THE TIME DOMAIN PERFORMANCE OF STANDARD LISTENING ROOMS: AN ASSESSMENT OF CURRENT ROOMS AND RECOMMENDATIONS FOR ACHIEVING IMPROVED COMPATIBILITY

BM Fazenda University of Glamorgan, Wales
KR Holland ISVR; University of Southampton, UK
PR Newell Consultant, Moaña, Spain
SV Castro Reflexion Arts SL, Vigo, Spain

1 INTRODUCTION

Since the mid 1970s, many efforts have been made to define recommended standards for listening rooms and audio control rooms for the broadcast industry. The intended uses for such rooms have largely been to aid the uniformity of decision making by the broadcasting organisations, where much interchange of programme material occurs, and for the evaluation by manufacturers of loudspeaker and audio equipment in general. However, it has also been recognised that many of these rooms can sound very different, despite falling within the specifications of any of the recommendations. Some room recommendations purport to emulate general domestic room characteristics, although their declared purposes of use do not cite their typical domestic nature as suggesting their use for assessing how things may sound in people's homes. They are generally intended to be common testing environments which are reasonably self-similar.

Moreover, the various standard listening rooms also get pressed into use for many purposes for which they were never intended. Nevertheless, for many uses it would seem to be desirable to reduce the colouration which most of the 'standard' rooms are acknowledged to possess[1]. Indeed it is the level of colouration which has dissuaded people from using them as sound control rooms in the music recording industry, where much more emphasis is placed on monitoring the recorded signal as closely as possible, and where the perception of fine detail in the sound is a high priority for quality-control purposes.

Historically, partly due to economic considerations, partly due to the relative absence of low decay end-user environments, and partly due to the low frequency time response problems of analogue recording systems, emphasis has been placed on low frequency pressure amplitude flatness. This has often been attempted by defining room aspect ratios that will then lead to the most even frequency spread of resonant energy. However, musical signal content is a combination of the time and frequency responses and it has been demonstrated that addressing frequency response solely is not in itself sufficient to render acoustically inoffensive the resonance problems in rooms[2]. It has also been shown that loudspeakers exhibiting less flat pressure amplitude responses, but faster time responses, can reveal more accurate information responses (in terms of modulation transfer function) than other loudspeakers with flatter pressure amplitude responses but longer decay times[3].

As digital recordings have yielded more accurate low frequency responses, and because much consumer listening has shifted to cars and headphones, the response errors given rise to by mixing music in resonant monitoring conditions have become more apparent. It would seem to be desirable that modern listening rooms and loudspeaker systems used for reference purposes should exhibit low-frequency decay responses which would not colour the subjective sound quality.

The evidence from much experience in the design and construction of low decay-time sound control rooms and corresponding monitor systems has shown that faster transient responses from the room/loudspeaker combination lead to easier decision making and less uncertainty in the sound

recording and mixing processes. Indeed, most of the criticism of these fast-decaying rooms is normally pointing to the uncomfortable feeling of being a bit lost with this surprising sensation of accuracy and transparency (lack of resonances), which is new to most people. Once accustomed to such conditions, professional users rarely seem to wish to revert to more traditional conditions[4].
It has also been recognised that a great number of professional sound mixing personnel have, over the years, empirically chosen to use monitor loudspeakers which exhibit fast time responses on the basis that they have found these to lead to more robust and reliable mixes when played in a variety of domestic circumstances[5]. Holland et al[6] have shown that the low frequency information content falls as the loudspeaker and room decay times rise, and that rooms with low frequency decay times well within many of the current ‘standard’ room recommendations, such as EBU, ITU and IEC [7,8,9,10], are well capable of blurring low frequency information.

