Analogue Hearts, Digital Minds

Michael Uwins

Innovations in audio are invariably met with mixed responses and this was never more evident than during the early 1980s, when the compact disc was first heralded as the successor to vinyl[1]. Further developments have enabled music to be captured with increased sampling rates and resolutions, meaning that digital systems are now capable of reproducing at least in theory, audio of a quality that surpasses any of their predecessors. Despite this, the 'digital versus analogue' debate is one that shows no sign of abating, continuing today in journals, music magazines, hi-fi periodicals and other forums. Over the past thirty years I have purchased, borrowed, recorded and enjoyed many records on vinyl, reluctantly making the transition to digital in the 1990s. After first training as an audio engineer and now currently lecturing in the subject, I both have participated (and watched from the side-lines) in many such debates and whilst the theory and the science clearly suggest that the compact disc version of an album should be superior, I still find myself favouring the sound of my old (and rather worn) vinyl. This article describes my attempt to find out why my head is at odds with my heart, by discovering if and how perceptions of audio fidelity are influenced by factors other than what is coming out of the speakers.

1. Vinyl attraction

Figures published by the IFPI [1] at the beginning of last year showed that since 2006, sales of LP records have increased every year, with an average annual increase of approximately 43%. Remarkably, this currently makes the vinyl LP the fastest-growing music format, including all electronic formats. Whilst vinyl still only makes up a small proportion of total music sales, interest for the format has clearly been reinvigorated. This revival can be most clearly witnessed on 'Record Store Day', a now well-established date in the music industry calendar. Its purpose is to celebrate the culture of the independent record store, bringing together fans and artists, at events organised at thousands of stores across the world. Unsurprisingly the main focus of the day is vinyl, with fans converging to buy new or rare records, many of which are especially issued for the day. To begin my study, I spent a few hours at my local event in Nottingham, asking those present (and there were hundreds) to tell

[1] The first album to be released on CD was Billy Joel's 52nd Street, which was released in Japan on 1st October 1982, along with Sony's CDP-101 CD player.

me in their own words just what it is about vinyl, which had brought them into town at seven o'clock on a chilly April morning.

> *'What is it about vinyl? It's everything about vinyl isn't it? It sounds amazing. It feels good [...] It's what I'm used to! I've been buying it since I was that high. The artwork is beautiful. Everything about it. Everything that comes with it. The shops. The people. The communities. Oh! Where do I start? It's just brilliant isn't it?'*

Sentiments such as these (from local legend, DJ Rick Donohue) highlight that in addition to any extra qualities that vinyl may have as a means of sound reproduction; there are other 'hidden' virtues of vinyl, which are not so easily measurable. In addition, all of those I interviewed were more than happy to describe (in detail) not only their record collections but also the hi-fi setups on which they played them. A more detailed study of vinyl enthusiasts carried out by Nokelainen and Dedehayir [2] concurs, stating that

> *'The typical LP user today is highly interested in the technology itself... Moreover he or she appears to be (or perceive him- or herself as) technically proficient, probably a willing tweaker.'*

Coining the phrase, 'technostalgia', Jose Van Dijck [3] also discusses the importance of tweaking, suggesting that interactions with our playback devices are as much a part of our memory as the sound of the recordings themselves.

> *'Technologies and objects of recorded music are an intrinsic part of the act of reminiscence. Personal memory evolves through our interactions with these apparatuses and material things (records, cassettes, digital files), as both are agents in the process of remembering'.*

Thinking back to the albums I listened to as a teenager, I can still visualise the system on which they were played and even smaller details like the clicks that the radio buttons made and how I used to 'fiddle' with the tape counter whilst the records were spinning.

Vinyl is well known to be subject to imperfections (e.g. crackle, pop, warping, scratches, etc.) and it is suggested that these flaws somehow help to humanise the format and as the listeners are 'flawed human beings themselves', that they are therefore sympathetic to the medium. In the words of another respondent at Record Store Day, 'When you've got the crackle, when you've listened to your favourite vinyl so many times, it just makes it sound so much better.'

In the interests of a balanced argument it should be stated that these feelings are by no means universal and there are equally those that believe in the contrary, one commentator proclaiming vinyl as a 'fatally out-dated distribution system for the benefit of those who still have a stake in it'. [4]

However, the fact that at least some of the population seem to prefer the imperfect sound of vinyl, appears to completely contradict the notion of 'better' sound. How is it possible to expect superior sound quality, to emanate from a medium with all those probable flaws? Such a dichotomy suggests that we need to pose another key question; 'What do we mean by better sound?'.

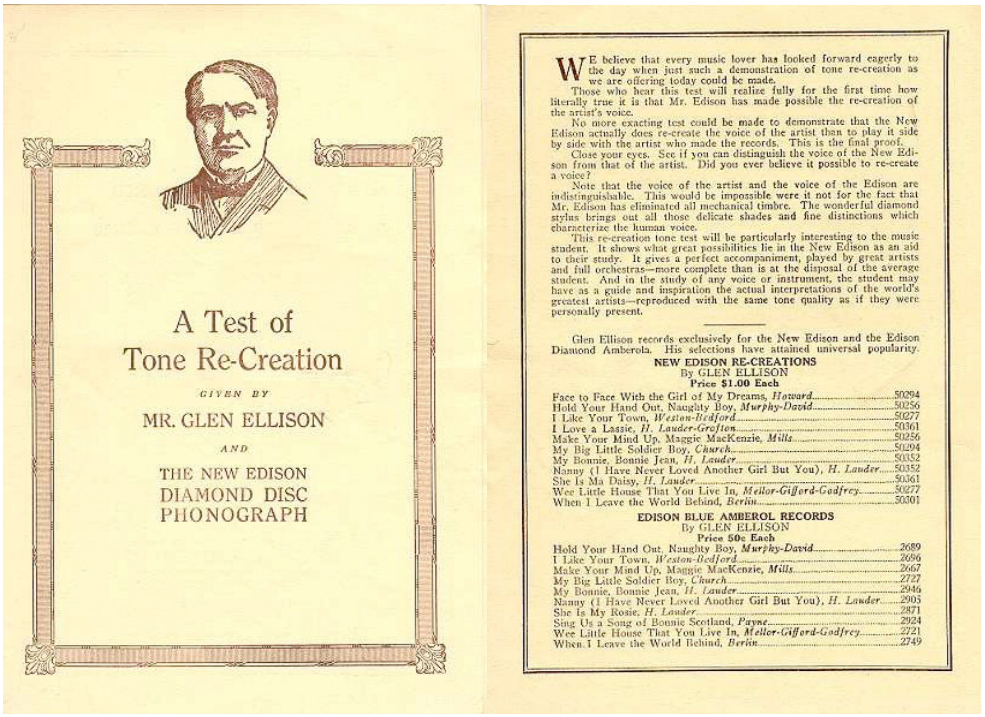
2. Sound: A Subjective Quality

'People will hear what you tell them to hear', Thomas Edison, c.1910. [5]

