Audio Engineering Society

Convention Paper 8310

Presented at the 129th Convention
2010 November 4–7 San Francisco, CA, USA



# First Results from a Large-Scale Measurement Program for Home Theaters

Tomlinson Holman[1] and Ryan Green[2]

[1] University of Southern California
Viterbi School of Engineering and
School of Cinematic Arts,
and Audyssey Laboratories
tholman@usc.edu

[2] University of Southern California
Viterbi School of Engineering
and Audyssey Laboratories[1]
ryang16@gmail.com

**ABSTRACT**

The introduction of one auto-equalization system to the home theater market with an accompanying reporting infrastructure provides methods of data collection that allows research into many practical system installations. Among the results delivered are histograms of room volume, reverberation time vs. volume and frequency, early arrival sound frequency response both equalized and unequalized, and steady-state frequency response both equalized and unequalized. The variation in response over the listening area is studied as well, and sheds light on contemporary use of the Schroeder frequency.

[1] (when this work was done, as a Master's Candidate in Electrical Engineering)

# 1. INTRODUCTION

## 1.1. Background and Environment

### 1.1.1. Multichannel nature

Multichannel audio is now thoroughly established in the marketplace as a principal delivery means for sound, especially sound accompanying a picture. It is difficult to even purchase a “pre-pro” or a receiver without surround sound capability, and the two-channel only market for equipment is greatly reduced from its prior marketplace dominance. That fact notwithstanding, most sound-only formats are still two-channel stereo in their delivery despite the known advantages of surround sound.

### 1.1.2. Loudspeaker/Room Equalization Systems

Many manufacturers provide some means of automatic calibration of multichannel systems, of varying levels of sophistication, systems that are typically not available on two-channel equipment. Calibration procedures potentially include identification of channels as to their bandwidth (satellite or subwoofer), time of flight for direct sound to the principal listening location, polarity, level trim for balance, and equalization. Certainly compared to uncalibrated systems these systems offer the potential to come much closer to the experience of the program material expected by the program producers than heretofore has been available.

One supplier of licensed technology in this field is Audyssey Laboratories, with its range of MultEQ offerings that include particular developments in the field of loudspeaker-room equalization [refs. 1–23]. All of the products licensed by Audyssey as MultEQ are supplied with a calibration microphone that has been qualified to production tolerances during manufacturing. Each microphone is compared in real time (single stimulus, calibration and device-under-test microphones measured simultaneously), unit by unit, on text fixtures designed and built by Audyssey. Because of the mass manufacturing nature of supplying all finished units with a microphone however, there must an allowed tolerance on level and frequency response. While small, these tolerances are nonetheless non-zero. These microphones have proved to be highly reliable and the measurements are repeatable over time, and have been shipped in the millions of units.

For custom installations (one manufacturer marks such products with a designator such as “CI” in their model numbers) a different microphone type may be employed typically by a custom installer, substituting for the calibration most end-users will perform with the microphone supplied with the equipment. Utilizing individual calibration in an anechoic chamber against a reference standard microphone, data is supplied with each microphone for its particular correction, and that microphone data file is loaded into software doing the equalization by serial number. Thus the tolerance on the equalization and level is made very tight, usually better than ±0.5 dB across frequency. The microphone capsule is an electret electrostatic type that has shown extremely good stability over time. Measured over a two-year interval the sensitivity and response change of a sample was within the measurement error. This was for a field unit, not a lab sample. If any question should ever arise about a particular microphone’s calibration, the unit under test can be compared to its original calibration. The type used has also proved to demonstrate virtually no change over much longer intervals.

### 1.1.3. Measurement Program

Accompanying the use of the calibrated microphones, designated home-theater installers in the Audyssey installer program follow a set of instructions for equalizing each sound system installation. This includes especially the choice of microphone locations for the multi-position clustering algorithm used within the program. For instance, the microphone positions are to be in representative listening locations. Note that these locations may not correspond to those that an acoustician might choose for measuring effects such as room modes, radiation load of subwoofers, or similar applications. Weighting the measurement program for likely listening locations does not seem to be on its surface problematic, but professionals measuring the same spaces with a different purpose in mind could come to a somewhat different set of measurements.

While it is not a requirement, the installer may choose to upload data files containing information about the installation, including the unequalized frequency response, and the equalization filters found by the software, to a web site in anticipation of this work. As of August 2010 more than a thousand installations have been documented. Each of these use a minimum of six loudspeaker channels, typically a minimum of four microphone locations and usually more (all within and at locations typically used in the listening area; more about this later).

#### 1.1.4. Selection of the Rooms for Analysis

The data is stored at the original resolution and contains more than 30,000 pre-equalization curves and corresponding filters, so the database requirements are challenging—this would have been impractical just a few years ago. Even now number crunching the entire set is extremely tedious and for this reason only a subset of the entire database was analyzed for this work, with various software filters used for the various outputs.

As practicality grows with increasing computer horsepower in coming years, perhaps a full analysis of this data can be done. For now, selection to those with a common target response curve[2], at least six loudspeaker channels, and at least five microphone positions resulted in a minimum of 275 of the available rooms being used for some of the following content. The fact that not all data was used means that statistical measures such as standard deviation in the data may be applied to understand the validity of it. In one representative case, reverberation time vs. frequency information shown below has a low standard deviation.

#### 1.1.5. Regional differences

Virtually all of the data collected at the time of this analysis was from U.S. homes. A program to perform the same measurements as a part of an installer program in Europe has just begun, so differences in construction between these continents can be quantified as time goes by. In particular, room acoustics of average U.S. listening rooms, which are larger than European listening ones and built with different underlying construction, are different. This idea is taken up further in section 3.2.

## 2. VALIDATING THE DATA

### 2.1. Measures taken to ensure the reliability of the measurements

#### 2.1.1. Microphone type, usage, positioning

**Microphone type**

A 7-mm face diameter (i.e., small), pressure (omni-directional) microphone provides a means to minimize variations due to angle of arrival. [ref. 24] The microphone is calibrated for pressure-mode response by comparison with a high-grade laboratory type microphone of similar diameter (Bruel & Kjaer Type 4938 capsule and associated electronics) supplied with its own electrostatic actuator curve and correction curve to 90° incidence. The microphone under test is driven in an anechoic chamber at grazing incidence across the microphone diaphragm. This makes negligible the pressure "congestion" at the microphone face in free air that forms a barrier and increases level at frequencies comparable to the face dimension of the microphone. The calibration includes both absolute level and frequency response. The response measurements in this case are carried out to the ultrasonic frequency of 40 kHz in order to be certain of the microphone's amplitude and phase response to 24 kHz.

Note that polarity can also be an issue (as it may be employed as part of a test on loudspeaker polarity), and that measurement microphones, through long-standing practice and convention, are wired in the opposite polarity to recording microphones: positive pressure at the diaphragm results in a negative voltage excursion at the output, and this fact must be accounted for.

**Microphone usage**

The microphone bodies are connected to the stand with a shock mount to filter low-frequency vibration from entering the measurement (which although common on studio microphones is almost unknown for measurement microphones), and on a small-diameter flexible arm arranged so that reflections off the stand are minimized.

**Microphone positioning**

The first microphone position in a series has a special meaning: it is the one that according to the instructions is to be placed at the principal listening location, and used for time-of-arrival data. Since this data can only be right for one position, this is why the first position is chosen as prime for measurement timing. In all other ways however, level and response, it is clustered given appropriate weighting to the other locations.

Locations are designated as those that represent the listening locations to be employed by users of the space. Unlike acoustical measurements made in a general way to characterize spaces, this method weights for listening locations. This is probably why some of the findings are

[2] Selected by the installer from a catalog of possible in-room response curves.

contrary to commonly held wisdom about small-room acoustics, which will be described below.

