Room Acoustics for Listening

James Heddle

James Heddle Acoustical Consultants, Brisbane, Australia

ABSTRACT
A number of recent projects involving audiences and critical listening have been designed based on a 'perceptually efficient' concept. An outline of this evidence-based design strategy will be given together with a discussion of outcomes and insights from the completed projects.

1. INTRODUCTION
Our understanding of auditory processes and the functioning of the brain is far from complete, however, recent attempts at biomimical modelling of the neo-cortex of the brain are providing some useful insights (Byrne, 2015). Coupling the incoming stream of information (current observations) with learnt patterns (auditory experience) results in cognition, the brain generates the most plausible explanation of the inputs and updates the model as necessary based on new information (e.g. looking around a room, head turning). In this conception, our experience of an auditory event is a reconstruction of reality based on pattern recognition.

What we perceive is therefore reliant on the observations from which the model has been constructed (inputs from the whole range of senses, our cumulative experience of the world) with capabilities for aspects having a survival advantage (food finding, predator avoidance, etc.) likely to be given emphasis.

Consistent with this premise are the precedence effect and other findings, amongst many, such as the McGurk Effect, speech intelligibility in a reverberant room improving with consistent listening exposure to the room, relative to situations with inconsistent room exposure (Zahorik & Brandewie, 2016), and speech-in-noise perception improvement after 2 years of group music training (Slater et al., 2015).

Recently it has been demonstrated that our ability to simultaneous judge the frequency and timing of a sound exceeds the limits of Fourier time-frequency analysis (Oppenheim and Magnasco, 2013), particularly by conductors and composers and particularly in relation to temporal acuity, which casts doubt on models of hearing based on linear operators, such as filter banks and spectral analysis. Additionally, important auditory capabilities, such as localization, rely on the differences between the signals arriving at the two ears. Acoustic design targets based on data and parameters derived from single microphone measurements are therefore likely to have inadequacies and may be misleading. Overall this implies our current understanding and models of perceptual processes, including those involved with listening, have significant room for improvement and are likely to more fully develop over time. We should, therefore, be cautious of design based solely on room acoustic parameters given in current international standards (Lokki, 2013).

The following outlines some elements of a design scheme used on recent projects intended to achieve an improved accord between the physical properties, conditions and treatments in the room and desired perceptual qualities.

2. LISTENING CONDITIONS – GENERAL REQUIREMENTS
A well recognized primary requirement of good conditions for listening is an absence of competing sound, that is, it is desirable that the sound signal of interest has minimal levels of interference and competition from other sound. The larger and more listening-focused a space is required to be, the more critical this becomes. This means that the levels of sound generated by mechanical and electrical services, hydraulics and due to other internal and external noise sources and noise events ideally need to be suppressed to below the threshold of audibility, no noise event distractions or masking of the source sound. The difference this makes to hearing fine detail and subtleties in performances and in recordings is marked but can be easily undermined by lack of awareness and indifference. For less than critical acoustical spaces, and given real world budgetary limits, the ambient noise goal can be relaxed, without significant detriment, to those found in Standards provided the noise spectrum is constant and neutral.

Another desired general goal, more difficult to give a precise design target for, is for the space to be non-fatiguing. In short, this means the minimization of cognitive load, not overly acoustically dead or reverberant.

3. RESONANCES AND ROOM LOW FREQUENCY RESPONSE

Resonances in a listening space have the potential to degrade the fidelity of the sound by spectral colouration, distorting the timbre of the sound, and temporally, by degrading the transient response. These aspects can also apply to sound sources in the listening space, such as loudspeakers. The audibility of resonances has been found to relate primarily to three factors: the resonant frequency, the Q-factor or decay time of the resonance and the level of resonance relative to that of the signal (Mateljan et al., 2007; Toole & Olive, 1988).

3.1 Room Resonances - Audibility

Room resonances occur at all frequencies, increasing in severity as the room surfaces become more reflective and in density as frequency increases. The audibility of these resonances tends to be greatest at lower frequency. A transition frequency, below which the room is increasingly dominated by separate modes, was proposed by Manfred Schroeder at the point of a three-fold overlap of adjacent modes, given by Fs =$2000(T_{60}/V)^{0.5}$ Hz.

An investigation of the perceived quality of low frequency audio reproduction in rooms (Wankling et al., 2009) found that increased preference was associated with a decrease in perceived room resonance and that increased excitation of modes, such as due to positioning a monopole source in a room corner, was detrimental to preference.