Sound is described as a 'periodic disturbance in the pressure or density of a fluid or in the elastic strain of a solid, produced by a vibrating object'. However, a sound is not heard until the pressure wave hits the eardrum and the resulting vibrations are 'decoded' by the brain. Applying the simple premise that no two people (and therefore no two brains, skulls, or pairs of ears) are alike, it follows that every listener will experience sound differently. Therefore, evaluation of sound quality is by definition a subjective matter, as it will not only be influenced by a person's physiological attributes, but also psychological factors which are shaped by their prior knowledge, experience, cultural surroundings, lifestyle choices and so on. [6]

One person in particular who understood and exploited the subjectivity of the listener was the inventor of the phonograph, Thomas Edison. Edison was rightly regarded as a genius of invention but he was also a shrewd businessman and clever marketeer. As part of the promotional campaign for his Diamond Disc Phonograph, he devised a set of public demonstrations, during which a renowned artist would take to the stage and perform a rendition of a song, alongside a phonograph playing back a recording previously recorded by the same artist. At a given point in time, the artist would pause and allow the phonograph to 'take over', much to the amazement of the assembled audience [7] The first of these 'tone tests' took place at the Monclair Club, New Jersey on 17th September 1915. Prior to the performance, Verdi E.B. Fuller (head of phonograph division at Edison's laboratory) took to the stage, proclaiming 'I shall demonstrate to you that the characteristic tone of every musical instrument can be faithfully re-created', adding that '…the reproduction of the human voice is equally faithful.' [8] Aside from Fuller's convincing rhetoric, there were other aspects of the test which were also carefully managed. Adverts were placed in local papers to publicise the events but in what would now be considered as market segmentation, attendance of the concerts was by invitation only, with Edison deliberately selecting specific, interested and influential members of society. To give the demonstration the feel of a special event, a programme for each concert was printed (figure 1) and the mood and atmosphere in the auditorium was also carefully managed, with lights kept deliberately low to add to the air of intrigue and excitement.

The aspect that should however come under the most scrutiny is the manner in which the audience's perception of sound was managed. As the phonograph's only means of amplification was a horn attached directly to the stylus, the discrepancy in amplitude between the machine and an operatic



A Test of
Tone Re-Creation

GIVEN BY

MR. GLEN ELLISON

AND

THE NEW EDISON
DIAMOND DISC
PHONOGRAPH

WE believe that every music lover has looked forward eagerly to the day when just such a demonstration of tone re-creation as we are offering today could be made.

Those who hear this test will realize fully for the first time how literally true it is that Mr. Edison has made possible the re-creation of the artist's voice.

No more exacting test could be made to demonstrate that the New Edison actually does re-create the voice of the artist than to play it side by side with the artist who made the records. This is the final proof.

Close your eyes. See if you can distinguish the voice of the New Edison from that of the artist. Did you ever believe it possible to re-create a voice?

Note that the voice of the artist and the voice of the Edison are indistinguishable. This would be impossible were it not for the fact that Mr. Edison has eliminated all mechanical timbre. The wonderful diamond stylus brings out all those delicate shades and fine distinctions which characterize the human voice.

This re-creation tone test will be particularly interesting to the music student. It shows what great possibilities lie in the New Edison as an aid to their study. It gives a perfect accompaniment, played by great artists and full orchestras—more complete than is at the disposal of the average student. And in the study of any voice or instrument, the student may have as a guide and inspiration the actual interpretations of the world's greatest artists—reproduced with the same tone quality as if they were personally present.

Glen Ellison records exclusively for the New Edison and the Edison Diamond Amberola. His selections have attained universal popularity.

NEW EDISON RE-CREATIONS
By GLEN ELLISON
Price $1.00 Each

Face to Face With the Girl of My Dreams, *Howard*	50294
Hold Your Hand Out, Naughty Boy, *Murphy-David*	50256
I Like Your Town, *Weston-Bradford*	50277
I Love a Lassie, *H. Lauder-Grafton*	50361
Make Your Mind Up, Maggie MacKenzie, *Mills*	50256
My Big Little Soldier Boy, *Church*	50294
My Bonnie, Bonnie Jean, *H. Lauder*	50352
Nanny (I Have Never Loved Another Girl But You), *H. Lauder*	50352
She Is Ma Daisy, *H. Lauder*	50361
Wee Little House That You Live In, *Mellor-Gifford-Godfrey*	50277
When I Leave the World Behind, *Berlin*	50301

EDISON BLUE AMBEROL RECORDS
By GLEN ELLISON
Price 50c Each

Hold Your Hand Out, Naughty Boy, *Murphy-David*	2689
I Like Your Town, *Weston-Bedford*	2696
Make Your Mind Up, Maggie MacKenzie, *Mills*	2667
My Big Little Soldier Boy, *Church*	2727
My Bonnie, Bonnie Jean, *H. Lauder*	2946
Nanny (I Have Never Loved Another Girl But You), *H. Lauder*	2903
She Is My Rosie, *H. Lauder*	2871
Sing Us a Song of Bonnie Scotland, *Payne*	2924
Wee Little House That You Live In, *Mellor-Gifford-Godfrey*	2721
When I Leave the World Behind, *Berlin*	2749

Figure 1 Edison's Tone Test Demonstration Programme (Courtesy of http://nipperhead.com).

voice at full volume, would have been considerable. The live performance would also have benefitted from the additional natural reverberation from the auditorium. Under these circumstances, it would have been very unlikely that the phonograph would have made much of a favourable impression. Anna Case (opera singer and Edison recording artist who performed at the most famous tone test at Carnegie Hall, New York in 1920) describes how she overcame the problem:

> *"The audience was there, and there was nobody on stage with me. The machine played and I sang with it. Of course, if I had sung loud it would have been louder than the machine, but I gave my voice the same quality as the machine so they couldn't tell." [8]*

It was not only the level which the singers imitated, but they would also strive to imitate the sonic characteristics of the record, such as the 'pinched' quality it lent to voices, due to the system's limited frequency response. This subtle inversion (or rather, perversion) of the supposed test, would of course have been intended to further impress the audience, rather than give them an opportunity to make a fair and informed comparison. This said, if one were to listen to a Diamond Disc phonograph recording now, it is hard to envisage how it could ever have been mistaken for the 'real thing' [9]. That is until one considers the listener's lack of cognitive experience of the situation they were witnessing. Al-

though the phonograph (and its competitor, the Victrola) were gaining popularity at that time, the audience were unlikely to have witnessed recorded music played from a stage and therefore, would have had no frame of reference. Add to this the afore mentioned aural illusions and his powers of persuasion, Edison was able to exploit the audience's naivety and his Diamond Discs were (at least according to the reports), very warmly received.

Inevitably, the success and notoriety of the tone tests resulted in many similar challenges being issued, most notably by audio-tape manufacturer Memorex. As part of their 1970s "Is it Live or is it Memorex?" campaign Nelson Riddle (band-arranger for Ella Fitzgerald) proclaims that he can't tell the difference between a tape recording and her live band. Unsurprisingly there are no details of the test, only the TV 'reconstructions' [10]. Both examples provide insight into how the listener's perception of sound quality can be subtly, yet deliberately influenced by non-auditory events, whilst at least purporting to be authentic audio tests. As we will discover, tests developed for the purpose of this study, were based upon a similar premise.

