track interface was the same as for the corrective EQ user interface
  - Additional novel methods for the visualisation and direct manipulation of the interface to perform mirrored EQ were provided in the form of ghost nodes representing peaks from the other track on a track's spectral plot. A visual method of interacting with the spectral display to perform band pass and low and high shelf filtering were also included.
- More broadly this thesis hints at how the adoption of HCI theories and consideration of modern interface technology could benefit and transform the design of future recording studio equipment user interfaces.

9.3 Further Work

Further work includes the following:

A multi-track EQ user interface will require more channels to be displayed to be fully comparable with the functionality of multiple EQ sections in traditional interfaces. The preferred multi-track EQ user interface will start to take up significant screen space if simply extended to more tracks and accommodating 24 tracks on a single screen will require further consideration. More designs will need to be considered and prototyped to achieve this objective.

Extension of the multi-track EQ user interface to include other mixing tasks, such as pan, will also require further consideration. A version of the two-dimensional and three-dimensional stage metaphor would appear to offer potential as suggested by Gelineck *et al* (2013). Furthermore, it is anticipated that the design process will identify further metaphors that have not yet been considered in multi-track interfaces which will require development and evaluation.

Developing the user interface to allow stereo input sources in addition to mono sources and to output a stereo mix will again require development and evaluation.

A wider range of scenarios and types of recorded instruments should be considered in further usability tests. Furthermore future tests should be conducted with a greater number of test subjects. Ideally a different group of test subjects should be selected for each scenario and subjects should be tested over extended time periods.

The proposed novel EQ user interfaces could be technically improved by employing a more sophisticated EQ algorithm and further consideration of implementation details such as the best choice of FFT size to provide the most usable resolution for the spectral plots and whether the use of FFT bin maximums is better than a gated FFT average for the spectral plots.

10 References

Ableton. (2013). Ableton Reference Manual Version 9 [Brochure]. Retrieved from https://www.ableton.com/en/manual/welcome-to-live

Allen & Heath. (2008). ZED R16 User Guide [Brochure]. Retrieved from http://www.allen-heath.com/media/AP7236_iss_2.pdf

Apple. (2013). Logic Pro X User Guide [Brochure]. Retrieved from http://support.apple.com/manuals/#professionalsoftware

AMS Neve. (2013). 88RS Product Brochure [Brochure]. Retrieved from http://www.ams-neve.com/products/music/88-rs

Behrens, C. (2008). *Design Pattern categories*. Retrieved from http://www.niceone.org/infodesignpatterns/index.php5#/patterns.php5

Behringer. (2012). User Manual EURORACK UB1202 /UB1002 [Brochure]. Retrieved from http://www.behringer.com/assets/ub1202_ub1002_ub802_ub502_M_en.pdf

Benedek, J., & Miner, T. (2002). *Measuring desirability: New methods for evaluating desirability in a usability lab setting*. Proceedings of Usability Professionals Association, 2003, 8-12.

Bertin, J. (1983). *Semiology of graphics: Diagrams, maps networks*. Madison WA: University of Wisconsin Press.

Brixen, E. (2010). *Audio Metering: Measurements, Standards and Practice*. Burlington MA: Focal Press

Card, S. K., English, W. K., & Burr, B. J. (1978). *Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT*. Ergonomics, 21(8), 601-613.

Carrascal, J. P., and Jorda, S. (2011). *Multitouch interface for audio mixing*. In Proceedings of New Interfaces for Musical Expression (NIME 2011) (pp. 100–103). Oslo, Norway. Retrieved from http://www.nime2011.org/proceedings/papers/B19-Carrascal.pdf

Charles, J. F. (2008). A tutorial on spectral sound processing using Max/MSP and Jitter. *Computer Music Journal*, 32(3), 87-102.

Chuchacz, K., O'Modhrain, S., & Woods, R. (2007, June). *Physical models and musical controllers: designing a novel electronic percussion instrument*. In Proceedings of the 7th international conference on New interfaces for musical expression (pp. 37-40). ACM.

Crystal, A., & Ellington, B. (2004, August). *Task analysis and human-computer interaction: approaches, techniques, and levels of analysis*. In AMCIS (p. 391). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.76.3295&rep=rep1&type=pdf

Cycling 74. (n.d.). *Max.* Retrieved from http://cycling74.com/products/max/

Diamante, V. (2007). *Awol: Control surfaces and visualization for surround creation*. Technical report, University of Southern California, Interactive Media Division. Los Angeles CA. Retrieved from http://www.surroundshape.com/diamante_awol_thesispaper_18.pdf

Eastty, P. C., Thorpe, P. D., & Sleight, C. (2002). U.S. Patent No. 6,359,632. Washington, DC: U.S. Patent and Trademark Office.