Particularly for smaller rooms, research by Stephenson (2012) indicated that the frequency below which modal effects are audible tends to be higher than that indicated by the Schroeder frequency. This was, in part, found to be due to interference between modes. As the interference between modes is dependent on both their spacing and their relative phase, neither the density nor the spacing of modes by themselves are determinant parameters for room mode inaudibility. Figure 1 shows a Schroeder frequency calculation in comparison to the subjective mode audibility limit found in his tests. Based on the somewhat limited test data, a provisional relation for the subjective transition frequency can be given by F_{ms} = 92577/(V-62.72) Hz.

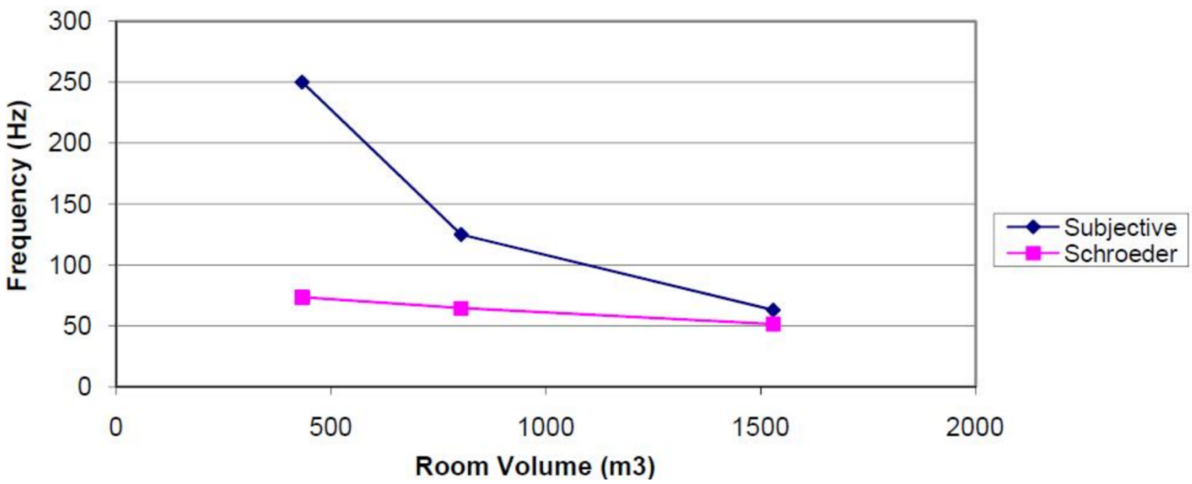

Figure 1: Subjective Transition Frequency for Modal Audibility vs. Schroeder Frequency (Stephenson, 2012)

A review of room mode control methods (Fazenda et al., 2012) found that the perceived improvement in reproduction sound quality was strongly correlated to those control methods causing a reduction of low frequency modal decay times. In a later study (Fazenda, 2015), both the absolute thresholds of the audibility of modes and audibility thresholds based on music signals were determined.

Figure 2 shows the threshold of audibility for room modes found in the study for both music and technical test signals (absolute). The absolute threshold mode decay times and the modal Qm (maximum) associated with the threshold decay time may be expressed approximately as,

$$Rtm(absolute) \approx 0.11 + \frac{617}{fm^2}\ seconds; \quad Qm(absolute) \approx \frac{fm}{20} + \frac{280.5}{fm} \qquad (1)$$

$$Rtm(music) \approx 0.15 + \frac{755}{fm^2}\ seconds; \quad Qm(music) \approx \frac{fm}{14.67} + \frac{343.2}{fm} \qquad (2)$$