3. Listening Tests

3.1 Double-blinded by science?

Given the seemingly unpredictable and unreliable nature of our listener and that they may also be prone to other non-auditory biases, the next logical step for me was to try and establish once and for all, whether the sound of vinyl made any difference to the way people feel about their records. In order to do this, it was necessary to devise a test in which sound was the only variable and where no other factors could influence the result. Stuart Yaniger's excellent article, Testing One, Two, Three (Linear Audio Issue 2) discusses the challenges involved in designing and implementing suitable listening tests. Although it may appear obvious, the first step to devising any successful experiment is to be clear on exactly what you want to find out. In our case, I wanted to know whether the sound (and only the sound) of vinyl records was preferred to other digital formats, so I needed a test in which several examples could be auditioned and compared. To minimise the chance of bias being introduced the test also needed to be double-blind, hiding any clues as to the identity of the source from not only the participant but also the observer. Unlike Edison, I would have nothing to gain by the final outcome but my own preferences, pre-conceptions or expectations could unintentionally shape the way that I interpret results (the so-called observer-expectancy effect). In other words, in the interests of science, it's better that I don't know either!

So what were my options? Tests such as the ABX - in which the listener is invited to compare an unknown sample (X), with two known reference stimuli (A and B) – are often used to support or refute the claims of those advocating various technologies, formats and playback systems [12]. If the subject cannot tell the difference between two stimuli, then any assertion of a preference of A over B (or vice versa) must be declared null. However, the fact that ABX testing does not enable the listener to express a preference, means that it will be an unsuitable method for use in this study.

The ABC-HR methodology (ITU-R BS1116-1) enables comparisons to be made, whist validating the quality of the data. This is achieved by introducing two references, A (known) and HR (unknown) - the 'hidden' reference. The two samples under test are B and C, both of which will be variations of the original, a result of an audio process (e.g. down-sampling, application of lossy compression technique, etc.) The subject is invited to compare B,C and HR (the order of which are randomised) to the reference A, scoring each on a continuous scale. The interface used would typically be a computer programme or applet, which enables the participant to audition samples in any order and as often as required, before entering their responses. A reliable subject, should be able to find the HR and score it (within an acceptable margin of error) the same as A. If this condition is met then the remaining scores given, can be regarded with a much greater degree of confidence. A drawback with the ABC-HR methodology is that the highest rating on the scale is defined where the difference between the test and reference is 'imperceptible'. This is suitable for testing compression algorithms, as it is safe to assume that the original, un-compressed file represents the ideal standard. However, there is no scope for a sample to be scored higher than the reference – an outcome which must be offered as an option, if our subject were to perceive that the vinyl pressing sounded 'better than' the source files.

The MUSHRA (Multiple Stimuli Hidden Reference and Anchor : ITU-R BS.1534-1) methodology represents a further improvement as it can incorporate more than two test files, again allowing for the order to be randomised, whilst allowing the participant to switch freely between test and reference files. It also incorporates another hidden file, the 'anchor' - a deliberate, clearly identifiable sample of a quality lower than any of the other test files, which serves as an additional check for unreliable data. The premise here is that if our participant scores the anchor higher than any of the other samples, then one would need to question the validity of their responses. However, for our scenario choosing the anchor proved to be something of a conundrum, as this would require us to make assumptions about which format our subjects perceived to be the worst. As we have already learnt, some people prefer the sound of crackly records! It was decided to omit the anchor file and instead use two consecutive tests, both of which included a hidden reference and a copy of one of the test files - which was in our case taken from vinyl. If the participant scored the vinyl significantly differently in the first and second tests – we could again question the veracity of their responses. This post-selection process in which such unreliable subjects are removed, improves the overall quality of our data. With our additional modification to remove the anchor, the MUSHRA (or should that be MUSHRANU? – anchor not used!) seemed like a good choice for our tests.

There was however one major, contentious issue to be resolved. Unlike ABX (which have been configured using physical relays), MUSHRA tests need to be deployed digitally. Although theoretically possible, it would in practice be almost impossible to implement a 'truly analogue' test because either the participant or the observer would need to physically interact with the vinyl (or other formats). This would of course reveal the identity of the format being tested and the possibility of bias being introduced.

Would it be acceptable then, to use a digital test to evaluate the sound of an analogue format? On the face of it the idea may appear preposterous but consider the flow diagram below (**figure 2**), which represents the different paths which two test signals would take, in order to be included in such a test. Disregarding for now the additional digital-to-analogue-to-digital conversions required, it can be seen that differences between our two test files will arise from various processes required to cut our record (amplifier, lathe[2], diamond disc-cutting stylus) and then play it back (turntable, stylus and hi-fi amp). Also included in this 'system', is the vinyl itself - as any changes or defects (e.g. static, scratches, warping) would too alter the sound of our final files. The contentious issue is that of the additional analogue-to-digital conversions, which so often appear at the centre of the argument. How can we possibly overlook their influence on the perception of final audio quality?

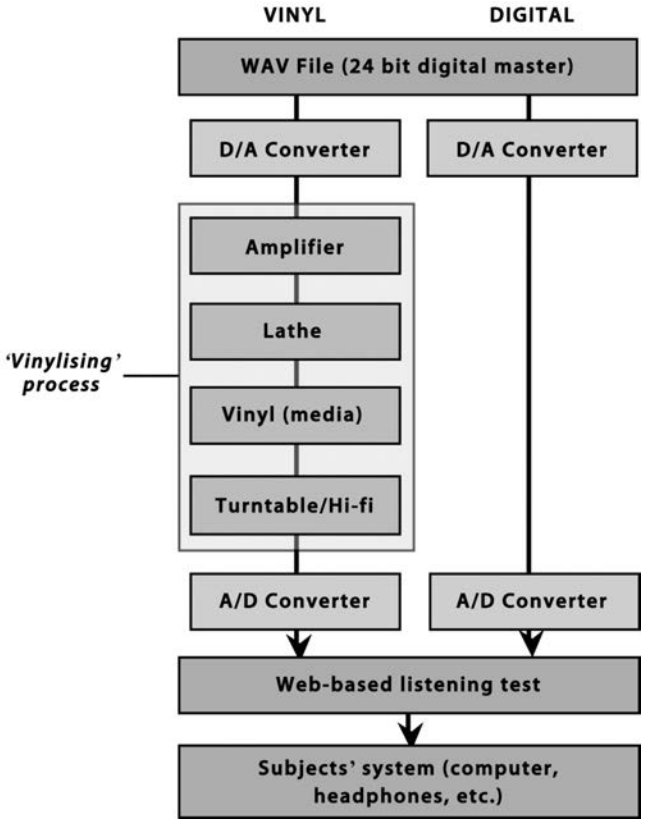


Figure 2 Differences in signal path between vinyl and digital test files.

[2] Vinyl manufactured on a small scale (i.e. one-off test pressings used in this study) will be lathe-cut as opposed to stamped. Records for this study were cut by the engineers at Vinyl Carvers (www.vinylcarvers.com).