Despite this new evidence, many loudspeaker systems currently in use for sound mixing purposes, perhaps due to their manufacturers feeling the marketing pressures for ever more compact boxes, are exhibiting low frequency decay times that are much longer than ideal, and this situation may be being ‘got away with’ because of the masking due to the largely unaddressed room decays. Research has shown that loudspeakers with slow low-frequency decays may lose significant amounts of information content[11], yet the sonic accuracy benefits of the faster responding loudspeakers may not be apparent when auditioned in rooms whose low frequency decay times are similar to, or greater than the length of the loudspeaker decay time. The reverse is also true, that the faster room decays may not be deemed necessary if the rooms are used with loudspeaker systems with commensurate low frequency decay times. This tends to lead to a subjective vicious circle, where low decay-time rooms and loudspeaker systems are each seen as unnecessary because they are each judged with reference to the other.

2 LOUDSPEAKER VS ROOM DECAY

Several current recommendations[10,12,13] give acceptable limits for loudspeaker transient decays as 2.5f, or even 5/f, where f corresponds to octave or third-octave band centre frequency, for decays down to levels of 1/e. These decays are shown as waterfall plots in Figure 1. It is clear that at increasingly lower frequencies the allowed decay becomes longer. A system with an appropriately extended low frequency range would be allowed to decay to -60dB in just under 600ms at a frequency of 40Hz. If the 5/f recommendation is taken into account the allowed decay time at the same frequency explodes to around 1s!

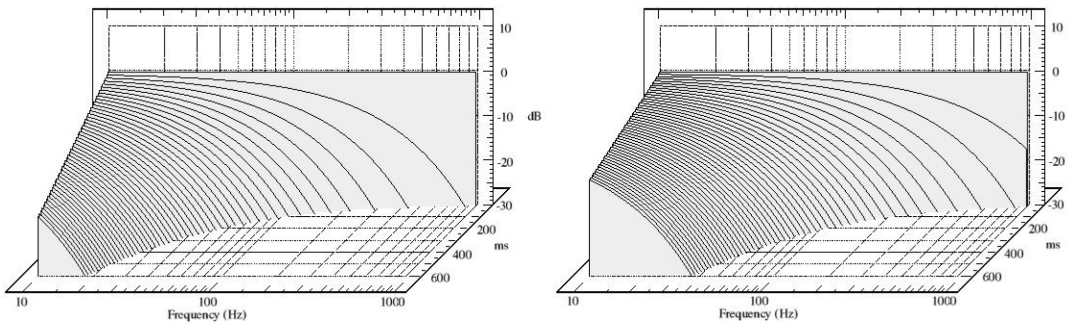

Figure 1 – Waterfall plots of recommended limits for loudspeaker transient decay response. a) Maximum decay time of 2.5/f for a decay level of 1/e; b) Maximum decay time of 5/f for a decay level of 1/e. f corresponds to octave band or third octave band centre frequency.

Assuming a linear decay, the recommended values for loudspeaker decay have been extrapolated to -60dB and are shown in Figure 2, plotted as decay-time against frequency. The responses for the aforementioned, standard, recommended listening rooms, assuming ‘maximum’ 400ms and ‘minimum’ 200ms mid-range decay times are shown plotted in a similar manner in Figure 3. From Equation 1 below, the lower limit of 200ms decay would correspond to a room of $50m^3$ and the upper limit of 400ms decay would correspond to a room of 400m^3, also suggesting a recommended maximum room volume.

Considering the 63 Hz octave band values in Figures 2 and 3 it is clear that the loudspeaker-decay recommendations are shorter than the maximum allowed room decay. This comparison suggests that rooms close to the maximum recommended levels may not be suitable for a critical evaluation of such loudspeaker responses.

The decay response of a 500m^3 dubbing theatre for the mixing of film soundtracks, installed with a very wide range, fast-decaying monitor system, is superimposed on the plot of Figure 3. This type of room is now widely accepted for the ease of evaluation of detail, despite the fact that according to EBU standards the recommended decay for such a room should be in the order of 425ms [7,8,10]:

$$T_m = 0.25\sqrt[3]{\frac{RoomVolume}{100}} \qquad \text{Equation 1}$$

If levels of RT and room volume are to be 'interlocked', as considered in the recommended standards, and critical subjective listening appears biased towards faster decaying low colouration rooms, it seems that highly accurate rooms should only be achieved using small volumes. It is widely acknowledged that this is not the case. This suggests thus, that definitions for such a critical factor as the time response of the room, and by association the reverberation time, should be influenced by experimentally defined subjective listening data.