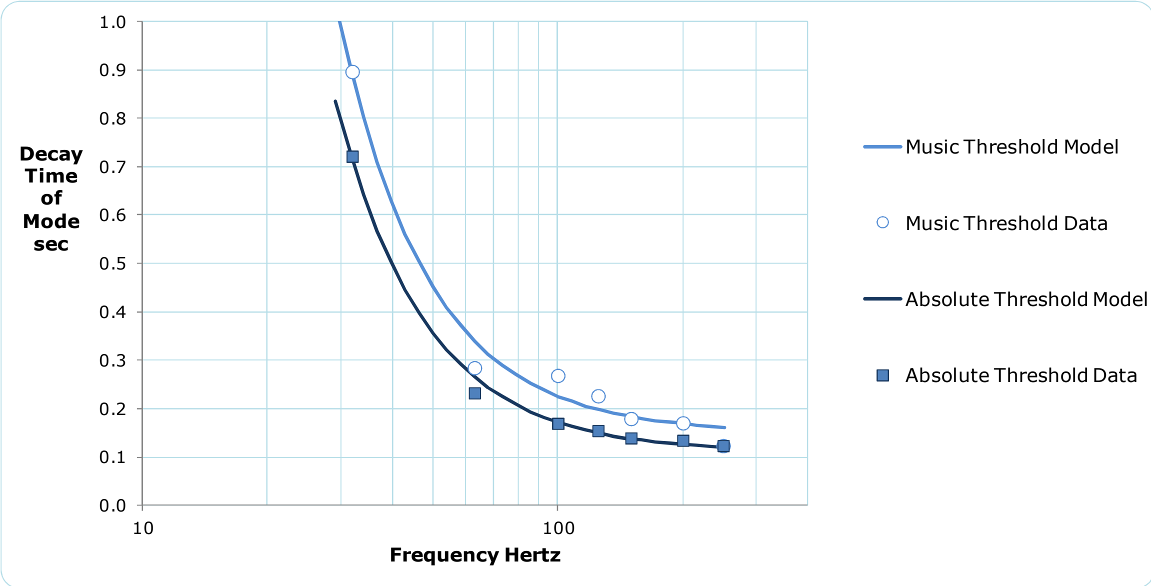

Figure 2: Threshold of Audibility Decay Times for Room Modes (Fazenda, 2012)

3.2 Room Treatment - Low Frequency

It should be noted that, in contrast to some other techniques, reducing the modal decay times via absorption improves the room globally rather than at specific locations and reduces spatial variance. Together with the findings on improved perceived sound quality this means that reducing modal decay to threshold levels is a perceptually efficient strategy.

The 60 dB decay time (reverberation time) of a given mode, R_{tm}, is determined from the effective travel path distance, metres, between the reflective surfaces involved in the mode, L_m, and the effective average absorption per reflection, α_m, of the involved surfaces at the mode frequency, fm.

$$Rtm = \frac{Lm\ln(10^{-6})}{c\ln(1-\alpha m)}\ seconds$$, where c is the speed of sound in air, metres/second. (3)

The Quality-factor of the mode, Q_m, and the decay time of the mode, R_{tm}, are related by $Rtm = 2.2\frac{Qm}{fm}$. Using this and (3) we may derive expressions for the average absorption coefficient of the surfaces involved with the mode required for a target mode decay time.

$$\alpha m = 1 - \exp\left(\frac{\ln(10^{-6})Lm}{c\ Rtm}\right) = 1 - \exp\left(\frac{\ln(10^{-6})Lm\ fm}{2.2\ c\ Qm}\right) \quad (4)$$

Various options exist for implementing bass absorption (Fuchs, 2013; Fuchs & Lamprecht, 2013; Everest, 1994) and the most suitable ones depend both on the constraints, the absorber performance and the design targets. It should also be noted that significant bass absorption may be present in the room construction, particularly in lighter weight walls and panelling (Bradley, 1997). The main design consideration is to allow enough physical space for the absorption in order to achieve the required decay rate. In general, it is quite difficult to retrofit additional bass absorption into existing rooms with problematic modes so this is a crucial design consideration.

In addition to the specific task of the attenuation of room mode resonances, project experience with music practice rooms, consulting, meeting and video conferencing rooms, lecture theatres and auditoria has confirmed that greater degrees of bass absorption in these spaces than is typical results in a significant improvement in functionality and user friendliness with subjective clarity and speech intelligibility notably improved. This seems to be related to the reduction in upward frequency masking associated with the reduction in decay time (Ehmer, 1959). Figure 3 shows the level required for a frequency component to be audible in the presence of a 250Hz tone at the stated sound pressure levels and how this causes masking effects well above the tone frequency.

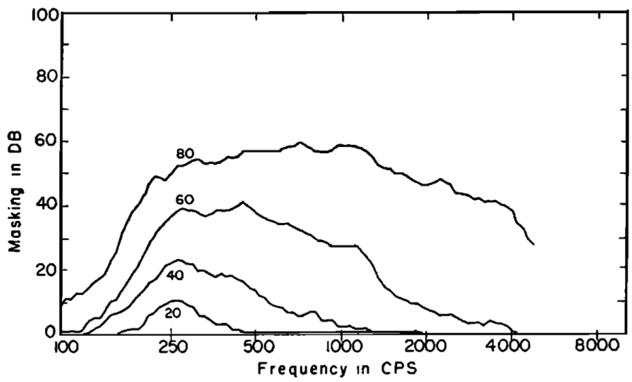

Figure 3: Average Masking Patterns due to a 250 Hz Pure Tone (Ehmer, 1959)

Consistent with these observations is the recommendations given in a review of acoustic treatment options for classrooms, the Essex Study (Canning & James, 2012), to adopt the criteria of British Association of Teachers of the Deaf limiting classroom reverberation times in the 125 Hertz and higher octave bands to 0.4 seconds, having found this gave the best listening and speaking conditions.

For the situation of acoustic performances in larger rooms, there is a trade-off in the amount of bass absorption implemented between clarity and the loss of strength and loss of late reverberant bass sound field level.