3.2 Analogue vs Digital? Please change the record!

> *'From the early 1980s up till now, and probably for another fifteen years to come – this is the darkest time ever for recorded music. We'll come out on the other end, and it'll be okay, but we'll look back and go, 'Wow, that was the digital age. I wonder what that music really sounded like?'. We got so carried away that we never really recorded it. We just made digital records of it.' (Neil Young, 1992)*

Since the introduction of the Compact Disc in 1982, many arguments questioning the quality (or even legitimacy) of digital recordings have been made. Assertions such as Young's and other well-respected music industry professionals, such as producer Steve Albini have been widely published and intentionally or otherwise, have perpetuated certain myths about digital audio. The interviews at Record Store Day have already revealed the strength of the emotions which vintage technologies can evoke in musicians, audio professionals, hi-fi enthusiasts and music fans alike. Having already acknowledged the subjective nature of sound quality measurements, we must also accept that an individual has every right to perceive one format as superior to another, even if the empirical evidence suggests otherwise. The problem with many of the 'anti-digital' arguments that have entered into the public domain, is that they are often based upon pseudo-scientific misinterpretations of digital audio theory and recording practice in general. We've learnt how TAE (Edison's recording company) and Memorex, were able to make claims that appeared to be supported by empirical testing, whereas in reality the tests were carefully constructed to further their own agendas. Such claims tend to hold more sway in the minds of audio enthusiasts, especially when they are made by influential figures. Let us consider in turn three of the most common negative assertions, made in relation to digital audio:

1 *As a digital audio signal is comprised of discrete samples, information which falls between the samples is somehow 'missing';*
2 *A specification of 16bit / 44.1kHz is inadequate;*
3 *Digital audio does not have the 'warmth' of analogue.*

Creator of the Ogg Vorbis audio codec Chris Montgomery, provides a comprehensive rebuttal of the first two points and several other supposed truths about digital audio. [13]

> *'All signals with content entirely below the Nyquist frequency (half the sampling rate) are captured perfectly and completely by sampling; an infinite sampling rate is not required. Sampling doesn't affect frequency response or phase. The analog signal can be reconstructed losslessly, smoothly, and with the exact timing of the original analog signal.'*

Although his statements are equally as bold, in contrast to those made by the proponents of analogue formats, they are supported by theory and empirical evidence, demonstrated by specific and repeatable scientific experiments. In particular he tackles the misconception that a digital audio signal is 'a jagged, hard-cornered stair-step facsimile of the original perfectly smooth waveform', ex-

plaining that it is just a convenient visualisation and one that has been taken far too literally, even by the community of audio engineers. The supposed limitations of sampling frequencies [14] are also addressed in reference to the capabilities of the human hearing system and whilst it is conceded that sample rates higher than 44.1kHz can be beneficial, it is only because they will allow for the use of gentler, less 'harsh' sounding anti-aliasing filters, rather than an increase of the system's frequency response.
At this point, it is worth restating that the aim of this work is to study only the listener's perception of the sound quality of vinyl records, in comparison with their digital counterparts and not to re-enter the deeper discussion of the merits of analogue and digital recording per se. However, the evidence offered here does strongly support our suggestion, that the additional conversions required to include our vinyl test files in an on-line audio quality test, will be entirely transparent - as long as equipment of a high enough fidelity is used and correct audio engineering practice is adhered to.

3.3 Warmth

Opponents of digital audio often bemoan the lack of mysterious, abstract quality that even the most experienced of audio engineers struggle to explain, 'warmth'. In relation to audio, Robjohns [15] definition is a useful one; 'the character that the analogue processing/recording equipment and the recording medium add to the sound'.

The character in question is usually some form of harmonic distortion added by microphone preamplifiers, vacuum-tubes, tape machines, etc.

> *'Harmonic distortion is the introduction of extra harmonics that are musically related to those already present, resulting in a change in timbre. Even-order harmonic distortion tends to sound musically sympathetic, smooth, and bright in a constructive way.' [16]*

That analogue recording equipment can add desirable sonic characteristics is not disputed (although it is recognised that digital audio workstations can now perform very convincing emulations) but there is often a misunderstanding of where in the production process, these characteristics are added.

If our records do appear to sound warmer (or fuller, heavier, fatter - or any other descriptor we care to substitute here), it is extremely unlikely to be a consequence of the final cut to vinyl. Unlike recording to analogue tape, where additional gain may be carefully applied in order to achieve tape-saturation, the process of cutting to vinyl is a controlled process in which the optimum signal must be sent to the cutting tip. If too much force is applied, then the cutting tip will simply cut too deeply and the resulting master disc may become unusable. If intended, it is therefore far more likely that distortion has been introduced elsewhere in the production chain. During playback however, the mechanism (the stylus on a turntable) will introduce mild harmonic distortion into the signal. Again, the distortion can be pleasing to the human ear but hi-fi enthusiasts tend to associate *lower* measurements of Total Harmonic Distortion (THD) with better quality equipment.

4. Preparing the Test files

4.1 The Quiet (and not-so Quiet) Life

Having chosen (and justified) our choice of test method, it was time to move on to the next stage, preparing the test files. The first stage of the audio production process was the creation of a new vinyl LP using one of my own production archives, ensuring that I could be absolutely certain of the quality and identity of the master used to cut the vinyl from. Would not the same have been true, if I were able to purchase a brand new LP on vinyl and the same album on CD? The answer is no. Without access to the engineers' notes (assuming they were 100% accurate), there would be no way of knowing whether the vinyl LP and the CD were produced from the same master mix, even if they were part of the same release schedule.

To illustrate this point I performed a small experiment, involving one of my favourite songs from the 1980s, 'Quiet Life' by seminal British new-wave band, Japan. I gathered every different release of the song in my possession, eight in all; the original 1979 (vinyl) release, six different versions on CD (three from different releases of the album and three from compilations) and an MP3 version, purchased from the iTunes store. The vinyl was digitised to 16bit/44.1kHz and CD tracks were converted to WAV files, as was the MP3. **Figure 3** shows the tracks in an audio editor, placed left-to-right in chronological order of release. Differences in the waveform can clearly be seen, especially in the later releases. Below each waveform are the LUFS[3] measurements for each file. As can be seen, loudness measurements vary by up to 4.1dB and only the first two files are within 0.2dB (the limit for subconscious awareness of each other [11]. This would mean that comparison of our vinyl sample with all-bar-one of the other files would, due to the influence of loudness, be unreliable.

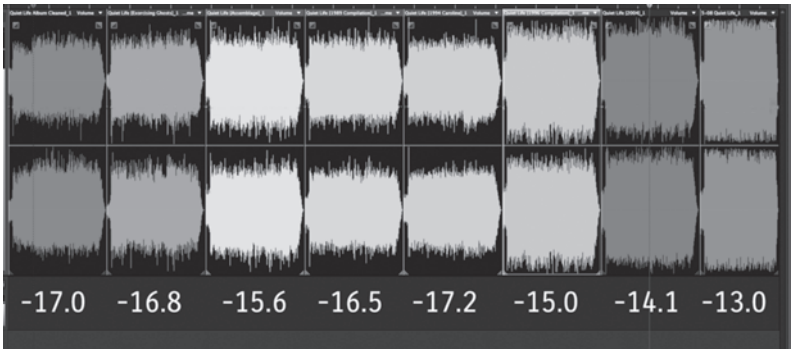


Figure 3 Loudness comparison (in LUFS) of vinyl and CD releases of Japan's Quiet Life (NB : All files were first peak normalised to -0.1dB, Vinyl copy at far left).