The number of microphone positions may range up to 32, depending on the endurance of the installer (five or six locations usually converges to a solution that is quite good). The microphone is only moved into each new position once, and the program circles among the channels for that location before pausing to ask for a new location.

### 2.1.2. Test signal

The test signal employed as a stimulus is a log-sine chirp. Since this is basically a swept sine wave, several benefits accrue. One salient one is that it is simple to tell if a frequency range is missing due to a bad loudspeaker driver for instance, or if the system is clipping, or there are serious, obvious distortion problems such as voice coil scrapes or rattles. While not specifically designed for this purpose, a listener becomes trained as to what the log sine chirp sounds like when properly executed by the system and one can almost tell the general shape of response curves simply by listening.

Interestingly when a variety of systems using various test signals parallel to those used in audio were tested for ghost cancelling operations for broadcast television, the Philips system based on a log sine chirp beat other schemes for ghosts caused by multipath in broadcast television. [29]

### 2.1.3. Signal-to-noise ratio

The software employed in MultEQ has an algorithm that ensures a high signal-to-noise ratio is available for analysis. Several measures are taken to ensure this:

> Each measurement at a position for a loudspeaker channel involves multiple log-sine chirps with averaging among them, which increases the signal-to-noise ratio over the number of averages since the chirps are correlated among each other, whereas the noise measured multiple times is uncorrelated. The summed signal level grows by 6 dB for doubling the number of averages, whereas the noise grows by 3 dB and thus s/n grows by 3 dB.
>
> Each measurement interval of the chirp signal spends one half of its time "looking" at the noise floor, not at the signal, and these two components (with and without signal present) are separated in the measurement system and used to test the s/n ratio.
>
> The system is interactive in that if too small an s/n is found in one set of chirps, then the chirp drive level is incremented upwards. If the s/n causes failure a second time, then another increment is taken. Should the measurement fail a third time at this increased level, then the fact that no measurement can be made that is reliable is reported to the operator.
>
> In a multichannel system the level for each channel is stored and returned to as one works through various microphone locations, thus speeding up the procedure.
>
> Given average power amplifier and loudspeakers sensitivities and room acoustics, the first chirp level will be at about 75 dB SPL[3] and will certainly not overdrive any reasonable system with which it is used. If the increments upwards result in too great a level for the drivers they will complain in such a way as is audible without destruction, although potentially a woofer could bottom for instance, it would only be briefly.

### 2.1.4. Bounds on equalization

The system provides bounds on equalization so that drivers are not pushed beyond their bandwidth limits, infinite holes are not attempted to be filled, and so forth. This is accomplished by performing a set of rules on the equalization after inversion of the source data

### 2.1.5. Data Taking and Reduction

The log-sine chirp is captured at high frequency resolution, typically much higher than the various target system's equalization resolution (at least partly for the anticipated purposes of this report). Each target product has a certain capability and all of the available capacity is utilized during equalization of the measured response to match the capacity of a particular model and channel to a target curve. The available DSP horsepower rated in MIPS affects the resolution of the final target filters, and thus their ability to follow small response details, especially to low frequencies. Improvements over standard FFT implementations[4] include the following:

---

[3] Reference 20 $\mu N/m^2$.
[4] Such as that in reference [25].

Frequency warping. Described in the references this improves low-frequency resolution by following equalization features more closely to lower frequencies than conventional FFT-based implementations and is present in all implementations [1, 6, 19].

Multirate. By splitting the spectrum up into regions, downsampling (decimating) the lower frequency parts, performing the equalization, and upsampling (interpolating) to the original rate and recombining correctly, great improvements over past implementations are possible. Subwoofer channels have had multirate applied with an 8× advantage for some years, while newly introduced algorithms (not included here because of the point in time where the data was extracted from the database) have increased resolution. It is difficult to put it into one simple number, but for a typical MIPS budget this technology can achieve a 32-fold improvement in resolution at the lowest frequencies.

What is stored is the high-resolution original pre-equalization response data, and the filter coefficients for equalization derived by the process at the various resolutions caused by the various capabilities of the target products. Since the product model information is stored as a part of the data, one could look at each filter type for its resolution and how that affects the post-equalized response curves. However, once the psychoacoustic data reduction described below is performed on the pre-equalized data, and on the post-equalized data by applying the found filter to the pre-equalized data, no comparison has been made across filter resolutions, although this could be done it was beyond the scope of the work here. Individual cases do show the advantage of higher filter resolutions and greater application of the principles outlined above.

**Psychoacoustic smoothing**

Both the original data, and the post equalized data obtained by applying the filter derived by the process to the original data, are smoothed with a rolling Equivalent Rectangular Bandwidth boxcar[5] filter developed by Prof. Brian C. J. Moore. [26] This filter is often used to describe the frequency resolution of the critical band filters in the ear, and it corresponds to equal increments of length along the basilar membrane in the ear of 0.89 mm.

Note that a rolling boxcar filter (gathering the underlying FFT frequency bins, variable in number with frequency according to an algorithm and summing them) at each output frequency provides data that looks surprisingly like high-resolution data, unlike "histogram" style presentations (usually 1/3-octave, but also other fractional bandwidth filters) that are common with real-time analyzers. The psychoacoustic justification for the rolling filter is that hearing does not employ a set of fixed bandwidth filters centered on particular frequencies, but rather the actual hearing filters are adaptable to various center frequencies and then have a particular bandwidth.

This particular filter (ERB) has been used in large-cinema applications by Holman over past papers [27–28], so the results may be compared across large- and small-room systems, restricted to the purpose of cinema program material (no concert halls or multipurpose rooms).

There can be a case made for capturing high-resolution data, and employing different smoothing schemes in separate frequency ranges to present the data. The frequency range of the lowest critical band of hearing is quite wide, whereas we definitely hear the note-by-note variations in level within the band caused by the interaction of the source and receive location with the room acoustics, in particular caused by low-frequency room modes. Thus a critical band filter is too broad to explain what is heard in this case. So in one way of looking at the problem, we might use fractional bandwidth filters of high resolution, say 1/24$^{th}$ octave[6], at low frequencies, below the Schroeder frequency[7], and cross over to ERB above the Schroeder frequency. This was one path we went down only to discard it when certain properties of the sound field became clear to us, which are described below.

[5] That is, all of the bins within the bandwidth of the filter contribute equally to the output, and all the bins outside the bandwidth are not counted. The number of bins varies across frequency according to the formula in reference.

[6] 1/24$^{th}$ octave because we already know we can hear note-by-note variations spaced at 1/12$^{th}$ octave, and we would like to have higher frequency resolution by some reasonable amount compared to the finest detail.

[7] See section 3.3.1.

# 3. ROOM ACOUSTIC MEASURES

## 3.1. Basic Room Characteristics

### 3.1.1. Room Volume

The mean room volume of 572 home theater listening rooms is 3,372 cu. ft. This is interesting because the rule of thumb for room volume of home listening rooms has been given as 3,000 cu. ft. in the past, which may now be seen as more than anecdotally correct. Since these rooms had custom installers with attendant higher budgets than average, it is expected they are somewhat larger than an overall mean of all rooms used for home theater. A room volume of 3,000 cu. ft. has been assumed for the design of equipment in the past, such as determining the required power output capability of amplifiers given specific minimum loudspeaker sensitivity and room acoustics. At that time, in the late 1980’s, an assumption on room volume had to be made, and it appears as though that was a very good assumption.

Figure 1 gives a histogram of home theater room volumes. A handful of rooms over 12,000 cu. ft. were discarded from the data as being outliers. The resulting curve shows a distinct skewness towards higher volume.