Erickson, T., & McDonald, D. W. (2008). *HCI Remixed: Reflections on Works That Have Influenced the HCI Community*. Cambridge, MA: The MIT Press.

Gelineck, S., & Serafin, S. (2009). *A quantitative evaluation of the differences between knobs and sliders*.

Gelineck, S., Büchert, M., Andersen, J. (2013, April 27-May 2). *Towards a more flexible and creative music mixing interface*. In proceedings of CHI 2013 Extended Abstracts (ACM 978-1-4503-1952-2/13/04) (pp. 733-738). Paris, France. Retrieved from http://vbn.aau.dk/files/74873966/gelineck_CHI_2013.pdf

Gibson, D. (1997). *The Art Of Mixing: A Visual Guide To Recording, Engineering, And Production. Boston*, MA: ArtistPro Press.

Gibson, D. (2008*). The Art Of Mixing: A Visual Guide To Recording, Engineering, And Production (Second Edition)*. Boston, MA: ArtistPro Press.

Gonzalez, E. P., & Reiss, J. (2007, September). *Automatic mixing: live downmixing stereo panner*. In Proceedings of the 10th int. conference on digital audio effects (DAFx-07), Bordeaux, France. Retrieved from http://dafx.labri.fr/main/papers/p063.pdf

Gonzalez, E. P., & Reiss, J. D. (2008, September*). Improved control for selective minimization of masking using interchannel dependancy effects*. In Proceedings of the 11th int. conference on digital audio effects (DAFx-08), Espoo, Finland. Retrieved from https://www.acoustics.hut.fi/dafx08/papers/dafx08_15.pdf

Green, M. (2006*). Toward a perceptual science of multidimensional data visualization: Bertin and beyond*. ERGO/GERO Human Factors Science, 1-30. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.378.3579&rep=rep1&type=pdf

Hansen, W. J. (1971, November). *User engineering principles for interactive systems*. In Proceedings of the November 16-18, 1971, fall joint computer conference (pp. 523-532). ACM.

Heer, J., Bostock, M., & Ogievetsky, V. (2010). *A tour through the visualization zoo*. Commun. ACM, 53(6), 59-67. Retrieved from http://hci.stanford.edu/jheer/files/zoo/

Holtzblatt, K., Wendell, J. B., & Wood, S. (2005). *Rapid contextual design: a how-to guide to key techniques for user-centered design*. San Francisco CA: Elsevier.

Huber, D. M., Runstein, R. E. (2014) *Modern Recording Techniques (8th Edition)*. Burlington, MA: Focal Press

Hughes, B., & Wakefield, J. (2012, April). *User Interface Evaluation for Discrete Sound Placement in Film and TV Post Production*. In Audio Engineering Society Convention 132. Audio Engineering Society.

Hunt, A., & Kirk, R. (2000). Mapping strategies for musical performance. *Trends in Gestural Control of Music*, 21.

Izhaki, R. (2008). *Mixing Audio Concepts, Practices and Tools*. Waltham, MA: Focal Press

Izotope. (2013). Mastering with Ozone [Brochure]. Retrieved from http://downloads.izotope.com/guides/iZotopeMasteringGuide_MasteringWithOzone.pdf

Jacob, R. J., Deligiannidis, L., & Morrison, S. (1999). A software model and specification language for non-WIMP user interfaces. *ACM Transactions on Computer-Human Interaction* (TOCHI), 6(1), 1-46. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.17.8771&rep=rep1&type=pdf

Kiefer, C., Collins, N., & Fitzpatrick, G. (2008, June). *HCI methodology for evaluating musical controllers: A case study*. In Proceedings of NIME (Vol. 8).

Lee, E. (2012, May). *A Taxonomy of Data Visualization*. Retrieved from http://visualizing.org/stories/taxonomy-data-visualization

Leyshon, A. (2009). The Software Slump?: digital music, the democratisation of technology, and the decline of the recording studio sector within the musical economy. *Environment and Planning* A 41(6), 1309 – 1331.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7(1), 57-78.