[3] 1 LUFS = 1dB. In this context, the units are interchangeable.

Subsequent frequency analysis also revealed some significant variations and although not documented in detail here, the indication was that on each occasion the engineer has responded to the track differently and has (either intentionally or unintentionally) produced a master, with differing EQ characteristics, which will in turn be subject to the preferences of the listener. There was one more surprising discovery to be made in amongst the digitised tracks. **Figure 4** shows stereo vector-scopes of two versions, the first from a 2004 re-release of the album itself and beneath, one taken from a version included on a 1996 compilation.[4] Remarkably, the latter was mastered with the right hand chan-

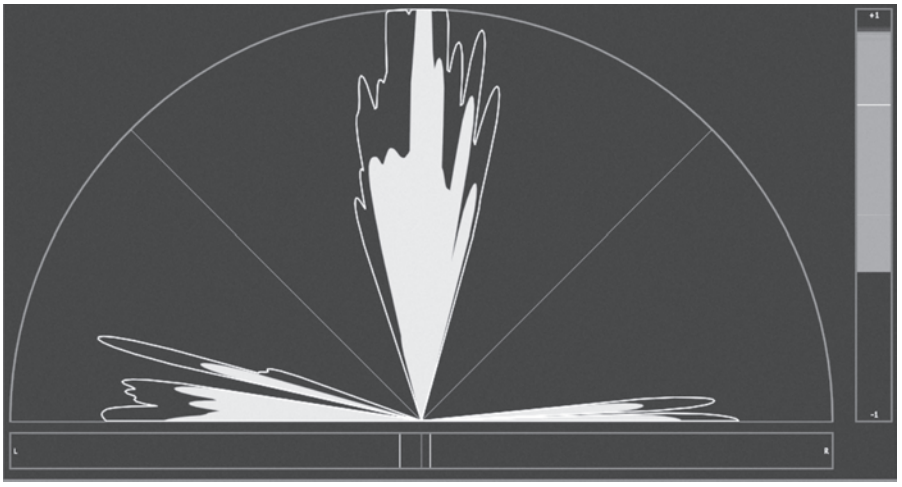


Figure 4a Stereo field of 'Quiet Life' (from remastered album re-release in 2004)

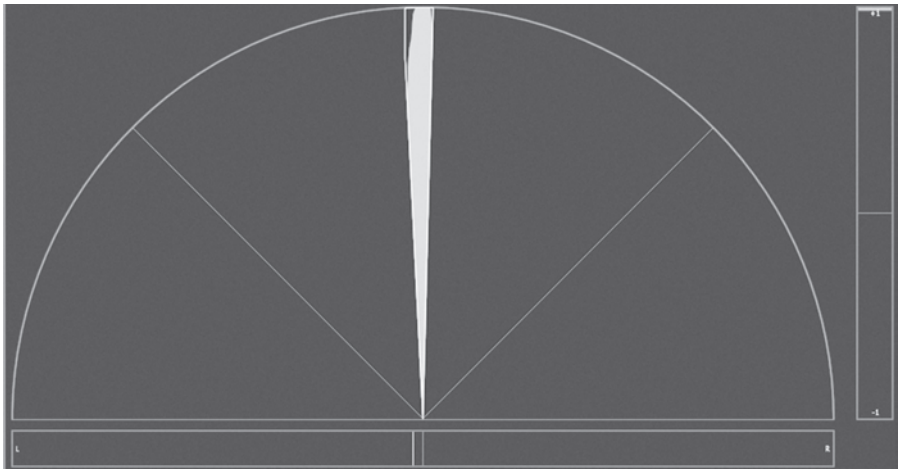


Figure 4b Stereo field of 'Quiet Life' (from compilation released in 1996).

[4] 'Japan: In Vogue' Camden/BMG 1996

nel, repeated in the left - a perfectly mono mix (albeit 'leaning' slightly to the left) and all tracks on the compilation suffered the same issue. Clearly, a serious error had been made during the production process (and completely missed at the QC stage too) but even so, it underlines how different two commercial recordings could be, even if the files are derived from the same master of the same mix!

4.2 Making records

A total of eight newly created masters were chosen for inclusion on a new vinyl LP. Ideally, the album would have been sent to a plant for a large-scale pressing, but with a minimum order of five-hundred (and my old record company not picking up the phone…) this was not economically viable.

Instead, a one-off vinyl LP (or dub-plate) was lathe-cut [17], a process in which the cutting tool is applied directly to a disc and the groove directly etched by the incoming, amplified audio signal. A dub-plate is actually made of a PVC compound, which is very similar to pressed vinyl. The material is durable and will withstand repeated plays but the sound quality of cuts can be very dependent on the actual lathe used. However the same can be said about any commercial-pressing, the sound of which can vary due to the properties of the vinyl used, the quality of manufacturing plant, etc. One clear advantage of using lathe-cut vinyl, was that I was able to ensure that no additional processing or equalisation was applied, other than the basic gain optimisation and RIAA pre-emphasis required to facilitate the creation of the record.

To re-digitise my vinyl, I booked a session at the mastering studio[5] which was used for the original CD release[6], where the equipment was all of a high specification and properly maintained. The mastering engineer was also vastly experienced and having worked with me during the original mastering sessions and was able to offer another pair of ears and further expert guidance. The turntable was calibrated using a set of test records and after an initial play though to check for any obvious problems (there were none and both the engineer and I agreed that it was a very good cut), both sides of the album were digitised at 24bit / 96kHz resolution.

4.3 Vinyl restoration

Two songs were then chosen for inclusion in the listening tests - 'Charm Offensive' and 'Protect and Survive' (this was a personal decision - I believe that they are the best mixes). On first playback of the digitised vinyl, it was noted that there was a slight, persistent noise at the low-end of the spectrum, which predominated on one side. This was suspected to be ground hum, which was confirmed by our meters, which showed a clear peak at 50Hz. The hum was removed using a suitable comb-filter EQ, which also reduced the effect at multiples of 50Hz (the harmonic series).

At this stage it was also decided that large, clearly noticeable crackles, clicks and pops should be removed. Although it has already been suggested that such imperfections may be part of the allure of

[5] Formation Audio - http://www.formationaudio.co.uk
[6] 'Protect and Survive' by Manuskript, Resurrection Records (2006)

vinyl records, it is a view that cannot be considered universal. Removing noticeable, momentary imperfections would enable our vinyl dub-plate to closer approximate a 'first play' from a pressed record. Crackles and pops are also rather obvious cues, which could influence perceptions.

The pitch of the record also appeared to be slightly higher than the source files and this was confirmed by tempo analysis using a sequencer. As the playback deck had already been carefully calibrated (using the test tone records), it was suspected that the lathe used to cut the disc was responsible. It was however possible to correct the speed of the captured files by tempo analysis, using an audio editing package (http://www.diamondcut.com) specifically designed for vinyl restoration. Hum removal, click removal and speed correction were all applied (in that order), to give us two corrected, digitised vinyl examples for inclusion in our listening test.