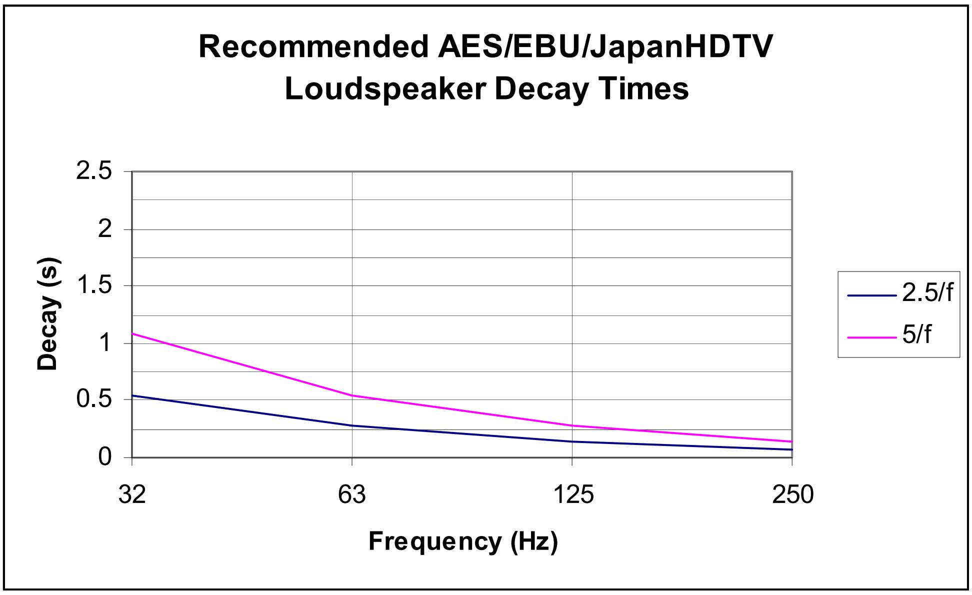

Figure 2 – Recommended decay times for loudspeaker systems.

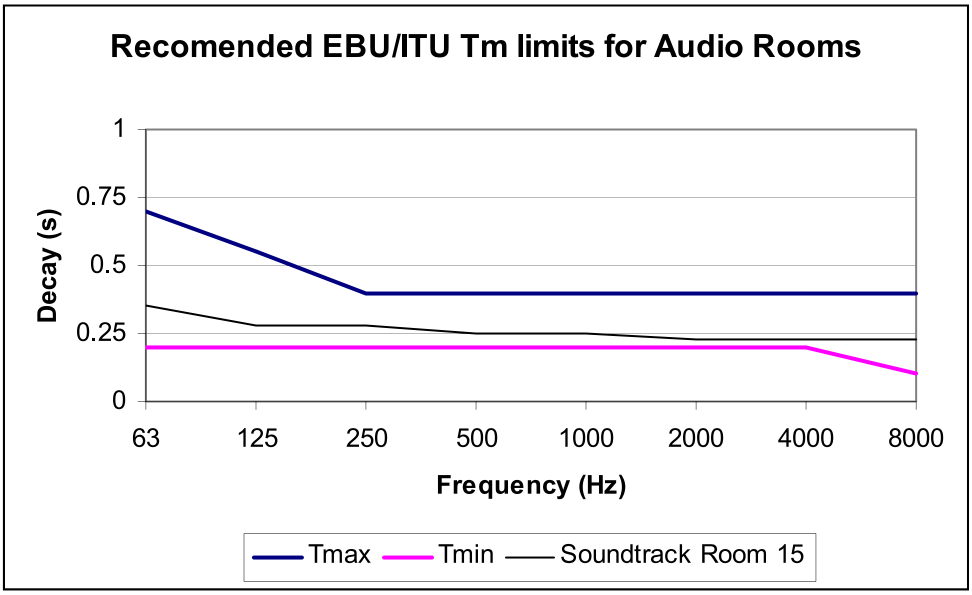

Figure 3 – Recommended reverberation times for listening rooms. Shown are the limits for reverberation time above and below an optimal RT of 0.2 seconds. The measured reverberation time of a 500m^3 multi-channel dubbing theatre (control room) for mixing film soundtracks is shown in the thin black line.