4. EARLY REFLECTIONS

Reflections introduce spectral, temporal, and spatial modifications to the direct sound. Reflections can be considered as early up to a time instant after the arrival of the direct sound when the reverberation tail cannot be subjectively distinguished from that at any other position, or listener orientation, in the room, generally taken as a fixed time of 50 msec to 80 msec. The time after the arrival of the direct sound that this occurs has been found to vary but appears to be in the order of the time taken to propagate at least 5 mean free path lengths (Lindau et al., 2012). Early reflections are generally stronger than later arriving reflections and contribute the most towards the acoustic signature of a space (Haapaniemi & Lokki, 2014).

In the case of rooms for audio monitoring, to avoid introducing biases into the localization of phantom images, both the sound sources and the room should be symmetrical about the central front/back axis.

4.1 Spectral Colouration

Spectral colouration can be thought of as perceived timbral distortion of the source signal and it is a primary factor influencing judgements of sound quality and preference (Clark, 1983; Rumsey, 2005). This distortion can be created in the components of an audio signal path and as a consequence of room reflections and diffraction. The interaction between the direct sound and secondary sources, such as room reflections, results in comb filtering at the ears. This comb filtering is different at each ear except in the case of median plane reflections where it may be highly correlated.

Gunther Theile (1980) found evidence of an inverse relationship between the ability to localize a source, detect its location, and the detection of colouration for binaural listening. He found that the colouration that was evident when listening with one ear, and which changed in character with movement, ‘disappeared’ with binaural listening when a source or phantom source could be localized. That is, it appears that monaural-like first order reflections in the median plane, or in the same direction as the source, are more prone to cause colouration of the sound and measurable spectral distortion, such as due to a reflection, is not necessarily perceived as colouration, particularly in the case where the source or phantom image can be localized (Clark, 1983; Toole, 2008; Theile, 1980).

Figure 4 shows the spectral colouration at the receiver for fixed source and receiver heights and various ceiling heights. This shows that as the ceiling height and the travel path via the ceiling increases, the colouration reduces with the maximum and minimum excursions reducing and moving to lower frequency. The shaded section represents the zone with subjectively inaudible colouration, having a spectral deviation not greater than 1 dB.

At the same relative distances this behaviour would also apply to other room surfaces.

The design strategy adopted gives emphasis to low colouration. The approach taken is to attenuate early median plane reflections sufficiently, via absorption, acoustic shielding or redirection, for there to be minimal changes relative to the source spectrum (although it is typically not possible to achieve this for the floor reflections).