### 3.1.2. Room Dimensions and Cost

Some installers have only entered the room volume. Others put in dimensions, reduced in the case of odd-shaped rooms to an equivalent estimated three-dimensional box. This is a limitation because of the nature of the data gathering. The mean room height for those entering the parameter was 9.00 ft. The average square footage is thus 375 sq. ft.

It should be noted that probably not a large fraction of these cases are in fact dedicated home theater rooms, but rather general-purpose media rooms combined, often in some kind of open plan in today’s architecture, with family rooms and so forth. Unfortunately in a survey of this scale, inputting real plans and room purposes are beyond the scope of what can be done practically.

At upper middle class housing construction prices averaging $300/sq. ft. (the range is large, from $100 (tract house, part of large development – $400 (higher quality but could become even more)), that 375 sq. ft. cost about $112k.

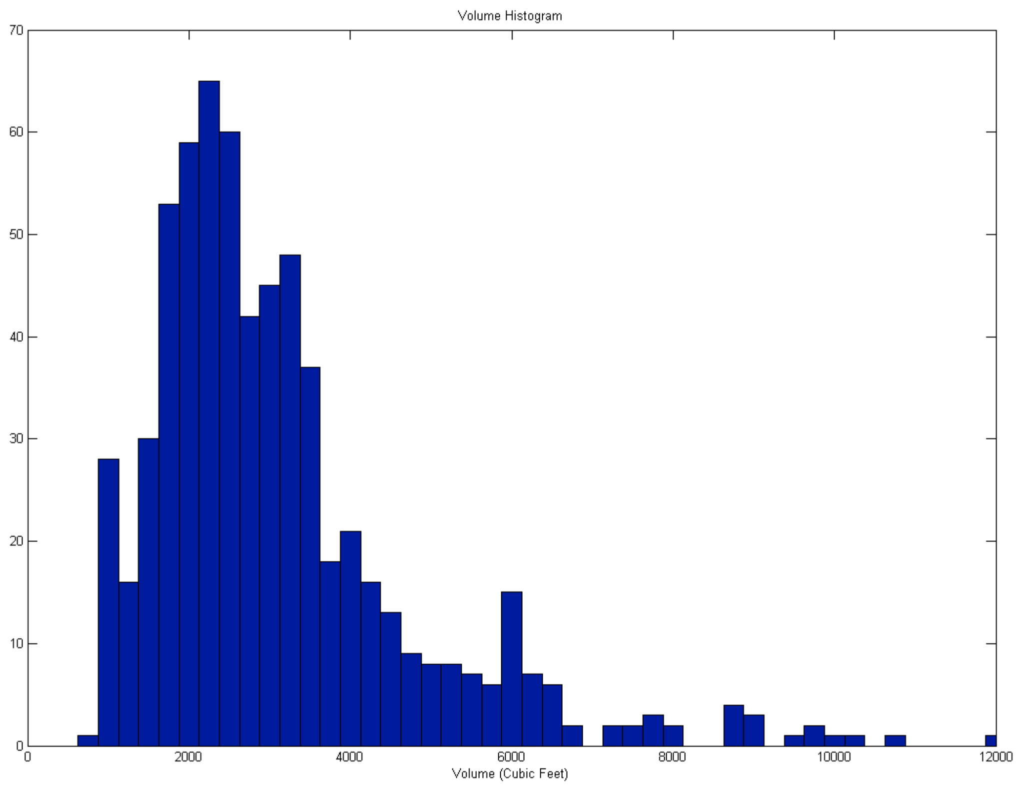


**Fig. 1. Histogram of Room Volume: Number of sites vs. Volume in cubic feet.**

### 3.2. Reverberation Time

Certainly one of the most important properties of room acoustics of large spaces is the reverberation time, as is widely known. RT60's impact on intelligibility, clarity, warmth, and other factors have been studied extensively. Attention is paid to concert halls and performing arts centers because they cost so much, and the acoustics are known to be critical, so that great attention is paid to the room acoustics.

Less widely known are the facts of home theater acoustics. This is partly because the naturally occurring reverberation time in small rooms is lower than that in large ones, and thus less of a problem for intelligibility. On the other hand, gains in localization and dialogue clarity do occur with increased front-channel loudspeaker directivity and thus less interaction with reverberation even in home theaters, with changes in directivity being plainly audible in control rooms, the size of home listening rooms.

> Changes of the mid-frequency directivity of ca. 3 dB were very noticeable, due to the change in definition, spatial impression, and presence. [34]

So reverberation time in rooms of the size with which we are dealing is a factor, even though perhaps less of one than it is in larger rooms. In fact, the reverberation time and the loudspeaker directivity interact—a fact that is well known in large rooms but not so obvious in small ones. [35–36]

Another factor comes to the fore in home theaters that is not so troublesome in large rooms: the effects of standing waves, at frequencies called modal ones. As is well known, all other things being equal, the modal density (number of room modes per unit of frequency) will be greater in a given band in a large room than in a small one. Thus home theaters exhibit more obvious modal behavior than concert halls. “Modal ring down time” is a more appropriate descriptor at many frequencies than is “reverberation,” but lumping classical reverberation and modal ring down together is commonplace, since the two are difficult to separate.

Figure 2 gives the measured Reverberation Time RT60 vs. frequency in octave bands from 63 Hz to 16 kHz, along with ±1 standard deviation of the RT for this survey.

Figure 3 gives a scatter diagram of Reverberation Time vs. Room Volume for this database.