With a view to defining more subjectively uniform room conditions for quality control purposes, investigations have been undertaken to try to define the thresholds for room resonances below which the levels of colouration produced by the room acoustics were deemed to be non-intrusive, and, as such, could render rooms more neutral and hence more interchangeable[14,15,16]. Rooms complying with such criteria would also be more standardised amongst themselves than are the majority of current 'standard' rooms, and would serve as more optimal references over a wider range of uses.

Even if such rooms were deemed to be less domestic in nature, this would hardly seem relevant given that the range of domestic listening rooms is so diverse. What is more, recent studies have shown that the average response of all domestic and professional rooms tends towards being flat[17,18]. As flatness would anyhow be a typical characteristic of a low colouration room, no conflict would appear to exist between the use of these rooms for psychoacoustic tests, as broadcast sound control or post-production rooms, or indeed for the evaluation of either professional or domestic audio products.

3 ROOM RECOMMENDATIONS VS SUBJECTIVE PERCEPTION OF ROOM RESPONSES

In order to achieve standard conditions that could be replicated across professional audio facilities, a seemingly reasonable approach would be to reduce the offending factors to values which are below the detection by its users. Whilst direct sound, early reflections and mid/high frequency reverberation times are easily controlled using common acoustic tools available, the modal sound-field and corresponding response of a room at low frequencies still tends to require solutions of considerable expense, both economically and spacially. In order to develop efficient room designs, the subjective significance of different factors in the room modes and their thresholds of detection need to be clearly defined. This in consequence will lead to design solutions with more effective use of money and space instead of the over-kill techniques which are frequently used in order to guarantee neutral response conditions.

The common knowledge is that high modal density and an even spread of the eigenfrequencies provides a frequency response which may be acceptable, i.e. close to flat. This is in general true, but only above a particular frequency value (usually defined as the *Schroeder frequency*), because at these frequencies the decays of resonances are fast enough to be imperceptible, and the natural decay of the musical stimulus far exceeds the resonant decays. However, the relationship between decay time and frequency is such that at lower frequencies, decays become longer and possibly detectable. This is demonstrated in Figure 4 where the associated decays of resonances with different Q-factors are represented against frequency.

Earlier studies on this subject have suggested that the detection of resonances may be closely linked with their temporal characteristics[14]. This is especially true in room responses where low frequency modes have high Q-factors due to low damping present in the room at these frequencies. Further studies have demonstrated that resonances distort the original signal in such a way that their otherwise fast transient responses are smeared in time. The extent to which a resonance will affect an input signal has been shown to be strongly dependent on the Q-factor of the resonance and the temporal characteristics of the stimulus[19]. There is therefore sufficient evidence to support the hypothesis that the perception of resonant problems is significantly associated with the time domain response and not just with the frequency domain response of a system as is often supposed. It appears thus that any resonance that exceeds the detection threshold would be undesirable even though it could contribute to the overall flatness of the room's amplitude frequency response.