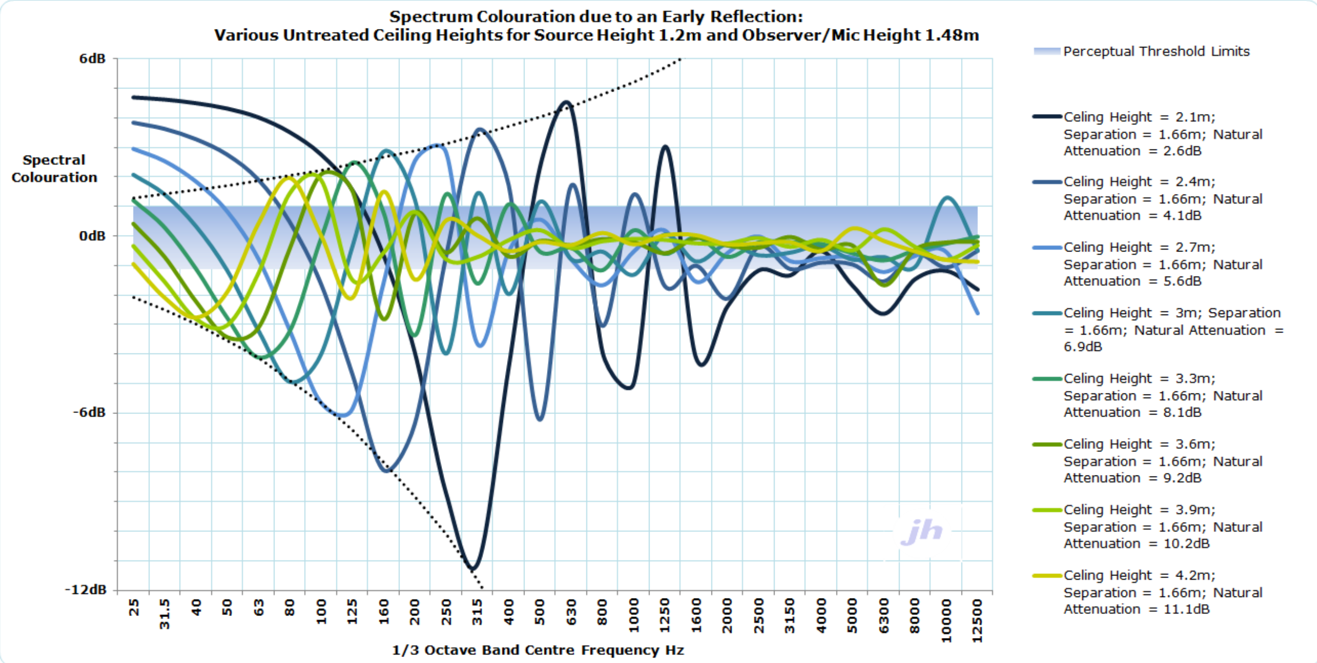

Figure 4: Spectrum Colouration due to a Reflection.

Figure 5 shows an implementation of a scheme used on a number of projects to reduce the first ceiling reflection. This involved determining the additional attenuation required to bring the colouration down to an imperceptible level, determining the critical treatment zone on the ceiling for the loudspeaker and listener positions and implementing the attenuation via baffle shielding. Baffle shielding rather than absorption was used to more directly target just the early ceiling reflection path. Further details outlining the treatment zone determination using Fresnel zones are given in a companion paper (Heddle, 2016).

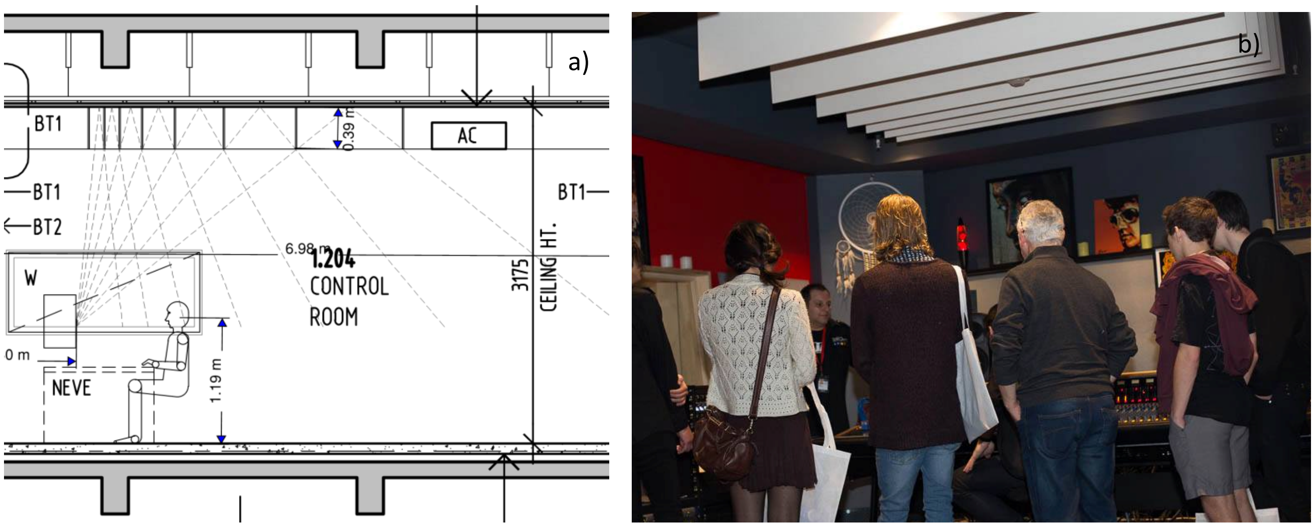

Figure 5: Control Room Ceiling First Reflection Treatment. a) Room section showing loudspeaker source and ceiling baffling. b) View of ceiling baffles in control room.

4.2 Lateral Reflections

Due to the location of the ears in the side of the head there are useful gains for sound in the upper frequencies that arrive laterally. Early lateral sound can enhance the apparent loudness of the direct sound and, particularly if the temporal envelope is not distorted by the reflection surface, provide a more open, clear, and enveloping sound (Lokki et al., 2011).

An early lateral reflection should ideally be a mimic of the direct sound with the exceptions of overall level, delay and polarity. It seems the auditory process is best able to focus on the direct sound in the presence of reflections if the reflections are recognized as being due to the direct sound without ambiguity or uncertainty.

This means that diffusive surfaces that scramble phase relationships, frequency variable absorption and sound systems that are not constant directivity are likely to degrade this effect and thereby degrade the ability to separate out the direct sound as the focus of attention, known as audio stream segregation. It would also seem that this aspect has some bearing on the attribute 'proximity', the most significant aspect associated with preference found in a study of concert hall acoustics (Kuusinen et al., 2014). It is notable that this attribute has no correspondence with standard acoustic test parameters.