### 3.3. Work by others

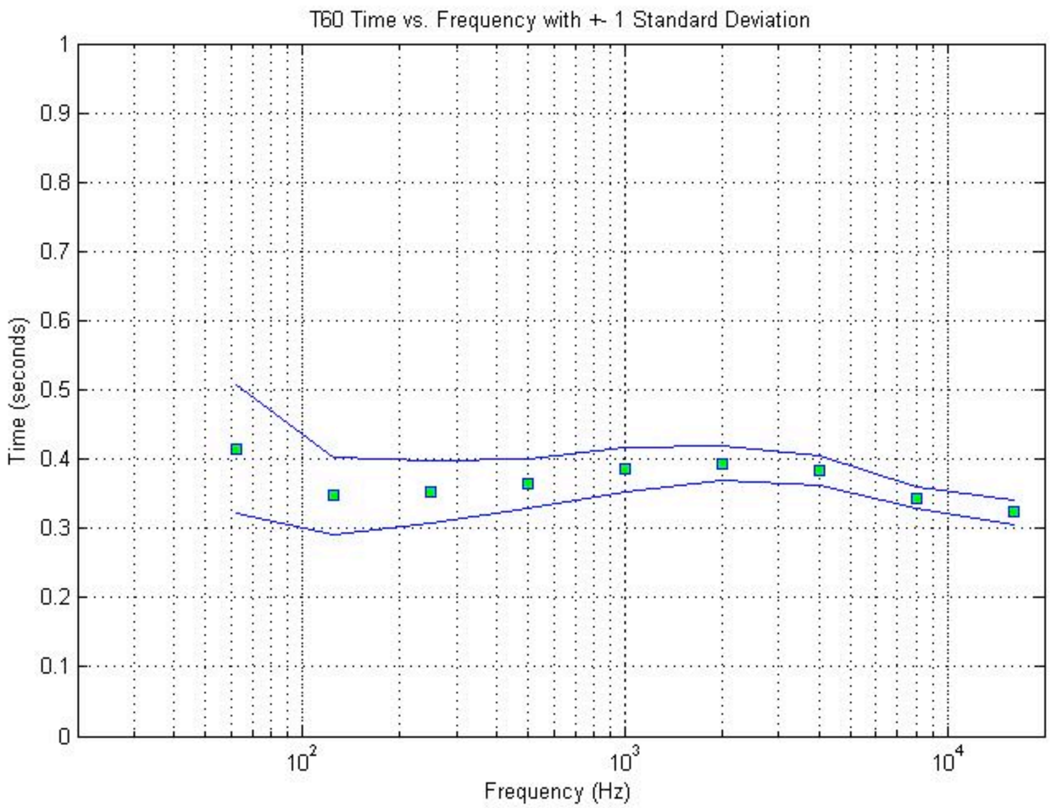


**Figure 2. Reverberation time RT60 in seconds vs. Frequency in Hz with error band of ±1 standard deviation.**

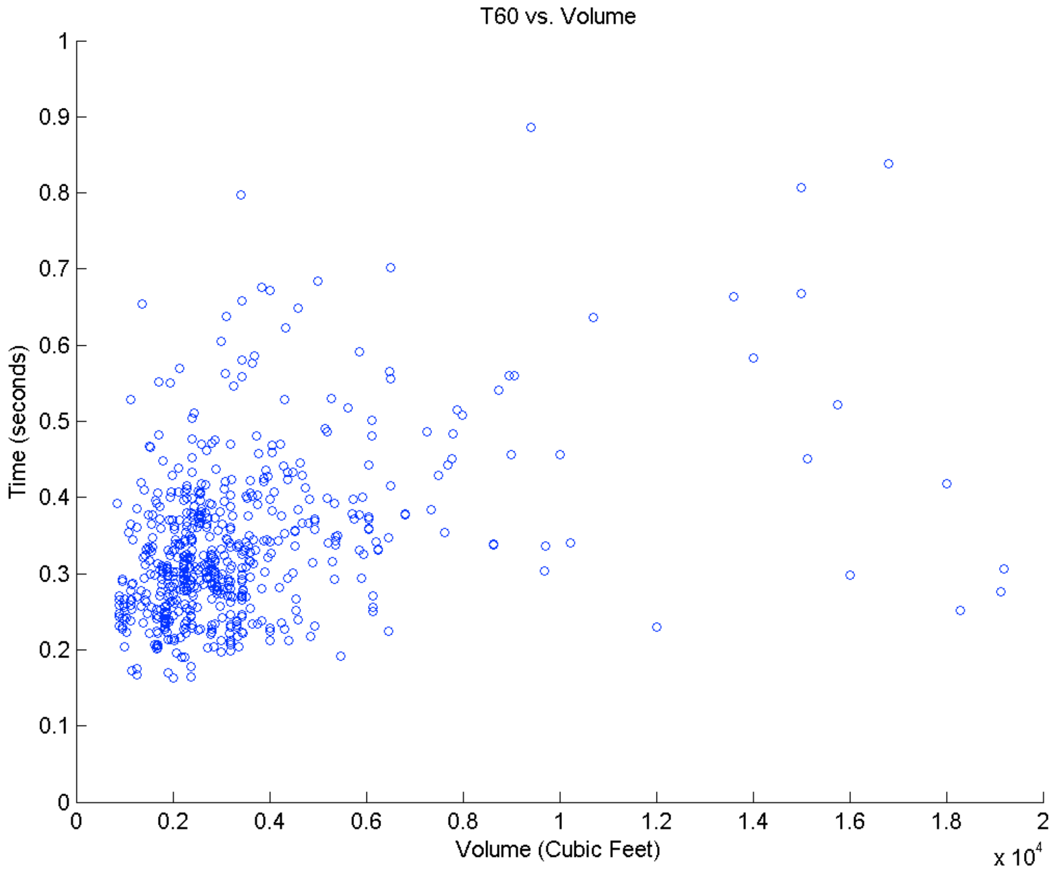


**Fig. 3 Scatter diagram of broadband RT60 in seconds vs. Volume of the room in cubic feet.**

EBU has recommendations for reverberation time for listening rooms, and interestingly the recommendation does not vary between stereo and multichannel rooms. An equation for RT60 in monitoring is given by EBU 3276–1999 "Listening conditions for the assessment of sound programme material: monophonic and two-channel stereophonic". This same recommendation is also made for multichannel sound in Supplement 1 (revised 2004) "Listening conditions for the assessment of sound programme material: multichannel sound." [31]

The EBU equation is:

$T_m = 0.25(V/V_0)^{1/3}$ s where:

$V$ = room volume in cubic meters and

$V_o$ = reference room volume of 100 m$^3$

For the average room volume of 3372 cu. ft. (95 cu. m) found in this survey, the recommended RT by EBU is 0.24 s, about 30% lower than that actually found in listening rooms.

A few studies of reverberation in home rooms have been performed, one of which used a very large database. [30]

Reference 30 shows that the furnished living room RT60 in 3211 living rooms in Madrid as 0.48 s at 500 Hz, about 25% longer time than found in this survey. Regarding the construction the authors say “All the rooms have heavy walls and ceilings. The interior partitions are hollow brick walls covered with plaster. The floor covering is mainly terrazzo or parquet finish glued onto a leveling layer of cement mortar.” [30] So it is not surprising that this survey shows somewhat longer RT60s than the current one, since most of the construction in this survey is walls of gypsum board on wood or metal studs, flooring typically consists of joists covered by several layers of plywood, and ceilings are gypsum board over joists.

The unfurnished rooms in reference 30 show the expected longer reverberation time caused by the lower absorption in the spaces than that in furnished rooms. By analyzing the report, the relative reverberation time between furnished and unfurnished conditions can be calculated in a rather small room size that was reported of 700–1,000 cu. ft., as a ratio of 1:4.3. An earlier survey [32] shows a ratio of furnished to unfurnished in the range of 1:5 and 1:3, depending on the data set (reporting on other’s measurements). The thesis of that work is that the average reverberation time came down over the period from the 1950’s to the 1980’s, so the ratio went up.

We have no way to separate the effects of furnishings from the effects of surface absorption in our measurements, although we note that furnishings do form a great deal of the absorption in rooms.

Some specifications for RT60 call for reverberation time to be monotonic with frequency, with a flat midrange, and permitted low-frequency rise and high frequency droop. The source of the low-frequency rise is given because of the lower absorption of materials in the room at low frequencies, and the high-frequency droop is caused by air absorption.

Most of the wall construction of these rooms is gypsum board on studs, and ceilings are gypsum board on joists. The absorption of such construction has been studied by Bradley [41]. None of the configurations of walls (varying cavity filling, gypsum thickness, etc.) show an increase in the absorption that would correspond to the dip in the RT60 frequency characteristic found here. Furthermore the longer reverberation times for empty vs. furnished rooms in [30, 32] shows the underlying construction may not be the source of this “dip” in RT60. This is a bit of a stretch since their construction is different, but it seems likely that unfurnished rooms of gypsum board construction would also have a rather longer RT60 than furnished rooms to the point that the unfurnished underlying construction is not the source of the dip in RT60. Thus furnishings are indicated to be the most likely source of this characteristic

Another set of authors has performed measurements in a professional control room for the effect on RT60 by audio equipment and furniture. [42] They started with the empty room, measuring its RT60, and added just the console and then the rest of the equipment and furniture. Although their RT60 is much shorter than for our survey (125 ms for the midrange), their RT60 vs. frequency characteristic shows a dip in RT60 in the 125 Hz band not unlike our own, in the condition of the room with just the console. Adding other items causes the overall higher frequency curve to come down, obscuring the dip.

The unknown absorptive effects of furnishings is the likely source of errors in the prediction of RT60 by various means. On one project the Sabine equation predicted RT60 from +17 to 39% across frequency (range 125 Hz to 4 kHz) from that observed, and the Arau-Puchades equations predicted it as from –4 to –33% of the actual reverberation time.

### 3.3.1. Schroeder Frequency

The Schroder frequency is calculated from the room volume and reverberation time. The Schroeder frequency for a large number of home listening rooms is graphed in Figure 4.