4.4 The magic of vinyl! Discovered?

Before starting the listening tests some additional audio analysis was carried out, as I wanted to investigate what the process of being pressed into (and played back from) vinyl had actually done to each of the two test songs. The premise was that if the variations between our source (original 24bit master) and destination (the digitised, corrected vinyl) files could be discovered, then it may also be possible to recreate the desirable sonic properties of vinyl by applying EQ and/or effects, without ever needing to leave the digital domain. Although there are several plug-ins and gadgets that purport to emulate the sound of vinyl, I have always found these to be rather gimmicky, the majority just re-introducing the imperfections which had just been painstakingly removed. Instead, we focused on three more fundamental parameters; frequency spectrum, stereo width and harmonic distortion.

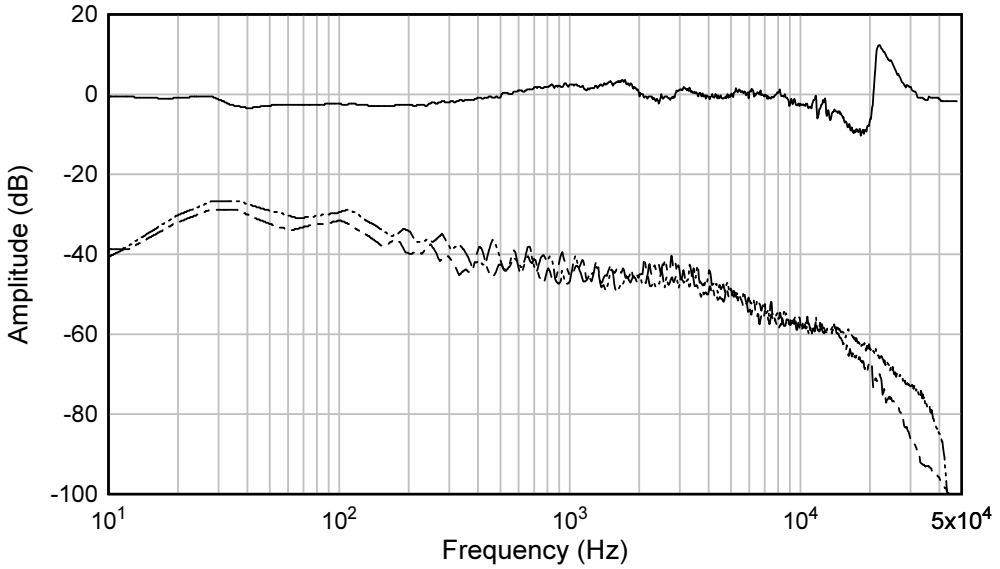


Figure 5 Use of the Spectral Difference function to match frequency characteristics.

Firstly, a spectrum analyser was used to measure the overall frequency responses of both the source and destination files, producing a 'difference curve' which represented the variation between the frequency spectra of the two files. This curve (**figure 5**) was then be applied to the source file enabling us to super-impose the EQ characteristics of the 'vinylising' process.
Stereo width was the next property to be analysed. A standard compact disc contains one or more interleaved audio files – yet the term interleaved only refers to the data structure of the file and the audio signal in the left and right channels remain independent. Therefore, a stereo mix-down onto CD can maintain maximum separation between channels (e.g. 'hard panning'). This is not the case with vinyl, as the pick-up coils for both the left and right channels are physically attached to the stylus [18]. This results in a certain amount of cross talk between the channels, which has the overall effect of narrowing the stereo width. Some advocates of vinyl argue that the restriction of stereo width helps to instil better mastering practices (e.g. not using excesses of bass in wide panning positions) to prevent problems being caused and generally help to focus the mix. As Dorsey describes, "Cut too much low frequency information with a wide stereo spread, and you get a lot of deep peaks and valleys in the groove and styli tend to pop out of the groove." [18] The stereo width of both the 24bit master and the digitised vinyl were measured using a stereo vector-scope, with the digitised vinyl showing a reduction in stereo width of approximately 25%.

Defined earlier as a result of harmonic distortion, we next consider that much vaunted characteristic warmth and the perceived effect that its presence (or absence) will have on sound quality. Harmonic distortion can be measured using frequency analysis, by passing a sine wave of a known frequency through a system and then looking for additional peaks in the spectrum, at exact multiples of the fundamental. For audio systems, the total harmonic distortion (THD) can give a very good indication of its linearity. It is calculated as a percentage as follows:

$$THD = \sqrt{\frac{V_1^2 + V_2^2 + V_3^2 + \ldots + V_n^2}{V_t}}$$

where:
V_n = RMS Voltage of each harmonic distortion
V_t = Total RMS Voltage

To investigate the harmonic distortion of our system, we used a series of pure tones (ranging from 62.5Hz to 15kHz) captured from test records. Figure 6 shows the spectrum analysis of a 300Hz pure sine tone generated in the mastering studio, cut to acetate and pressed to vinyl. The fundamental and the harmonic peaks can clearly be seen, all the way down to the eighth harmonic. The THD for this system at 300Hz was calculated as 5.295%. Readings were taken from the graph and tabulated, deriving the relative amplitude of the peaks for each on the nth-order harmonics. This was repeated for each of seven test tones, giving us measurements of distortion at a range of frequencies.

The resulting data was used to approximate the warmth of our vinyl system using software which could introduce distortion, with independent control of the gain for each harmonic (http://www.pcjv.de/vst-plugins). A multi-band distortion processor was built by using seven instances of the 'Christortion' plug-in and tested by applying it to a newly generated 300Hz pure-tone. The THD of this emulated 'warmth' was shown as 5.575% - a figure very close (within 0.28%) to the

THD measured from vinyl.

As a result of the analysis undertaken here, we found that even after removing the ephemeral nuances of vinyl, there were indeed clearly defined and measurable differences between the source and destination files. The measurements taken therefore made it possible to create 'emulated vinyl' versions of the masters by:

1. matching the frequency spectrum;
2. narrowing the stereo width, and
3. adding harmonic distortion.

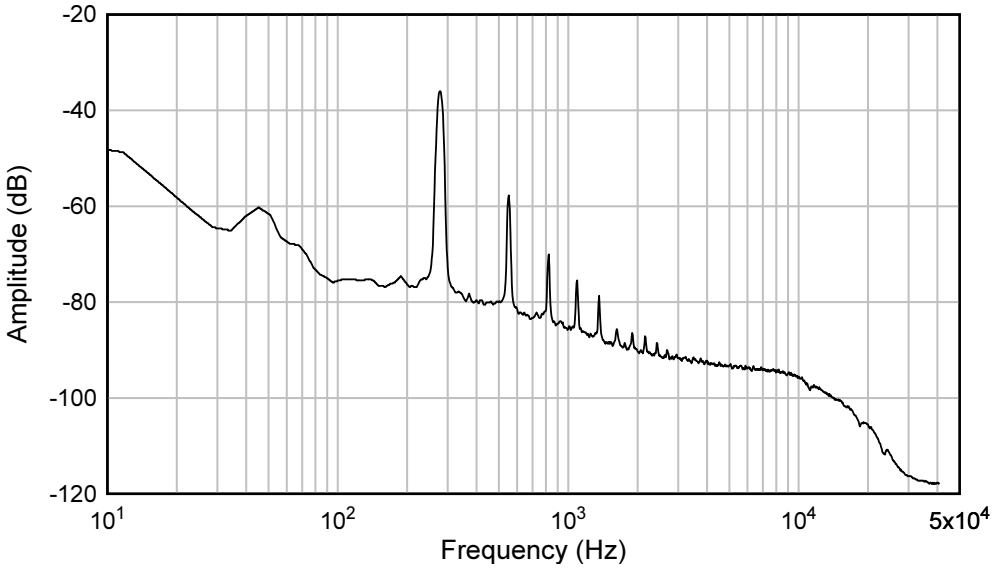


Figure 6 Spectrum analysis: peaks showing fundamental and harmonic distortion @ 300Hz.