Liebman, N., Nagara, M., Spiewla, J., & Zolkosky, E. (2010). *Cuebert: A new mixing board concept for musical theatre*. In Proceedings of New Interfaces for Musical Expression.

Lima, M. (2011). *Syntax of a New Language*. Retrieved from http://www.visualcomplexity.com/vc/book/posters.cfm

Line 6. (n. d.) *StageScape M20d*. Retrieved from http://line6.com/stagescape-m20d/features

Lopes, P., Ferreira, A., & Pereira, J. M. (2011). *Battle of the DJs: an HCI perspective of Traditional, Virtual, Hybrid and Multitouch DJing*. In Proceedings of the International Conference on New Interfaces for Musical Expression} (pp. 367-372).

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability interface*, 8(2), 3-6.

Ma, Z., Reiss, J. D., & Black, D. A. (2013, May). Implementation of an intelligent equalization tool using Yule-Walker for music mixing and mastering. In *Audio Engineering Society Convention 134*. Audio Engineering Society.

Mecklenburg, S. & Loviscach, J. (2006, April 22-27). *subjEQt: Controlling an equaliser through subjective terms*. In proceedings of Extended Abstracts Proceedings of the 2006 Conference on

Human Factors in Computing Systems (CHI 2006). Montréal, Canada. Retrieved from http://www.j3l7h.de/publications/subjeqt_web.pdf

Moens, B. (2011). *DJogger*. Retrieved from http://www.ipem.ugent.be/bartmoens/index.php

Mixerman. (2010). *Zen and the Art of Mixing*. Milwaukee, WI: Hal Leonard Books

MixGenius. (n.d.) Retrieved from www.mixgeni.us

Myatt, G. J., & Johnson, W. P. (2009). *Making sense of data II: A practical guide to data visualization, advanced data mining methods, and applications*. John Wiley & Sons.

Mynett, M., Wakefield, J. P., & Till, R. (2010). *Intelligent equalisation principles and techniques for minimising masking when mixing the extreme modern metal genre*.

Mynett, M. (2012). Achieving intelligibility whilst maintaining heaviness when producing contemporary metal music*. Journal on the Art of Record Production*, (6).

Nash, C. & Blackwell, A.F. (2011, July 31-August 5). *Tracking virtuosity and flow in computer music*. In proceedings of the International Computer Music Conference (ICMC 2011) (pp. 575-582). Huddersfield, UK. Retrieved from http://www.nashnet.co.uk/files/ICMC2011-Nash&Blackwell.pdf

Nielsen, J. (1994, April). *Enhancing the explanatory power of usability heuristics*. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 152-158). ACM. Retrieved from http://www.cs.helsinki.fi/u/salaakso/kl2-2002/lahteet/Nielsen94-Enhancing-Heuristics.pdf

Norman, D. A. (2002). *The design of everyday things*. NY: Basic books.

O'Modhrain, S. (2011). *A framework for the evaluation of digital musical instruments*. Computer Music Journal, 35(1), 28-42.

Perez-Gonzalez, E., & Reiss, J. (2009, October). Automatic equalization of multichannel audio using cross-adaptive methods. In *Audio Engineering Society Convention 127*. Audio Engineering Society.

Pohlmann, K. C. (2011). *Principles of digital audio (Vol. 6).* McGraw Hill

Presonus. (2013). Studio One 2.6 User Manual [Brochure]. Retrieved from

http://www.presonus.com/products/studio-one/download

Price, A. (2013, February) *StageScape M20d Review*. Retrieved from http://www.musictech.net/2013/02/stagescape-m20d-review/

Reiss, J. D. (2011). Design of audio parametric equalizer filters directly in the digital domain. *Audio, Speech, and Language Processing, IEEE Transactions on*, *19*(6), 1843-1848.

Reiss, J. D., & Perez Gonzalez, E. (2008, May). *An automatic maximum gain normalization technique with applications to audio mixing*. In Audio Engineering Society Convention 124. Audio Engineering Society. Retrieved from http://www-student.elec.qmul.ac.uk/people/josh/documents/PerezReiss-AES124-Anautomaticmaximumgainnormalization.pdf

Sabin, A., & Pardo, B. (2009, October). *A method for rapid personalization of audio equalization parameters*. In Proceedings of the 17th ACM international conference on Multimedia (pp. 769-772). Retrieved from http://music.cs.northwestern.edu/publications/sabin-pardo-acmmm09.pdf