The Schroeder frequency (SF) describes that frequency above which high modal overlap occurs, above about 3 in the modal overlap index. The modal overlap index is the average modal bandwidth divided by the average frequency spacing between modes. In this frequency region, the time-averaged fluctuations over space and frequency obey certain statistical laws. Below the SF, standing waves (modes) dominate the uniformity of the sound field. [39]

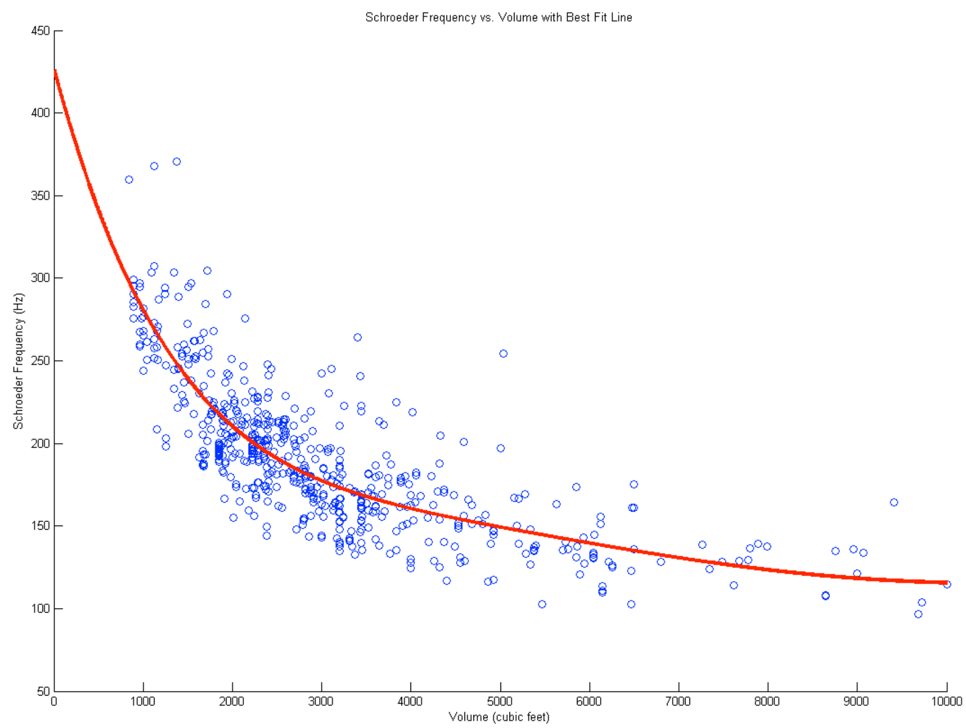


**Fig. 4 Schroeder Frequency Hz vs. Room Volume in cu. ft.**

### 3.3.2. Reverberation, directivity, and number of channels

One thought is that there should be a distinction between stereophonic and multichannel listening, in that surround sound provides the "missing" envelopment ingredient in two-channel stereo that is provided in stereo by the room acoustics interacting with the directivity of the loudspeakers to produce a particular direct-reverberant ratio, and that this ratio may be desired to be lower for stereophonic than for multichannel reproduction. This idea is supported in reference 34 with a statement that the directivity of a loudspeaker for monaural listening should be somewhat less than that for stereo listening. Extending the idea says that there should likely be a difference in directivity for surround sound applications compared to stereo.

Nonetheless an authority such as the EBU makes the same RT60 recommendations for both room types. The finding of reference 34 was based on experimental listening tests. It is not known how EBU came up with their recommendation.

## 3.4. Effects of early reflections and periodicity of reflections

Early reflections have been thoroughly studied in the context of small rooms. S. Bech among others has been active in this area. Periodic reflections (flutter echoes) are also well understood.

For our database we have not developed a procedure to evaluate the effects of early reflections or periodic reflections with statistical means across many rooms. Waterfall displays of frequency response over time are available room by room, but how to perform data reduction is unclear. The effect that does appear is in developing the reverberation time algorithm. As is typical, the direct arrival sound is found within the time interval, and then a time is added to it for the sound field to build up beyond the early reflections, before the reverberation time measurement begins, and it is terminated before the level vs. time curve descends into the noise floor. This mimics methods used manually by acousticians for many years in examining decay curves.

### 3.5. Background noise

Certainly background noise plays a role in home theaters. However, the small size of the microphone used in the installer program, and required for low-diffraction accurate measurements in mixed sound fields, precludes making noise level measurements down to the actual noise floor of home theaters. This is because the small active area of the diaphragm (much smaller than studio mics) produces a higher noise floor. The combined acoustical and microphone noise floor is calculated during the measurements to ensure they do not affect the reliability of the response measurements with an s/n ratio test, but the data is not saved.

Work by Fielder and Cohen [33] however provides us with a 27-room database of home listening spaces (whether general rooms with a sound system or dedicated home theaters) from which to understand the most likely noise floor. Figure 5 gives the information found by Fielder/Cohen (the middle curve over most of the range) along with two other curves: the threshold of hearing (the lowest curve over most of the range), and a good control room's background noise level (the highest curve over most of the range).

The control room noise floor is dominated by air conditioning rumble at 50 Hz (note not the line frequency of 60 Hz—it's not hum), and by console microphone preamplifier fan noise at 500 Hz and several harmonics, which is plainly audible as the sources in the room. (These preamps may be remotely located but in this teaching environment being able to manipulate them in front of students is more important than the ultimate room noise floor.) More general and smoother noise spectra are due to the HVAC system. A cautionary tale is shown in this figure: the professional space is noisier than the average home listening space, and this is not uncommon. Thus low-level sound may be recorded on the track that wasn't heard in the professional space, but becomes audible at home. This is partially compensated by the fact that professionals monitor at louder levels than consumers, but is nonetheless a source of potential problems.

If we were to be faced with these measurements as professional acoustical consultants, we would recommend isolation of the fan causing the 50 Hz problem, and isolation for the microphone preamplifiers (perhaps in an isolation cabinet with low-speed cooling fans and baffles to its exterior). These two measures may get the noise floor down to the average home listening room level, which perhaps should be required of professional spaces. However, we are also aware that this noise level is difficult to achieve in office and industrial buildings.