For halls and lecture theatres, sidewall treatment and configuration needs careful consideration to encourage beneficial early lateral reflections to seating locations. In the case of control rooms for audio monitoring, lateral reflections are not likely to cause detrimental effects if the loudspeakers are constant directivity (not usually the case), if the reflections are temporal envelope and spectrum preserving and if the reflections are equivalent on the left and right. Consideration should be given to attenuate the early lateral reflections, using frequency independent (broadband) sound absorption or redirection, where these conditions are not likely to be met.

4.3 Fusion and Localization Dominance

When two similar sounds follow closely in time the auditory system fuses these into a single object located between the two sources. An example of this is the centre phantom image of a source fed equally to two loudspeakers. As the delay difference increases between the two sources, the auditory system starts to locate the source direction as being that of the first arriving sound and the strength of the more delayed sound is suppressed (in the presence of many delayed sounds or reflections, a kind of de-reverberation occurs (Hartmann, 1997)) .

Figure 6a shows the percentage of trials for which the sound was not fused into a single source versus the delay (milliseconds) of the second sound after the first, 5 test subjects. It can be seen that the delay required for complete separation into separate sounds (100%) varies between subjects but is largely achieved by 6 milliseconds (dotted vertical line)(Litovsky & Shinn-Cunningham, 2001).

When this condition is met, which rules out flush mounted monitors, particularly in the case of audio monitoring with constant directivity loudspeaker sources, it is possible for the auditory system to, in effect, ignore the room and focus on the direct sound from the loudspeakers (Linkwitz, 2009; personal observations).

Figure 6b shows a horizontal section of a room designed to just fit a 6 msec time window for the required stereo listening triangle. It can be seen that achieving this design requirement with minimum room dimensions, for a typical listening triangle, results in a room wider than it is deep. There is also evidence that better low frequency sound source localization results from maximizing the time difference between the direct sound and the first reflections (Hill & Hawksford, 2013).

4.4 Direct to Reverberant Level

In parallel to the low ambient noise/good signal to noise requirement, larger spaces require an adequate ratio of direct to reverberant sound field level (D/R) at listener positions in the audience. Griesinger (2009) reported that for two separated sources with a sound signal alternating between them, localization was lost when the direct to reverberant level fell to around -13dB, the sound was perceived to be in the middle between the two source locations. The D/R via its effect on localizability thus also has an influence on perceived colouration. In addition, above this threshold, a decrease in subjective distance was reported.

For high sound quality, acceptable localization and engagement, it is desirable that the D/R at a listening location does not fall below a minimum around -10dB (Griesinger, 2007), and this requirement should take precedence over reverberation time goals in most cases. Exceptions to this would be venues dedicated to music suited to, and originating out of, reverberant environments, such as organ, Gregorian chant and slow tempo choral music.

