"REALITY? OR SOFT FOCUS?"

Tony Andrews Funktion One Research Limited, Dorking, Surrey, UK

1 INTRODUCTION

I've called this dissertation "Reality? Or Soft Focus?" because events over the last few years have given me cause to believe that professional audio has lost its way. In this paper, I will endeavour to explain my reasoning for the position I take over this sad state of affairs. It is my concern that we are unconsciously heading in the direction of smeared indistinct sound (soft focus) as opposed to striving for crisp intelligible sound (reality).

The main difference between the two perceived conditions of reality and soft focus is a sound system's transient response. If the transient response is poor then the audio will be soft focus. Definition and dimension will be missing, speech will lack intelligibility, musical instruments will be confused and the sound source will seem distant, culminating in what can only be described as audio mush.

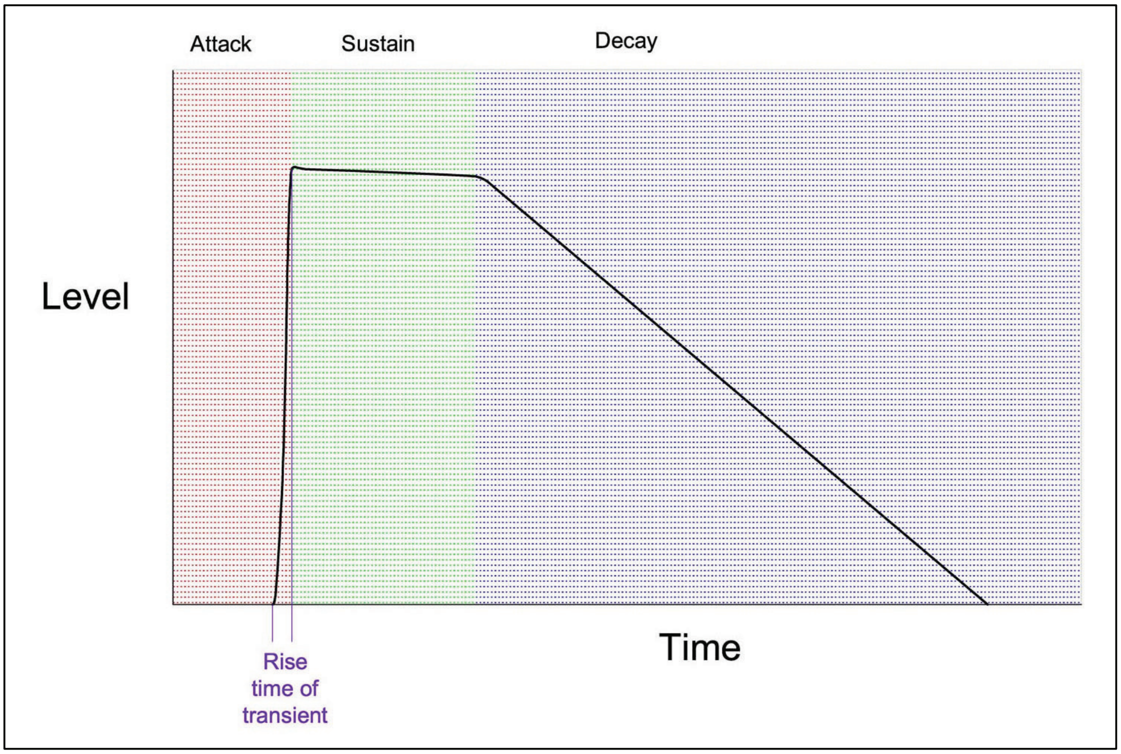
Transient information is the definition and excitement of audio and is as important a parameter as frequency response if not more important.

2 TRANSIENT RESPONSE

2.1 Definition

The word transient as used in audio usually means a short period of peak level with a fast rise time, such as a snare hit. To my understanding transient response is the ability of any part of the audio chain to accommodate the dynamic range of a given wave form and faithfully reproduce the leading edge. Attack is the nature of the leading edge and rise time is the time taken for the signal to go from minimum to maximum level. Incidentally transient response in amplifiers is known as slew rate and is measured in volts per micro second. Sounds are quite variable in their transient demands, contrast an "mmm" sound with a "t" sound. Percussion instruments are by their very nature demanding. A classic example of this would be the hit of a snare which goes from stationary to full on in an instant - it jumps. Audio components that do not jump fast will compromise definition of the audio. Therefore good transient response is contingent upon fast acceleration and deceleration. Devices need to stop as fast as they accelerate to avoid ringing on and interfering with the next part the audio signal.