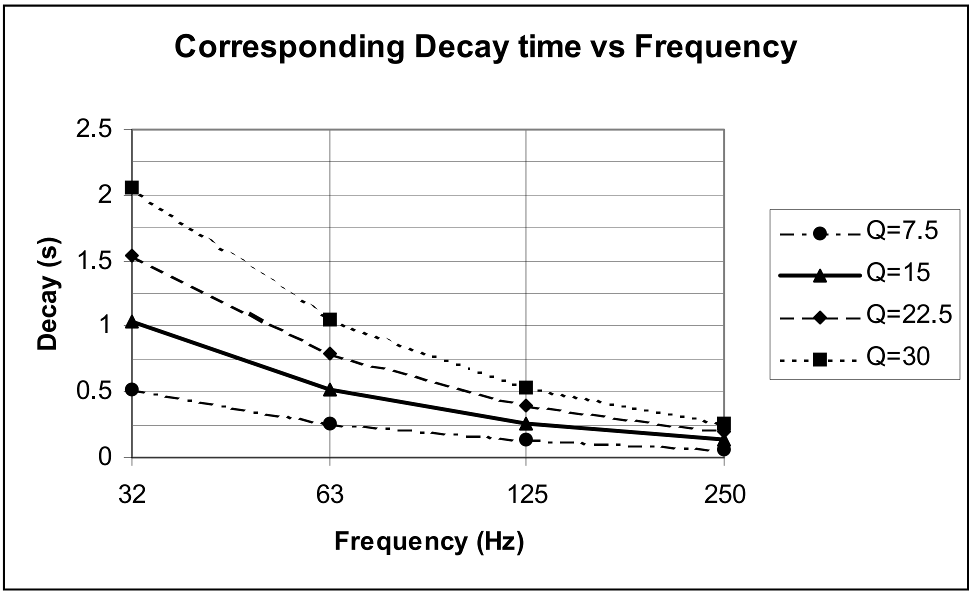

Figure 4 – Variation of decay time with frequency for different modal Q-factor levels. The thick black represents the threshold of detection for a group of room modes within 40Hz and 200Hz.

Recent studies concentrating on the audibility of changes imparted to the low frequency modal Q-factor, and as such to their decay, have defined a threshold of around Q=15 for a group of room modes between 40Hz and 200Hz [15]. These results suggest that any decaying energy existent in the room which is generally longer than the values represented by the solid line in Figure 4 will linearly distort the reproduced signal and hence be detected as colouration.

Comparison of Figures 3 and 4 shows that the current recommended levels for reverberation are well above theses values. For a frequency of 63Hz the recommendations for a standard room allow a decay time of around 700ms. However, subjects are able to detect decays as low as around 500ms. This differential further increases at higher frequency values, and at 125 Hz the recommended RT level is around 350ms higher than the detection threshold. Furthermore, the recommended value for RT varies with room volume (Equation 1) which leads to a rise in recommended low frequency decay for larger rooms, allowing for the occurrence of modes with even longer decays.

It should however be noted that other studies on the detection of room modes have identified thresholds which are somewhat higher, thus allowing more tolerant RT recommendations at the lower frequencies[17]. The experimental methods used in these experiments have differed somewhat in purpose and method to the ones presented here, and as a consequence the threshold results may have been defined too high.
Although indicative and at an initial stage, these new results on the perception of room modes certainly question the suitability of current room recommendations, especially at the lower frequency range.

4 CONCLUSIONS

Emerging results in the field of subjective perception of low frequency audio reproduction have led to some uncertainty on the relevance of current recommendations for room and loudspeaker performance assessment. The contrasting differences between current values for thresholds of audibility of low frequency reproduction and the standard recommended values for decay time of systems have shown that current recommendations may be too tolerant and somewhat remote from human listening capabilities. This realisation is certainly in line with practical experience in critical listening facilities where the sonic quality in the low frequency range varies drastically even between otherwise similar rooms (i.e. same volume and mid frequency RT characteristics). It has been demonstrated that this situation may be leading to a generalised acceptance of critical listening rooms exhibiting excessively long decay characteristics that, on the one hand prevent the correct assessment of loudspeaker decay responses, and on the other hand affect the correct monitoring of audio material.

This conflation of new evidence on the correct reproduction and evaluation of audio material calls for a revision and update of current recommendations that should consequently take into account the subjective levels of detection rather than allow the RT times to vary according to some arbitrarily percentage increase over an average mid-frequency value. The recommendations for critical listening rooms could therefore be revisited on order to update them with emerging research results based on levels of subjective perception.

The new recommendations would therefore lead to rooms with higher accuracy, which would not only allow the critical evaluation of audio material but also high accuracy electroacoustic equipment with short decay characteristics, even at the lower frequency range. Since the low frequency modal field in small rooms has long been recognised as one of the most problematic areas to solve, leading to large differences in sonic quality even between facilities with very similar architectural characteristics, rooms designed according to these new recommendations would also support a more significant standardisation of critical listening conditions.