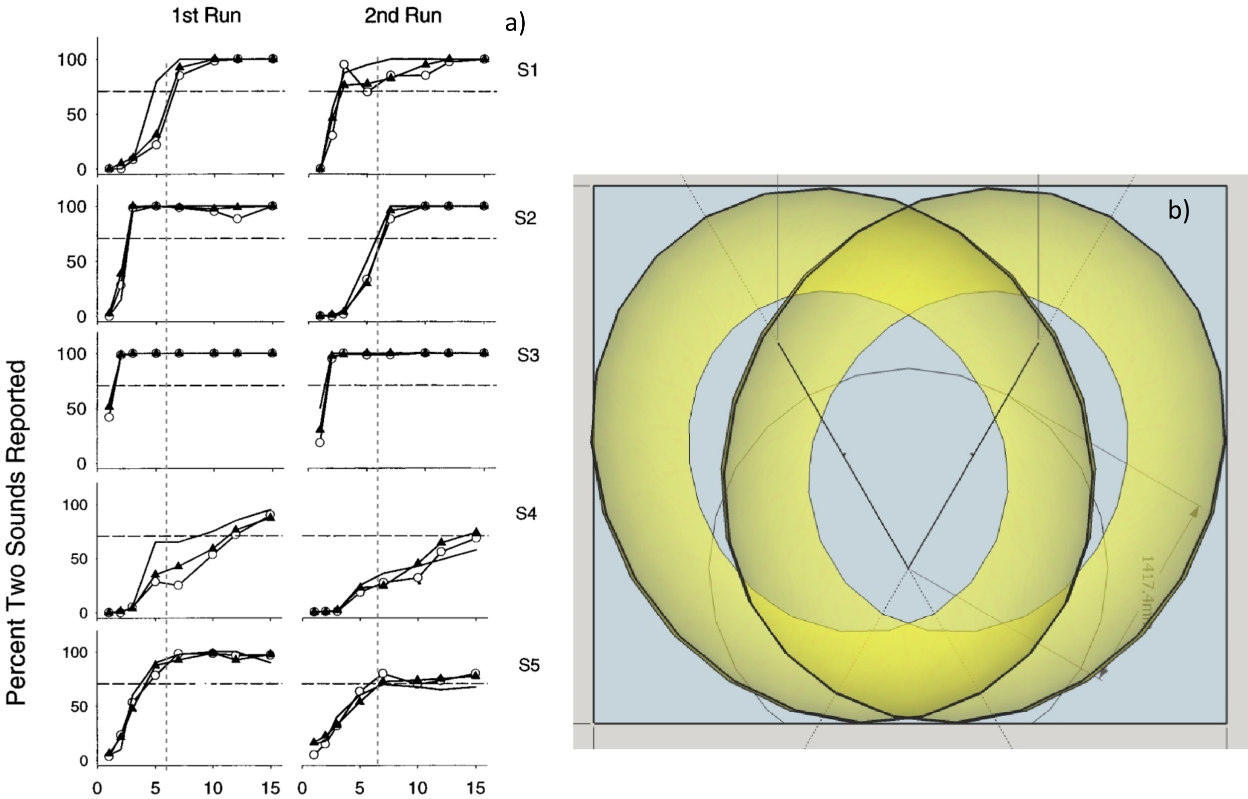

Figure 6: a) percentage of trials on which the subjects (S1-S5) reported hearing two sound images as a function of the lead–lag delay. b) Horizontal section of a room fitting 6 msec delay path spheroids focused on Left and Right Loudspeaker Positions to the Central Listener Location in a standard stereo listening triangle setup.

5. CASE STUDY

Figure 7 shows a recent 76 seat theatre design, used for lectures and audio demonstrations, incorporating the combination of acoustic elements discussed. Bass absorption has been installed on both sides of the theatre at ceiling height, running the length of the theatre and also across the front at floor level. Both side walls have smooth reflecting elements angled to provide lateral reflections at listeners in the range of 60 to 80 degrees relative to the stage source position. Ceiling baffles have been installed to block the first (median plane) ceiling reflection and, in this case, provide reverberation control. The theatre rear wall, a median plane element, is also absorptive (broadband). The measured room equivalent continuous ambient noise level was 35dBA and neutral in character.

User feedback is that the room is very clear with excellent speech intelligibility and ease of listening. It was also found that audio-visual program material can be run at high sound pressure levels without the room overloading, maintaining very articulate bass.

6. SUMMARY

The overall intent of the combination of measures discussed, designated as 'perceptually efficient' acoustic design, is to allow the focus of attention to be drawn to the source, to subjectively minimize the adverse effects of the room and external activity on the source signal and to improve engagement by lowering cognitive load and fatigue. Project experience and feedback on halls, lecture theatres, educational facilities and recording studio spaces designed using this approach has been very positive. The main features of this design approach include:

- Low frequency room modes and bass absorption are given higher priority and significantly more bass absorption than standard practice is used in smaller scale rooms.
- Early reflections that may otherwise have arrived at a listener in the median plane are discouraged or redirected to lateral, the acceptability criteria is based on their spectral distortion effect.

- Early lateral reflections are encouraged but with specific provisos.
- Early reflection absorption treatment is as frequency independent as can be feasibly achieved and is localized to the most effective treatment zones on the reflective surface (determined via Fresnel zones).
- A minimum delay between the direct sound and the arrival of the first reflection is required. This permits the sources, including audio sources, to be perceptually decoupled from the room. Because of this requirement flush mounting of loudspeakers for critical listening is not used.
- Adequate signal to noise and direct to reverberant levels are mandated, this taking precedence over reverberation time as a design target.
- These features to be viewed as a system to be implemented in combination.

Figure 7: Perceptually Efficient design example, Theatre. a) Section showing sidewall lateral reflectors, ceiling reflection control baffles, bass absorption a stage floor and at upper sides. b) View of Theatre.

REFERENCES

Bradley, J 1997, '*Sound absorption of gypsum board cavity walls'*, J. Audio Eng. Soc., Vol.45, No.4

Byrne, F 2015, '*Encoding Reality: Prediction-Assisted Cortical Learning Algorithm in Hierarchical Temporal Memory*', https://arxiv.org/abs/1509.08255. Also see https://www.youtube.com/watch?v=rrBhHDzmgUA

Canning, D & James, A 2012, 'The Essex Study: Optimised Classroom Acoustics for All', The Association of Noise Consultants, UK

Clark, D 1983, 'Measuring Audible Effects of Time Delays in Listening Rooms', 74th Audio Engineering Society Convention, New York

'*Classroom Acoustics – recommended standards*', British Association of Teachers of the Deaf

Ehmer, R 1959, 'Masking Patterns of Tones', J. Acoust. Soc. Am. 31

Everest, FA 1994, 'The master handbook of acoustics', McGraw-Hill, New York

Fazenda, B 2004, *'Perception of Room Modes in Critical Listening Spaces'*, Doctoral Thesis, University of Salford.