Figure 1: Sound Envelope Regions

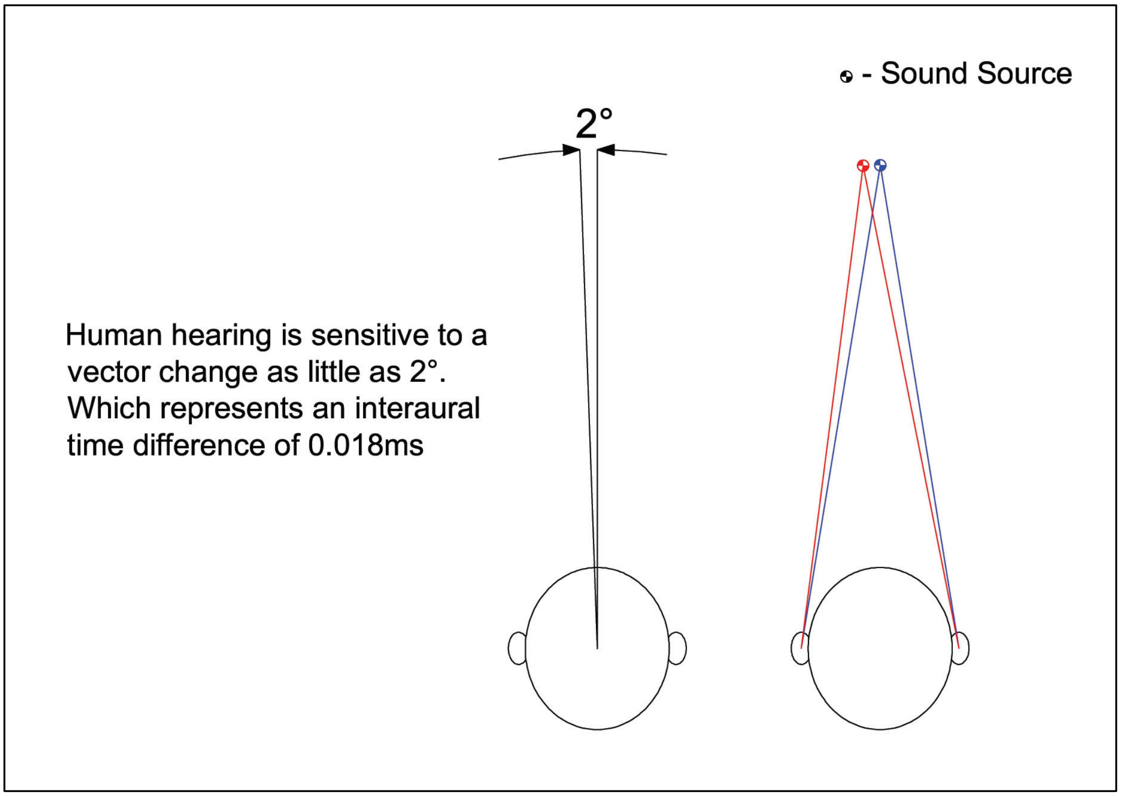


2.2 Resolution of Human Auditory System

Good transient response not only delivers definition, accuracy and excitement but also contains a wealth of information about the sound source's distance and direction. To retrieve this information the human auditory system is capable of discerning and processing an amazingly fine grained level of time resolution.

This is demonstrated by work done (I am informed by Dr Peter Lennox of Derby university) in determining the minimum detectable angle of vector shift. In other words how much a sound source has to move before human hearing detects a change in location, using the difference of arrival times between the two ears. This turns out to be as little as 1 or 2 degrees in the horizontal plane which represents a timing arrival difference at the ears of 13 to 18 microseconds which is 13 to 18 thousandths of one thousandth of a second. Contrast this with a movie frame rate of 28 fps or 36 milliseconds which is over two thousand times slower. I make this point to illustrate the highly refined nature of human hearing which begs the question as to why it is so often short changed in favour of the less refined visual sense. Why are we so ready to accept poor sound when we would never accept grainy or smeared visuals? To get back on track, this amazing sensitivity to time information is naturally focused on the boundary condition of the leading edge or attack which is where all the action is including information about distance.