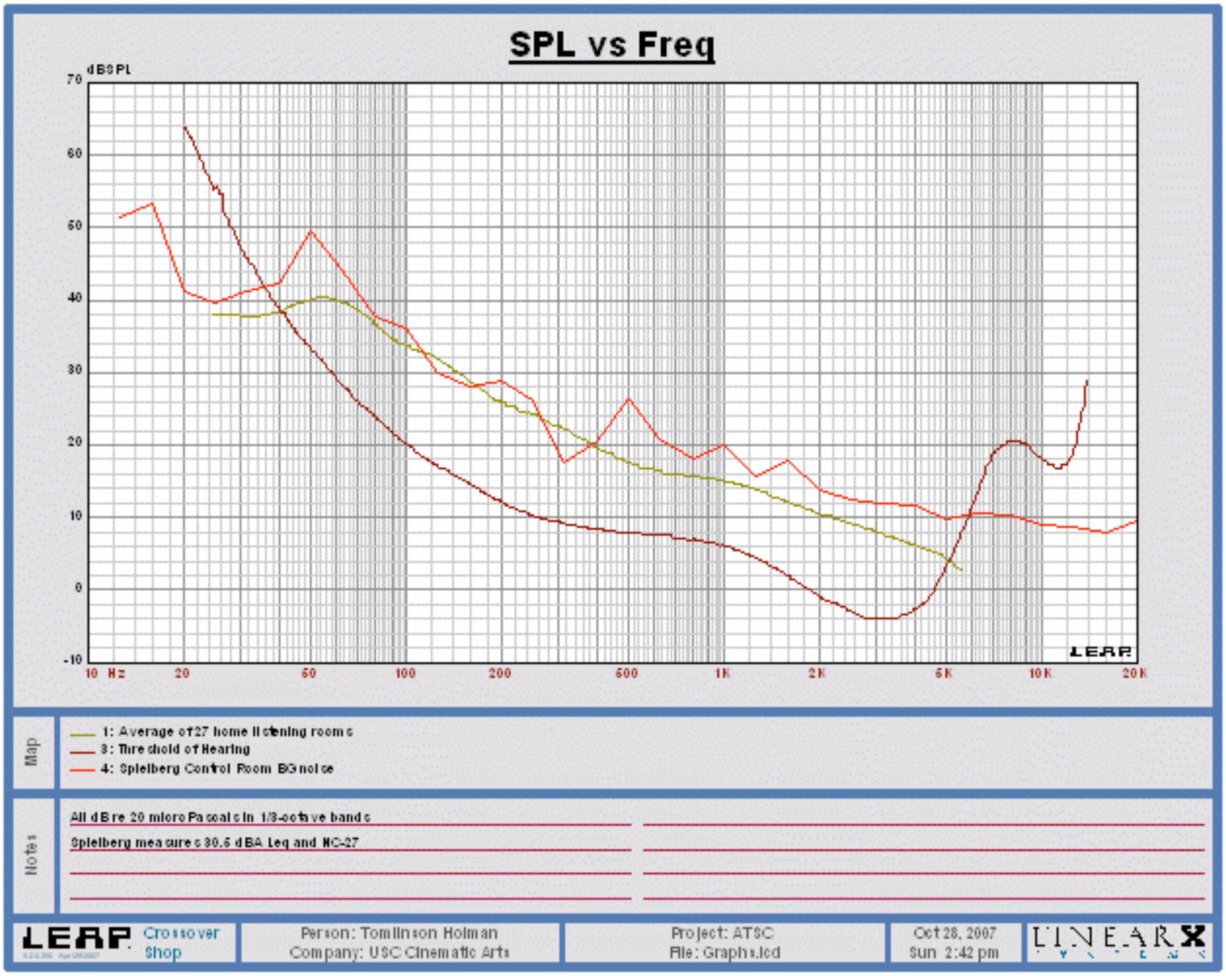


**Fig. 5. Background noise level and threshold of hearing in 1/3-octave band levels. dB SPL re 20 µN/m$^2$**

**vs. Frequency in Hz. The lowest curve (brown) over most of the range is the threshold of hearing. The smooth intermediate curve (green) is Fielder and Cohen's data for 27 home listening rooms. The curve that is highest in the 500 Hz to 5 kHz region (red) is the noise floor of a good control room measured at the position of the mixers.**

## 4. SOME SYSTEM-BY-SYSTEM DATA

### 4.1. Frequency response of several rooms pre- and post-eq

Three rooms were chosen at random from the database to illustrate a variety of issues. All of the data presented in this section is for the left-channel satellite loudspeaker of a multichannel system. The target response curves always contain a low-frequency high-pass response because satellites should not be driven beyond their bandwidth capability, and they will be used in a bass managed system with a common-bass subwoofer.

Figure 6, 7, and 8 are of the same room, pre-eq, post-eq, and deviation from the rms average equalization (take the deviation from the average at each frequency for each position, rectify to make all the error the same direction, and plot relative to the average). The mean of the deviations is shows as a heavier line. Figs. 9–11 are for a second room, and Figs. 12–14 for a third room.

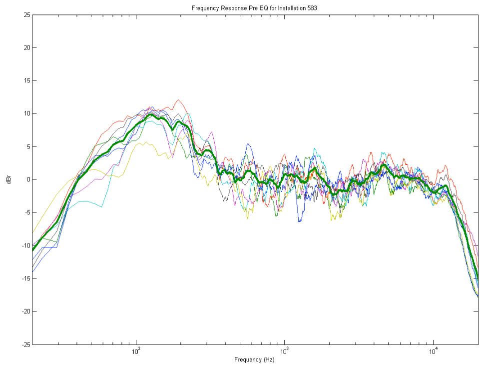


**Fig. 6 One room (583) measured in multiple positions and the average response. dB vs. frequency in Hz.**

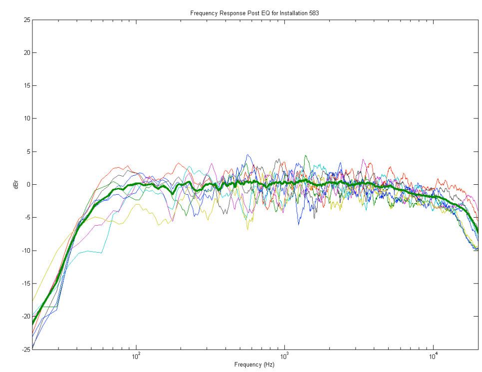


**Fig. 7 The same room post equalization. dB vs. frequency in Hz. This is the response of a satellite in a bass-managed system.**

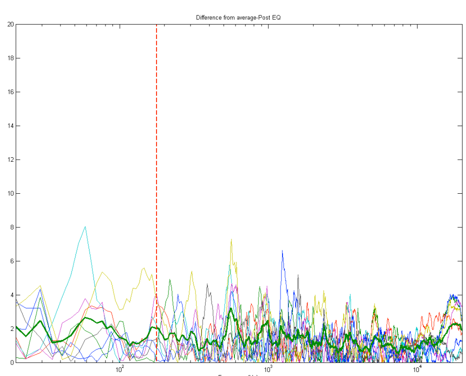


**Fig. 8 Differences from the average for the same room. dB vs. frequency in Hz. The red line indicates the calculated Schroeder frequency.**

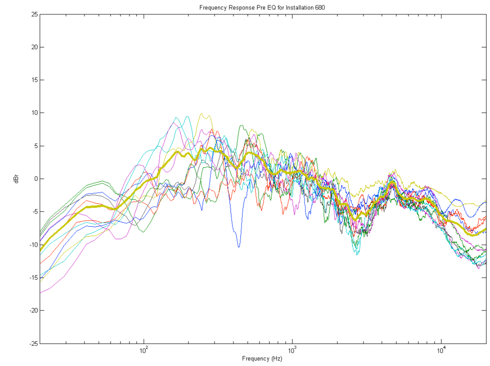


**Fig. 9 A second room (680) measured in multiple positions and the average response. dB vs. frequency in Hz.**

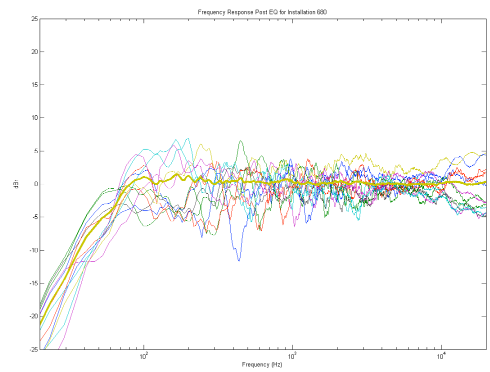


**Fig. 10 The same room as Fig. 9 post eq. dB vs. frequency in Hz. This is the response of a satellite in a bass-managed system.**

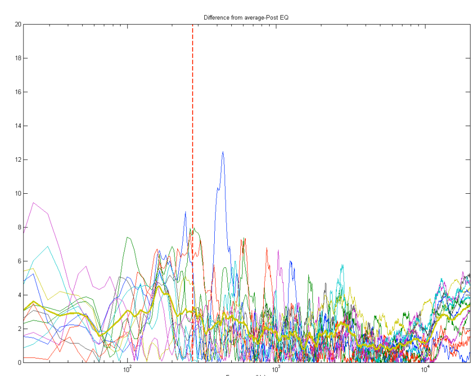


**Fig.11 Differences from the average for the room of Fig. 9. dB re average response vs. frequency in Hz. The red line indicates the Schroeder frequency.**