So, now it was time to put the files in the hands (or ears) of our participants and learn about their perceptions of sound quality.

5. Final Listening Test

5.1 Designing the interface

As mentioned above, the MUSHRA methodology stipulates that playback of test files should be randomised and that a second copy of the reference file must be hidden amongst the others. Another requirement is for the listener to be able to audition audio files in a critical manner, which means that repeated plays, rewinding, switching back-and-forth between files is not only permitted, but

encouraged. This type of operation is not possible using standard, on-line survey tools (e.g. Survey Monkey, etc.) but an existing web-based framework for MUSHRA testing – mushraJS - enabled a custom web-site to be built, meeting all requirements and allowing the test to be taken by anyone with a computer and pair of headphones (**figure 7**). [19]

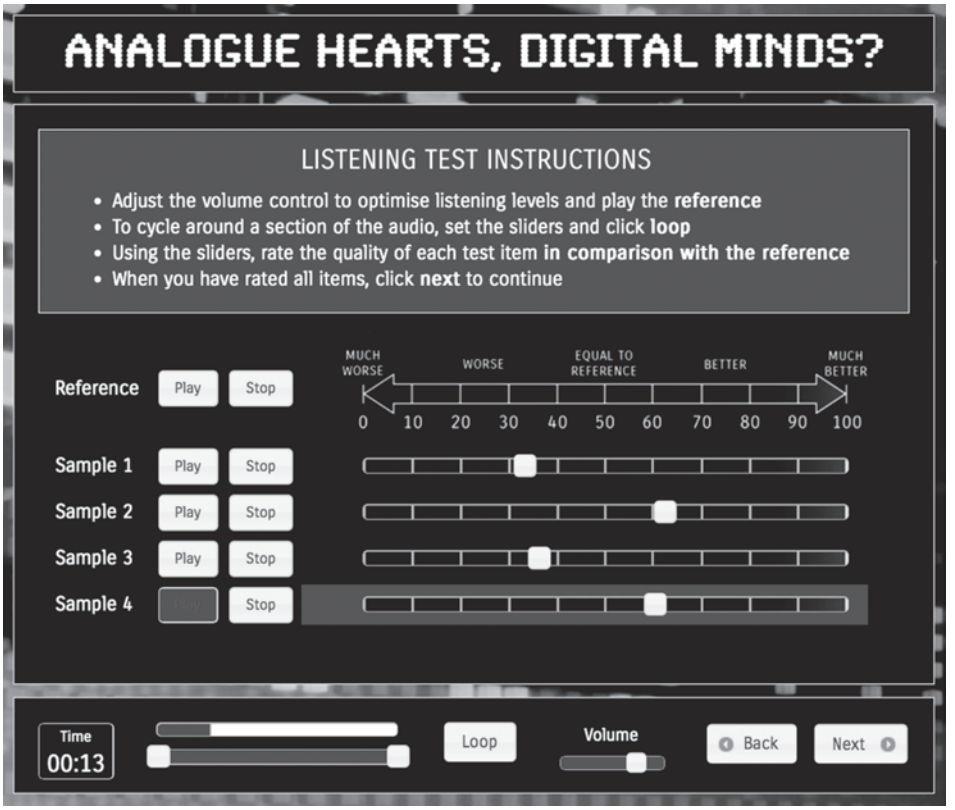


Figure 7 Graphical User Interface for MUSHRA-based listening test (http://analogueheart.co.uk).

An obvious disadvantage of using remote, internet-based blind testing, is that there is no possible way to ensure the correct configuration of the participants' hardware and therefore to be certain of the integrity of the audio signals being rated. It is also highly likely that headphones used will be of varying brands and quality, which would of course have direct consequences for sound quality assessments. However, there is growing belief that by making use of the subjects' own audio equipment, we are creating an 'ecologically valid environment' [20], producing results that translate best into 'real world' contexts. In other words, we allow the participants to listen to examples using their own system, as this is how they would usually listen to music. Listening tests were also based upon comparative ratings (with respect to the reference), so listening conditions would have been consistent for each subject, as long as they did not change their headphones or volume settings, during testing.

Song 1 : Charm Offensive		Song 2 : Protect and Survive	
Test A	**Test B**	**Test C**	**Test D**
Digital Master (Hidden Ref)	Digital Master (Hidden Ref)	Digital Master (Hidden Ref)	Digital Master (Hidden Ref)
Sampled Vinyl	Sampled Vinyl	Sampled Vinyl	Sampled Vinyl
Original CD Master	Emulated Vinyl	Original CD Master	Emulated Vinyl
MP3 (192kbps)	Original Studio Mix	MP3 (192kbps)	Original Studio Mix

Table 1 Audio files included in the web-based listening test.

5.2 Final list of files

The final list of audio files, which were included in the web-based listening test are given in Table 1.

In tests A and C, we investigate the perceptions of vinyl in comparison with its source file, and also two other common digital formats. In tests B and D, we instead seek to establish whether mastering the source file with our 'vinyl emulator plug-in chain' is deemed pleasing to the ear, or if the original studio version (the 'raw' studio mix) was preferred.

In order to decrease loading times and to reduce the overall amount of listening which each participant must undertake, it was decided to edit the tracks down to one-minute extracts. Already confident that there would be no perceptible consequences to the audio, files all were presented at 16bit / 44.1kHz. Finally, all files were loudness matched to the integrated LUFS value of the reference file.

6. Usability Test

6.1 Rationale

The aim of this part of the investigation was to test whether non-auditory factors could also influence the listener's perceptions of vinyl. I wanted to observe subjects' interaction with the format, the packaging and the artwork, and then ask follow up questions about their experiences of the test and also their general views on the format. Double-blind testing would have required both the turntable and digital source to be somehow hidden from view, which would have precluded the subject interacting with the records, thus negating our objectives. Even if I had engineered a situation in which the participant chose the record and took it from its sleeve before handing it over to an assistant to be played - the vital element of 'dropping the needle' would have been lost and as we have already learnt, interacting with technology is an integral part of the experience. So how does one play a record, without actually listening to it? The solution arrived in the form of technology developed for the most ardent supporters of vinyl.

6.2 Digital Vinyl. Better than the real thing?

A DVS (Digital Vinyl System) is one in which vinyl is still used but instead of the disc containing an analogue audio signal, it is pressed with a digital time-code. These discs can be played back on standard turntables, with the code being interpreted by a computer connected to the turntables via a suitable audio interface. The code reports the stylus' current position, its speed and the direction in which it is travelling, to a piece of audio software. In practical terms, this enables any digital audio file to be played and manipulated as though it was pressed on vinyl. For our intended purposes the use of a DVS was ideal, as it can provide an authentic vinyl user experience, but with an audio output theoretically identical to the digital system being compared. If you have not seen a DVS system in action, you should take a look at Mastahanksta [21] and see if you can see (or hear) the difference!