Figure 2: Vector Change Picture



High frequencies naturally have a fast rise time which correlates to transient information. Sound arriving from a distance suffers from natural high frequency loss and therefore impaired transient which is perceived and understood by the aural circuitry to be an indication of remoteness. Furthermore transient information can also be compromised by smearing due to secondary arrivals with slight delays due to their longer path lengths reflecting from ground, trees, buildings etc. This is more likely with greater distance so it is also taken as a cue for remoteness. Whatever the cause, compromised transient information has the perceived effect of placing the sound source at a distance and removing its dimensional qualities which in turn impairs the feeling of involvement. Therefore the ability of a chain of audio equipment to follow a given audio signal in all its detail is paramount if we are to achieve reality. This ability usually falls over when we get to loudspeakers.

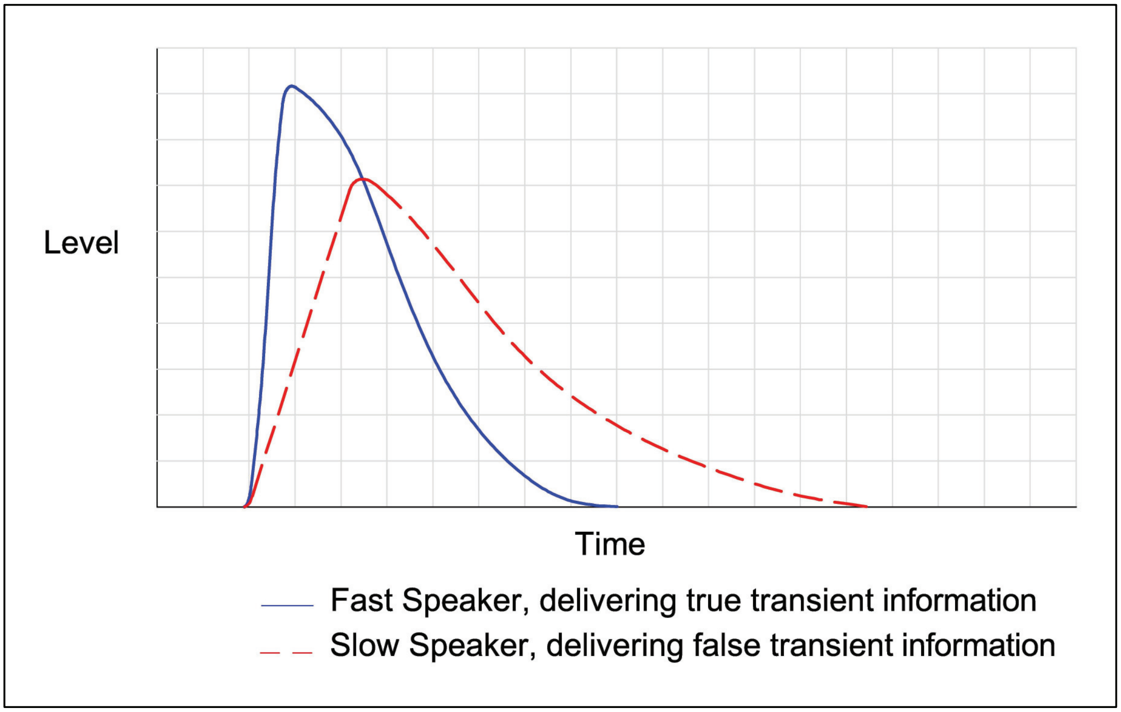
2.3 Resolution of Sound Systems

Loudspeakers being electro-mechanical devices have a tougher time than electronics in keeping up with the dynamics of wave forms. Loudspeakers are the traditional bottleneck which is the reason why my partners and I have spent 30 years researching and developing fast, accurate speakers.

Slow speakers compromise the perceived audio by being unable to correctly respond to the transient demands of the source. The rise time can be compromised to the point where the transient can be over before the loudspeaker ever reaches the intended peak level.

In this way the dynamic range is also compromised as well as resulting in the previously discussed problems.