Fazenda, B & Wankling, M 2008, *'Optimal Modal Spacing and Density for Critical Listening'*, 125th Audio Engineering Society Convention, San Francisco

Fazenda, B et al. 2012, *'Subjective Preference of Modal Control Methods in Listening Rooms'*, J. Audio Eng. Soc., Vol.60, No. 5

Fazenda, B M; Stephenson, M & Goldberg, A 2015, *'Perceptual thresholds for the effects of room modes as a function of modal decay'*, J. Acoust. Soc. Am. 137 (3)

Fuchs, H & Lamprecht, J 2013, 'Covered broadband absorbers improving functional acoustics in communication rooms', Applied Acoustics 74 (1)

Fuchs, H 2013, 'Applied Acoustics: Concepts, Absorbers, and Silencers for Acoustical Comfort and Noise Control', Springer-Verlag

Griesinger, D 2007, *'Perception of Concert Hall Acoustics in seats where the reflected energy is stronger than the direct energy'*, AES 122nd Convention, Vienna, Austria

Griesinger, D 2009, 'The importance of the direct to reverberant ratio in the perception of distance, localization, clarity, and envelopment, Part one', J. Acoust. Soc. Am. 125

Haapaniemi, A & Lokki,T 2014, 'Identifying Concert Halls from Source Presence vs Room Presence', J. Acoust. Soc. Am. 135 (6)

Hartmann, W 1997, *'Listening in a Room and the Precedence Effect'*, Chapter 10, Binaural and Spatial Hearing in Real and Virtual Environments, Psychology Press

Heddle, J 2016, '*Fresnel Zones and Spheroids for Room Acoustics'*, Proceedings of ACOUSTICS 2016, Brisbane

Hill, AJ & Hawksford, MOJ 2013, 'Low-Frequency Sound Source Localization as a Function of Closed Acoustic Spaces', Proceedings of the Institute of Acoustics, Vol. 35. Pt 2

Kuusinen, A et al. 2014, 'Relationships between preference ratings, sensory profiles, and acoustical measurements in concert halls', J. Acoust. Soc. Am. 135 (1)

Lindau, A; Kosanke, L & Weinzierl, S 2012, '*Perceptual evaluation of model-and signal-based predictors of the mixing time in binaural room impulse responses'*, J. Audio Eng. Soc., Vol.60, No.11

Linkwitz, S 2009, 'The challenge to find the optimum radiation pattern and placement of stereo loudspeakers in a room for the creation of phantom sources and simultaneous masking of real sources', 127th Audio Engineering Society Convention, New York

Lokki, T 2013,'Throw away that standard and listen: your two ears work better', International Symposium on Room Acoustics

Lokki, T et al. 2011,'Engaging concert hall acoustics is made up of temporal envelope preserving reflections', , J. Acoust. Soc. Am. 129 (6)

Litovsky, RY & Shinn-Cunningham, BG 2001, 'Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression', J. Acoust. Soc. Am. 109 (1)

Mateljan, I; Weber, H & Doric, A 2007,*'Detection of Audible Resonances'*, 3rd Congress of the Alps Adria Acoustics Association, Graz, Austria

Oppenheim, JN & Magnasco, MO 2013, 'Human Time-Frequency Acuity Beats the Fourier Uncertainty Principle', Phys. Rev. Lett. 110

Rumsey, F et al. 2005, *'On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality'*, J. Acoust. Soc. Am. 118 (2)

Slater, J et al. 2015, '*Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program'*, *Behavioural Brain Research*, vol. 291, pp. 244-252.

Stephenson, M 2012, 'Assessing the Quality of Low Frequency Audio Reproduction in Critical Listening Spaces', PhD Thesis, University of Salford

Theile, G 1980, *'On the Localisation in the Superimposed Soundfield'*, PhD thesis, Technische Universität Berlin

Toole, FE 2008, *'Sound Reproduction'*, Focal Press

Toole, FE & Olive, SR 1988, *'The Modification of Timbre by Resonances: Perception and Measurement'*, JAES, vol. 32, March, pp 122-142

Wankling, M & Fazenda, B 2009, *'Studies in Modal Density – Its Effect At Low Frequencies'*, Proceedings of the Institute of Acoustics, Vol. 31. Pt 4.

Zahorik, P & Brandewie, EJ 2016, '*Speech intelligibility in rooms: Effect of prior listening exposure interacts with room acoustics'*, J. Acoust. Soc. Am. 140 (1)