For the digital playback system, I opted for a standard CD player. Although declining in popularity, the compact disc is a well-established format and it was envisaged that all participants would be familiar with its operation. Playing a compact disc also has its own associated ritual and the comparison between putting on a CD and cueing up a record, is a valid one to observe.

From the subjects' perspective, the test was very straightforward. They were asked to play and listen to first a CD and then a DVS record and in the process, to observe and interact with the object and the playback system. The test was in essence a modern, interactive variation of Edison's tone tests. The main difference with our situation being that we already knew that the audio playback on both occasions would be identical, suggesting that any observations to the contrary must be due to the influence of other non-auditory factors. When Edison staged his demonstrations, he made every effort to ensure that they were perceived in the same manner that a 'real' concert would be, issuing tickets, setting the stage lighting and publishing programme notes. I also took steps in order to recreate a genuine user experience, reworking the original artwork and designs, to make an album sleeve, liner notes and disc labels. These were created in the exact style of the CD, although scaled-up for the 12inch format. The audio used on this occasion, was an exact copy of the track already on the CD, imported as a 16bit / 44.1kHz stereo WAV file.

I built my usability test around an open-source DVS system (http://www.mixxx.org), which was compatible with my existing audio hardware and a variety of time-code vinyl. Figure 8 shows the set-up used in the tests. The turntable fed the DJ mixer (used to amplify and apply the RIAA de-emphasis and also to add a little extra visual authenticity), which in turn fed the audio input of the iMac. The output level of the mixer was adjusted to provide the input required by the software. The chosen track was then loaded onto the software's first 'deck' and the vinyl control (i.e. time-code) function was turned on. Pilot runs were successful, although the tempo of the track on vinyl was observed to be approximately four beats-per-minute faster than the CD. Again the tempo was corrected, this time using the pitch control on the turntable. The importance of loudness matching has already been discussed and this was also checked at the beginning of each test.
During the tests, I waited until the subject had started playback and then manually switched the

control mode in order to 'lock out' any further pitch and tempo changes and thus preventing needle skips. This also helped to stabilise the pitch on the turntable used (which was belt-drive and slightly more prone to pitch variations). One other aspect that did perturb, was that the signal triggered by the time-code vinyl was so completely free of static, that the lead-in at the very beginning of the record actually sounded too clean. To remedy this, I resorted to some Edison-style trickery and added a very slight amount of un-modulated groove noise (the lead in recorded sampled from true vinyl) to the very beginning and very end of the each track stored in the DVS playlist.

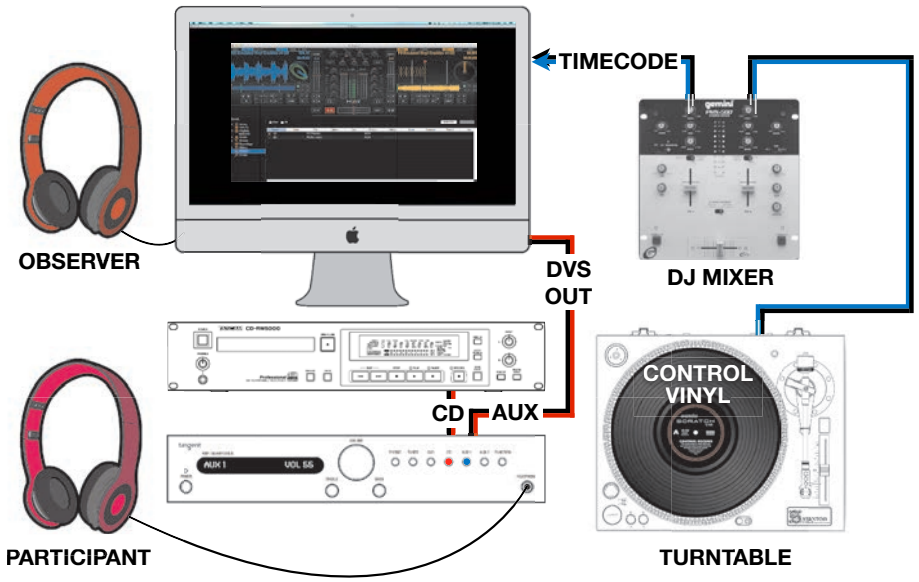


Figure 8 System schematic of DVS-based vinyl usability test.

At the beginning of each test, the subject was introduced to the equipment and briefed, before being given both versions of the album and invited to continue with the test at their own pace as per the guidance sheet and when they had finished, I asked some simple questions about what they heard… and what they felt.

7. Results

7.1 Listening Test

The listening test was deployed on-line, with a total of 59 responses gathered during the test period. The inclusion of the hidden references in our test methodology, enabled post-selection of subjects to be carried out, disregarding the results from those subjects whose responses fell outside a set of rejecting criteria. There were two rejection criteria set:

- *If there was greater than 1 point (or 20%) difference between the listener's scores for the hidden reference and actual reference. This criteria was applied for all four tests independently.*

- *If there was greater than 1 point (or 20%) difference between the listener's scores for the sampled vinyl references, in each pair of tests which featured the same listening material - e.g. Test A and B ('Charm Offensive'); and Test C and D ('Protect and Survive').*

Following post-selection, the results for the four listening tests were presented as mean values of the ratings for each of the included audio examples, with confidence intervals either side of the mean. A normal distribution was assumed with z-score chosen as 1.96, to give a 95% confidence level. [22]

Post-selection results are given as a set of plots, showing mean and upper and lower confidence bounds (**figure 9**).

7.2 Usability Test

The questioning which followed user interactions with the DVS vinyl system and CD player, provided further quantitative and qualitative data. Again, the answers given (full transcripts of which are available on the project web-site (http://analogueheart.co.uk) revealed the listeners' perceptions and preferences, with regards to the consumer formats being studied. In total, thirteen subjects took part in the lab trials of whom eleven also participated in the listening test, which allowed further comparison between the two tests to be made.

7.3 Analysis

Results from the listening tests overwhelmingly suggest that for our sample population, the digitised vinyl was the least favoured of all the formats. Across the four tests, its mean ranged between 2.44 and 2.79, clearly corresponding to the classification 'worse than reference', as defined on our modified MUSHRA scale. Those who favoured vinyl were in the minority and the format was also ranked last more often than any other format. In all cases, the highest mean score was attributed to the hidden reference, the newly re-mastered digital version. Note that confidence intervals for the hidden reference are narrower but this we can attribute to the post-selection process, in which respondents who made inaccurate ratings of the hidden reference, were removed from the results.

MP3 versions were judged favourably in both of the tests in which they appeared, with the mean values being 4.95 and 4.68 for Tests A and C, respectively. Surprisingly, the MP3 versions also outscore the CD versions (and by a clear margin for Test A), as both were created using an identical source file. Data would appear to suggest that this unexpected result may be attributed to Test A only. If taken on face value, then our result indicates that the process used to encode the song 'Charm Offensive' to a 192kHz MP3, yielded a better (if not equivalent) result than the dithered version down-converted to 16 bit/44.1kHz (for CD). This result was seen as something of an anomaly and therefore, the CD and