Figure 3: Effect of Slow Speaker on Transient

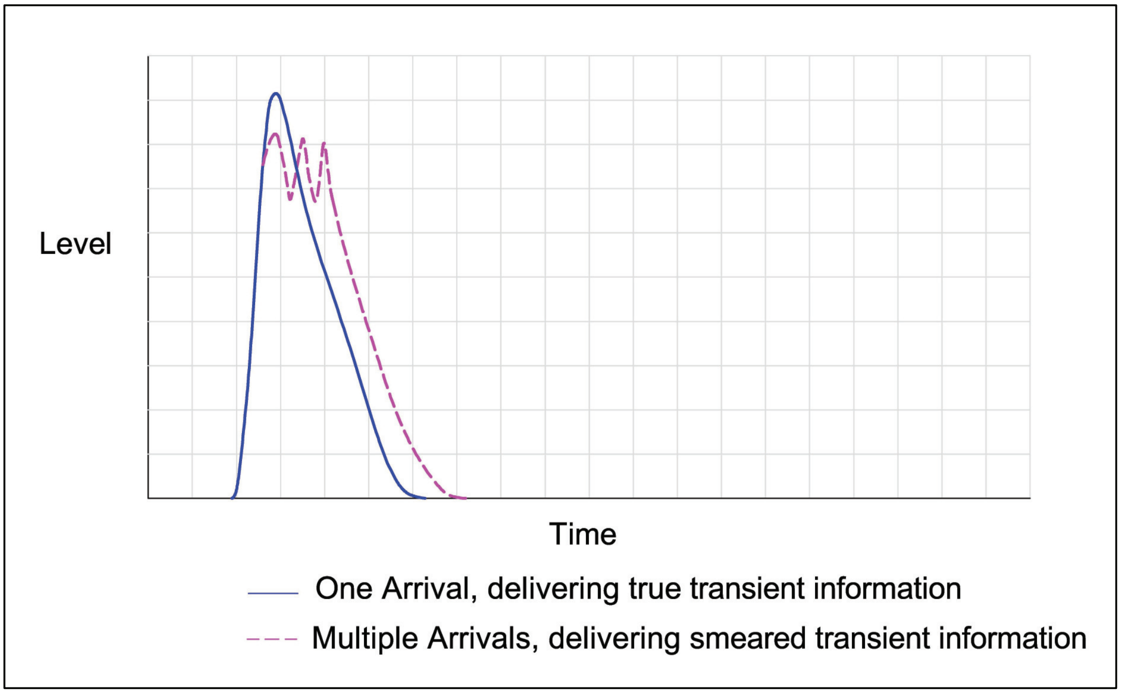


For good transient response it is essential for the loudspeakers to be individually fast. However, when combining speakers to make a larger system if their configuration introduces multiple arrivals resulting in time smearing this also compromises transient.

System configurations that produce multiple arrivals will have the effect of smearing the attack and even affecting the perceived frequency response. I have found this to be very evident in the case of line arrays, particularly in the voice range, where the obfuscation of defining consonants renders entire vocal performances unintelligible. In other words one cannot understand a single word being sung.

Although transient smearing is a different mechanism to slow rise time, the effect nevertheless compromises the perception of definition.

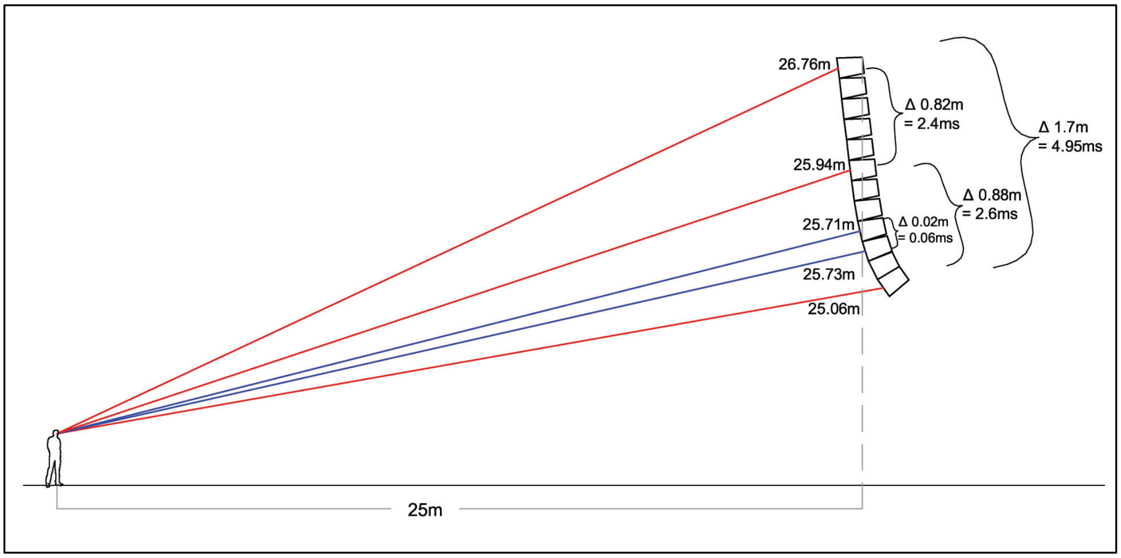
Figure 4: Effect of Multiple Arrivals on Transient