5 REFERENCES

1. Fazenda, B.M., Davies, W.J., “The views of control room users”, Proceedings of the Institute of Acoustics, Vol 23, Pt 8, 2001
2. Fazenda, B.M., Avis, M.R., Davies, W.J., Jacobsen, F., “Perception of modal distribution in critical listening spaces”, Proceedings of the 11th International Conference on Sound and Vibration, July 2004, St. Petersburg, Russia
3. Holland, K. R., Newell, P.R., Mapp, P., “Modulation depth as a measure of loudspeaker low-frequency performance”, Proceedings of the Institute of Acoustics, Reproduced Sound 20, Vol.26, Part 8, pp. 107-115, Oxford, UK, 2004
4. Newell, P., “Recording Studio Design”, Focal Press, Oxford, UK, 2003

5. Newell, P.R., Holland, K.R. Newell, J. P., “The Yamaha NS10M: twenty years a reference monitor. Why?”, Proceedings of the Institute of Acoustics, Reproduced Sound 17, Vol.23, Part 8, pp. 29-40, Stratford-U-Avon, UK, 2001
6. Holland, K.R., Newell, P.R., Fazenda, B., Castro, S.V., “Excess phase effects and modulation transfer function degradation in relation to loudspeakers and rooms intended for the quality control monitoring of music”, Proceedings of the Institute of Acoustics, Reproduced Sound 21, Vol.27, Part 8, Oxford, UK, 2005
7. IEC Publication 268-13, “Sound system equipment, Part 13, Listening tests on loudspeakers”, 1983
8. ITU-R BS 1116-1, “Methods for the assessment of small impairments in audio systems including multichannel sound systems”, p.11, Switzerland, 1994-1997
9. Hoeg, W., Christensen, L., Walker, R., “Subjective assessment of audio quality – the means and methods within the EBU”, EBU Technical Review, pp.40-50, European Broadcasting Union, Winter 1997
10. EBU Tech 3276 – 2nd Edition, “Listening conditions for the assessment of sound programme material: Monophonic and two-channel stereophonic ”, p. 14, European Broadcasting Union, May 1998
11. Holland, K. R., Newell, P., Mapp, P., “Steady state and transient loudspeaker frequency responses”, Proceedings of the Institute of Acoustics, Reproduced Sound 19, Vol.25, Part 8, Oxford, UK, 2003
12. AES, Technical Document AESTD1001.1.01-10, “Multichannel surround sound systems and operations”, p.12, New York, USA, 2001
13. Yoshikawa, S., et al (11 more), Japanese HDTV Multichannel Sound Forum, “Proposal for the specification of control room for HDTV multichannel sound program production”, Audio Engineering Society 100th Convention, Preprint 4253, Copenhagen, 1996
14. Olive, S., P. Schuck, Ryan, J. G., Sally, S. L., Bonneville, M. E., "The detection thresholds of resonances at low frequencies", Journal of the Audio Engineering Society, Vol. 45, No. 3, March 1997
15. Fazenda, B., Avis, M., Davies,W., “Difference limen for the Q-factor of room modes”, presented at the Audio Engineering Society Convention, New York, USA, October 2003
16. Karjalainen, M., Antsalo, P., Makivirta, A., Valimaki, V., “Perception of temporal decays of low-frequency room modes”, Proceedings of the 116th Audio Engineering Society Convention, Berlin, Germany, May 2004
17. Toole, F., “Art and science in the control room”, Presented at the Reproduced Sound 19, Proceedings of the Institute of Acoustics, Vol. 25, Part 8, 2003
18. Mäkivirta and Anet, “The Quality of Professional Surround Audio Reproduction – A Survey Study”, 19th Conference, Audio Eng. Soc. (2001 June).
19. Fazenda, B., Avis, M., “Perception of low frequencies in small rooms”, Proceedings of the Acustica 2004 (IV Ibero-American Acoustics Congress, IV Iberian Acoustics Congress, XXXV Spanish Acoustics National Congress -Tecniacústica 2004-, Acoustics European Symposium), Guimaraes, Portugal, September 2004