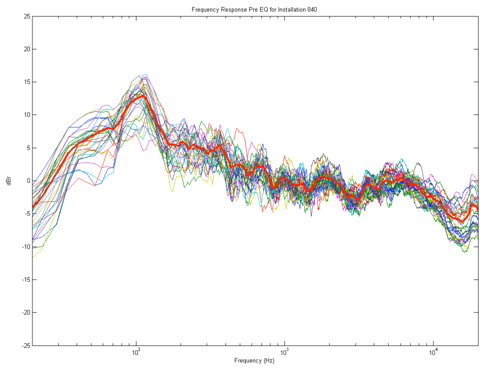


**Fig. 12 A third room (840) measured in multiple positions and the average response. dB vs. frequency in Hz.**

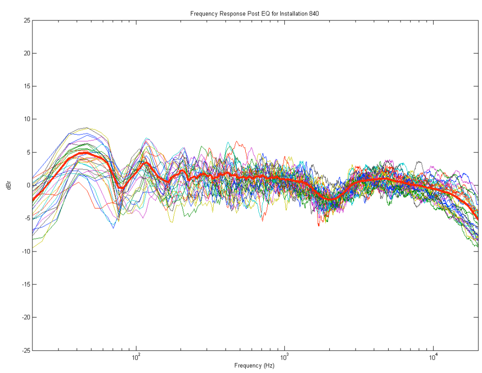


**Fig. 13 The same room post eq. dB vs. frequency in Hz. Note that a different target response curve was employed in this room, having a deliberate dip in the response.**

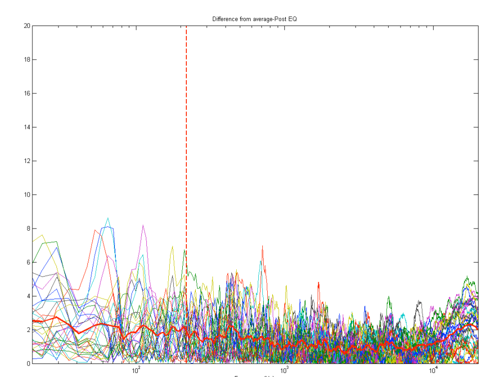


**Fig. 14 Differences from the average for the room of Fig. 12. dB re average response vs. frequency in Hz. The red line indicates the Schroeder frequency.**

### 4.1.1. Discussion of results

Several things may be noted by observing this data, although it is only on a very few rooms. Consolidated data is shown in the next section.

**Measures of quality**

1. A first measure of quality of a sound system is conformance to a target response (more about this later). Equalization clearly helps this conformance in all cases.

2. A second measure of the quality of a sound system is its uniformity of coverage of the listening space. Data for this has not been available in the past, and this is newly presented here insofar as is known. Uniformity of coverage is best illustrated by the average deviation from the average response, shown as the heavy curves in Figs. 8, 11, and 14. A surprise is that a system that starts out very non-flat and thus wouldn't be considered to be very good, Fig. 12, room 840, is greatly improved with equalization, and it has a small deviation vs. frequency, Fig. 14, although the same thing could be said of room 583.

**Other observations**

1. Fig. 12 illustrates why some may believe that only one parametric band of equalization might be adequate. Known from the introduction of the Dolby Cinema Processors since 1975, this idea still has some traction, and indeed, the one worst region might be able to be fixed with one parametric band, so long as the frequency, level, and Q of that required equalizer can be found. However that would leave major problems untreated.

2. Some features can be identified in the various deviation curves. One is that at high frequencies, above about 8 kHz, the data takes on an orderly nature. This is clearly due to tweeter dispersion. A second is that in some cases (not shown here) crossover frequencies can be identified. Since the measurement data is taken at listening locations, a wide variety of heights are not typically included in our data. Since most loudspeakers use a vertical array of drivers and thus have better horizontal than vertical uniformity of directivity through

crossover, it is the ones having woofer and tweeters side-by-side with less than ideal crossovers that demonstrate this effect, a fairly small number, especially in left channels (it may be more common in center channels).

3. These three rooms demonstrate mid-bass boosts. It seems likely that the loudspeakers are designed in anechoic chambers for flat response, and then when used in a room, suffer from build up in the mid bass due to the increased radiation load on the speaker's output caused by the nearby room boundaries, and the modal response of the room. This certainly contributes to a lack of dialogue intelligibility [38][8], which is a frequent complaint about Hollywood mixes heard at home. The combination of the frequency response, far-field listening conditions [36], and background noise frequently above that of the average quiet listening room, all combine to lead to intelligibility problems. The recommendation would be to design loudspeakers at a minimum with an average room load in mind, and in better cases to employ loudspeaker-room equalization.

4. A BBC Research Report [37] gives information on the required frequency response match between stereo loudspeakers required to maintain phantom image stability over frequency. Very few loudspeakers have ever been made in production to such tight tolerance, so equalization is necessary for imaging. Multichannel systems with a center channel possibly reduce this requirement, but it would apply to interstitial images between left and center and center and right, possibly to a different degree. The BBC results are shown in Fig. 15. This was not specifically studied here.

5. In this small sample, the Schroeder frequency is not identifiable with respect to the spread of the average responses increasing at low frequencies, as would be expected. In fact in some data clearly the worst frequencies are well above the Schroeder frequency. This may be due to the fact that we have restricted the measurement positions within the room to listening locations, which deliberately avoid the corners where modal effects are more prominent, for instance.

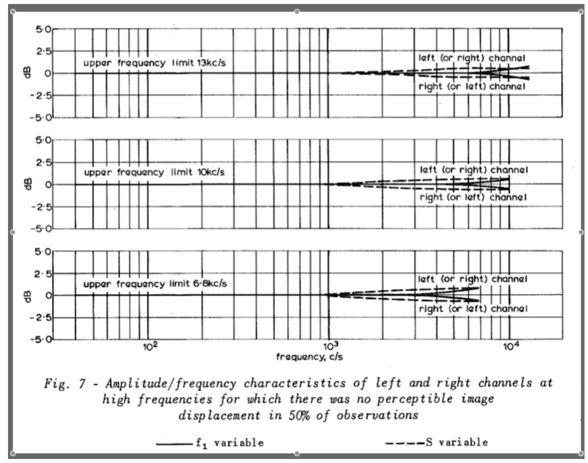


**Fig. 15 BBC Research Report results cited in the text. It shows the requirement for channel matching in order to maintain phantom imaging. For an upper frequency limit of 13 kHz, a channel-to-channel match for a left-right pair is about ±1 dB for low frequencies up to 13 kHz. dB in relative level vs. Frequency in Hz.**

## 5. CONSOLIDATED DATA

### 5.1. Steady State Response Data

All channels of all rooms that met certain criteria for data integrity including a minimum number of loudspeaker channels (5 satellite channels in this case) and a minimum number of microphone locations (5) were analyzed.

Fig. 16 gives the average pre-equalized steady-state frequency response of 275 installations. The number of underlying data curves is at least 34,375. (At least because some used more than the minimum number of loudspeaker and/or microphone locations; the actual number is higher.) The curve (Fig. 16) is smooth because of the number of rooms involved and averaging over those rooms—no added smoothing was done over the ERB filters so each underlying room appears to have the deviations such as those shown in Figs. 6, 9, and 12.

Note again that these are satellite loudspeakers intended to only be used with a bass-managed common bass subwoofer.