In the last 15 years, line arrays have become the paradigm of professional loudspeaker systems. By and large they are all inspired by Christian Heil's development of the classic line array arrangement and mostly employ multiple direct radiators and wide dispersion HF devices which, ideally, collectively reinforce each other by constructive interference although of course they also suffer from destructive interference. Direct radiators are inherently of low efficiency because of the substantial impedance miss match between the moving diaphragm and the air. Although the low efficiency is improved by the mutual coupling it is at the expense of introducing multiple arrivals.

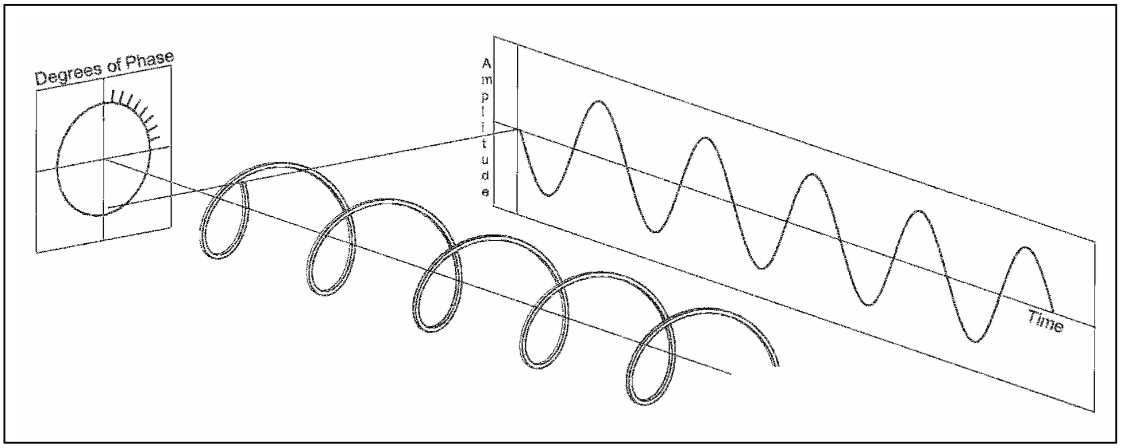
As can be seen from the diagram (*Figure 5*) the difference between the top and bottom boxes is as much as 5 milliseconds but more disturbingly the difference between adjacent boxes of 60 microseconds is at least three times greater than the lower limit of human time perception. To make matters even worse any frequency whose half wavelength is smaller than the distance between the centres of the drivers will not couple at source and so will be prone to off axis comb filtering which is evidenced by the large amount of phasing and instability that occurs when line arrays are working in even slightly windy conditions. I find this very ironic as in the early days of line array they were held to be superior to point source because they were less prone to phasing when one walked across them. Now you can enjoy the same effect without even moving. Badly smeared attack can never achieve a sound stage with depth, dimension and good vocal intelligibility. To add insult to injury, to my knowledge most if not all line array systems employ corrective EQ. Some of them as much as 18dB boost or cut at half a dozen or so different frequency points introducing yet more time or phase issues.

Figure 5: Distance and Time Difference between various boxes in a Line Array



Phase shift is measured in degrees of arc around a circle which looks flat but mathematically the situation is in fact three dimensional, with the circle of phase angle being an end on view of a spiral or helix, the sideways view giving us the familiar sine wave as shown in the diagram (*Figure 6*). This means that phase angle also represents a very small increment of time. As EQ introduces phase or time shift, corrective EQ is not only an admission of defeat but adds further confusion to the all important leading edges of sound. In fact line array resolution is so bad that it has resulted in a generation of audio engineers using very blunt tools which has provided the perfect smokescreen for all sorts of dreadful digital equipment to gain a very questionable hold on FOH in the touring industry.