Sabin, A., & Pardo, B. (2009, October). *2DEQ: an intuitive audio equalizer*. In Proceedings of the seventh ACM conference on Creativity and cognition. (pp. 435-436). Retrieved from http://music.eecs.northwestern.edu/publications/sabin-pardo-acmcc09.pdf

Samson. (2003). Samson S-Curve Dual 231 User Manual [Brochure]. Retrieved from http://www.samsontech.com/site_media/legacy_docs/S-curve%20231_ownman_v1s.pdf

Schmidt Horner, S. (2013). *Chasing Sound Technology, Culture & the Art of Studio Recording from Edison to the LP*. Baltimore, MD: The John Hopkins University Press

Schnell, N., Cifuentes, M. A. S., & Lambert, J. P. (2010). *First steps in relaxed real-time typo-morphological audio analysis/synthesis*. Retrieved from http://smcnetwork.org/files/proceedings/2010/57.pdf

Scott, J., Chapman, J., Moorhead, M., Schwabe, R. & Kim, Y. (2013). *Personalized Song Interaction Using a Multi-Touch Interface*. In Proceedings of New Interfaces for Musical Expression (NIME'13). Daejeon, Korea. Retrieved from http://music.ece.drexel.edu/files/Navigation/Publications/Scott2013.pdf

Senior, M. (2011). *Mixing secrets*. Burlington MA: Focal Press

Shneiderman, B. (1996, September). *The eyes have it: A task by data type taxonomy for information visualizations*. In Visual Languages, 1996. Proceedings., IEEE Symposium on (pp. 336-343). IEEE. Retrieved from http://drum.lib.umd.edu/bitstream/1903/5784/1/TR_96-66.pdf

Shneiderman, S. B., & Plaisant, C. (2005). *Designing the user interface 4th edition*. Pearson

Soundcraft. (2001). *Soundcraft the History*. Retrieved from http://www.soundcraft.com/news/soundcraft.aspx

Steinberg. (2013). Cubase 7 User Manual [Brochure]. Retrieved from http://www.steinberg.net/en/support/downloads/downloads_cubase_7.html

Stowell, D., Plumbley, M. D., & Bryan-Kinns, N. (2008, June). *Discourse analysis evaluation method for expressive musical interfaces*. In Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)(pp. 81-86).

Studer. (2012). *Studer Product History*. Retrieved from http://www.studer.ch/news/history.aspx

Supper, B. (2011). *Report on Intelligent Audio Editing Technologies Lecture by Josh Reiss*. Retrieved from http://www.aes-uk.org/forthcoming-meetings/intelligent-audio-editing-technologies/

Travis, D. (2008). *Measuring satisfaction*. Retrieved from http://www.userfocus.co.uk/articles/satisfaction.html

User Focus. (n.d.). ISO 9241 *Part 11: Guidance on usability*. Retrieved from http://www.userfocus.co.uk/resources/iso9241/part11.html

Van den Haak, M. J., & de Jong, M. D. (2003, September). *Exploring two methods of usability testing: concurrent versus retrospective think-aloud protocols*. In Professional Communication Conference, 2003. IPCC 2003. Proceedings. IEEE International (pp. 3-pp). IEEE.

Wanderley, M. M., & Orio, N. (2002). Evaluation of input devices for musical expression: Borrowing tools from HCI. *Computer Music Journal*, 26(3), 62-76.

White, P. (1997). *Music technology, a survivors guide*. London, UK: MPG Books Ltd.

White, P. (2012, April) *Line 6 Stagescape & Stagesource, Modular Live Sound System*. Retrieved from http://www.soundonsound.com/sos/apr12/articles/line-6-tech-interview.htm

White, P. (2012, November) *Line 6 StageScape M20D Digital Mixer With Multitrack Recording.* Retrieved from http://www.soundonsound.com/sos/nov12/articles/line-6-m20d.htm

Wiklund, M. E., Thurrott, C., & Dumas, J. S. (1992, October). *Does the fidelity of software prototypes affect the perception of usability?* In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 36, No. 4, pp. 399-403). SAGE Publications.

Yamaha. (2004). *Yamaha Releases V2 Software for 02R96 Digital Mixer*. Retrieved from http://usa.yamaha.com/news_events/live_sound/yamaha-releases-v2-software-for-02r96-digital-mixer/

Yamaha. (2010). O2R96 V2 Product Manual [Brochure]. Retrieved from http://www.yamahaproaudio.com/global/en/products/mixers/02r96vcm/downloads.jsp