Figure 6: Phase Shift



It's not that digital equipment is intrinsically bad, it's that it is actually very difficult and expensive to get the last few percent that really counts, so with line arrays masking all the deficiencies very few engineers have demanded that digital manufactures do better. Digital signal processing has been with us for at least 25 years but it has risen to prominence hand in hand with line arrays in the last 15. Early digital even allowed valves a come back to "warm up" the sound because early digital was so wrong that even valve characteristics were preferable to the lifeless distorted digital result. I do not exactly know what is wrong with digital but I have heard only a few devices which sounded real to me. The manufacturers of these devices tell me that some of the difficulties with the design and manufacture of digital equipment are:-

- Good quality analogue to digital converters are expensive.
- Within a given piece of digital equipment there is a lot of high frequency interference which is hard to keep away from sensitive areas.
- The mathematics used has to be very accurate or the audio is compromised.
- There is no such thing as a jitter free clock. Jitter results in distortion.
- All digital equipment has processing time known as latency which is fair enough but different input signals have different treatment which gives varying latencies resulting in yet more time smearing.

Digital only achieves full resolution at maximum input or zero dB so by the time you have allowed for transient peaks your average mix level has lost over half its resolution and it gets worse because with 2nd, 3rd etc. harmonics being 35 to 45 dB below operating level there are not enough digital bits left to reproduce them at all. This is part of the reason that digital lacks richness and sounds dreary and lifeless.

There are plenty of ways to get it wrong and as such digital is still work in progress. Thank Heaven for the hardcore engineers that will still only work with high quality analogue mixers such as XL4s or Heritage because they want the audience and themselves to enjoy what they do. With a few notable exceptions the current crop of digital mixers do not achieve those levels of enjoyment in terms of the reality of sound. Some in particular sound incredibly wrong. Line arrays do not have the resolution to allow engineers to discern the difference. It is difficult to believe that the audio community would have allowed these developments merely on the grounds of convenience. There are, however, political and psychological advantages in that indistinct and two dimensional sound could be likened to soft focus photography for smoothing out the blemishes. This reduces the dynamic range between the good and bad engineers preventing good ones from excelling and allowing bad ones to get away with it. This combination of line array and inadequate digital is what has brought professional audio to its current sorry state

People adapt to what they live with which is why generations growing up exclusively with MP3 audio have no idea what they are missing, and their auditory perception is being badly programmed. Anybody who is auditioning their rig on MP3s is either working with substandard equipment or they can't hear the difference in which case they really should be doing something else! By the same token the sound of line arrays has become accepted as 'normal' so that few even notice that it is actually substandard. For example I believe that we at Funktion One strive to produce a very clean, full bodied, accurate and fast mid range. It is very different from what is generally occurring in the touring world. Some engineers react by applying EQ to the lower half of the mid range just leaving behind the upper mid they are so used to. Others find it too unfamiliar and stark despite good transient response being crisp and revealing, they prefer to blunt the sound with compression.

An additional factor is that sound rental companies have a major influence on concert audio quality. In the 35 years that I have been involved it has always been very hard to achieve more than financial equilibrium owning and running a rental company. In the early 70s when the rental industry began in earnest the people involved were either audio enthusiasts or into the apparent “glamour” of touring or both. The readiness of people to do it for the emotional return resulted in pricing well below a fair rate and it has pretty well remained that way to the present day. This combined with traditional rental company price competition under values audio. This is nice for production managers and promoters, but it has bred a climate of unseemly behaviour and hidden alliances with self referential agendas. Apparently, the quest for excellence in audio has been buried in the bun fight years ago and the quality of experience for the audience is forgotten. Now it is just plant hire.

3 CONCLUSION

To sum up I hope that I have communicated how important transient response actually is and how the neglect of this important parameter has led to a substandard professional audio industry.

I understand the *spirit* of rock and roll to be making for the far horizons, while the *actuality* is a lowest-common-denominator, bland ‘mush’. The sad outcome of all this is that performance audio has not properly progressed in decades and the audiences (the people who *actually* pay for the music industry) are being short-changed in audio quality on top of ever more draconian sound level limits and miserable weather. There are many reasons, but no valid excuse for mediocrity and so, for those professionals who actually care and want to take pride in the quality of their workmanship, I ask:

> **“Do we want reality or are we more comfortable with boring, undemanding two dimensional soft focus?”**