[8] From the paper's abstract: "Previous work has highlighted deficiencies in the ability of the STI metric to satisfactorily recognise the subjective loss of intelligibility that occurs with sound systems having poor frequency responses, particularly in the presence of reverberation."

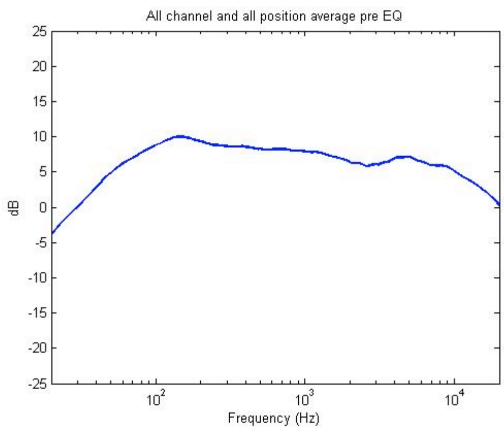


**Fig. 16 Pre-equalization frequency response for 275 installations consisting of >1375 total satellite loudspeaker channels measured at a minimum of 5 microphone locations for >34,375 underlying data curves. dB (arbitrary 0) vs. Frequency in Hz.**

Figure 17 is the post equalized data for rooms that employed one target curve, having a deliberate high-frequency rolloff and a mid-range dip that has been found to be a subjective improvement on many loudspeaker models.

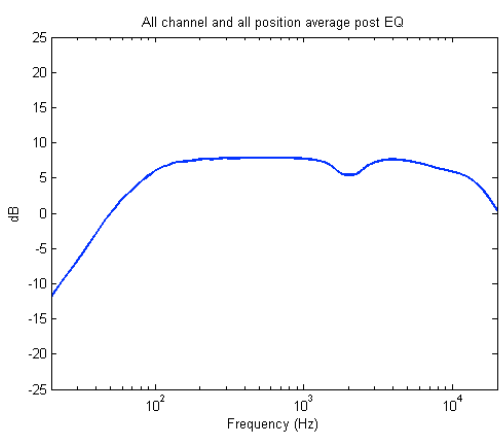


**Fig. 17 Post equalization frequency response to a target curve with a high-frequency rolloff and a deliberate mid-range dip. All other factors are the same as Fig. 16. The high-pass low-frequency characteristic is made a part of the bass management system. dB (arbitrary 0 but same as Fig. 16) vs. Frequency in Hz.**

Figure 18 is the average deviation from the average like the data shown in Figures 8, 11, and 14. The database selection is the same as for Figs. 16 and 17

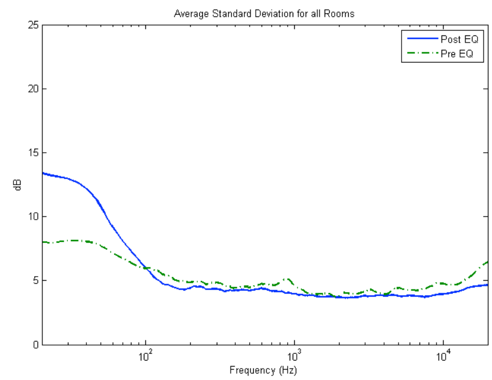


**Fig. 18 Average standard deviation across microphone positions using all channels (one at a time, not summations) for 275 installations of at least five satellite channels. Low frequency data is outside the band of the systems. Post equalized deviation is the solid line; pre-eq the dashed one. dB re Frequency in Hz.**

### 5.2. Transient Response Data

Corresponding to the Figs. 16–18 are the transient, first-arrival cases illustrated in Figs. 19–21. An algorithm finds the first direct sound in the interval of the time record version of the FFT and begins analysis. The analysis runs for 5 ms, and then a 1 ms decay using a Blackman-Harris window function is employed (no windowing is needed at the beginning of the interval since the direct sound is so much greater than level than the background noise before it: thus a symmetrical filter is not needed). A low frequency restriction occurs because of the short time of the filter, so the scale is truncated to 200 Hz.

Many other choices could be made including ones that are frequency adaptive as to truncation length. However, for this preliminary work with such a large database, calculation time was at a premium, so the particular windowing as described was used for the present.

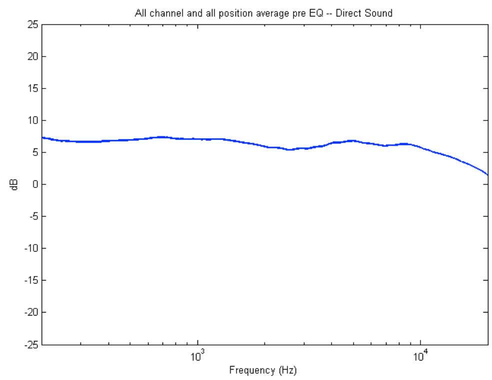


**Fig. 19 Early arrival (5 ms) frequency response, pre-equalization. dB (arbitrary 0) vs. Frequency in Hz. Note differing scale range compared to Figs. 16–18.**

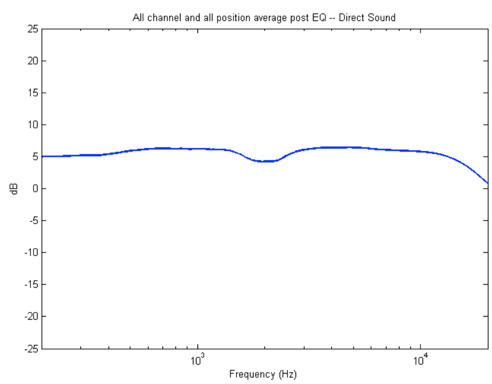


**Fig. 20 Early arrival (5 ms) frequency response, post equalization. Selected for this analysis were systems employing the deliberate mid-range dip described above. dB (arbitrary 0) vs. Frequency in Hz. Note differing scale range compared to Figs. 16–18.**

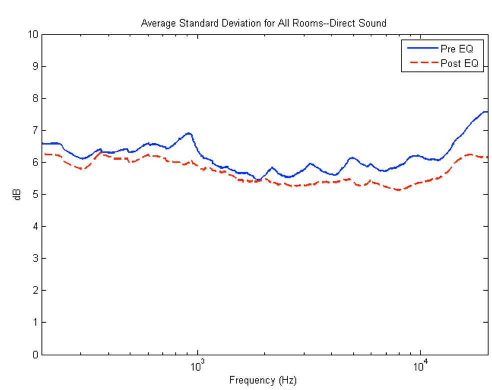


**Fig. 21 The average standard deviation of the early arrival sound (5 ms) in both unequalized (solid line) and equalized (dotted line) configurations. dB (arbitrary 0) vs. Frequency in Hz. Note differing scale range compared to Figs. 16–18.**

## 5.3. Subwoofers

Pre- and post-eq subwoofer responses are given in Figs. 22–24, parallel with the data presented above. Post equalization shows a broader, flatter response on average. The low-frequency high-pass characteristic occurs because the equalization algorithm is trained to not boost at the very lowest frequencies in order to avoid over excursion.

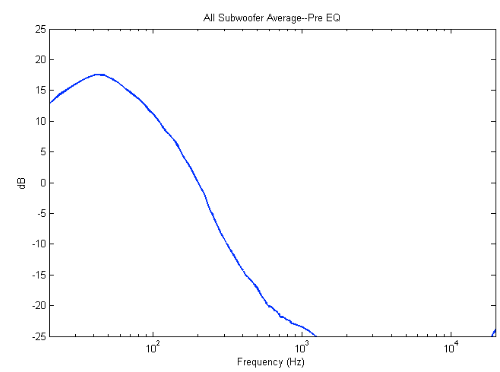


**Fig. 22 Unequalized average subwoofer response, dB (arbitrary 0) vs. Frequency in Hz.**

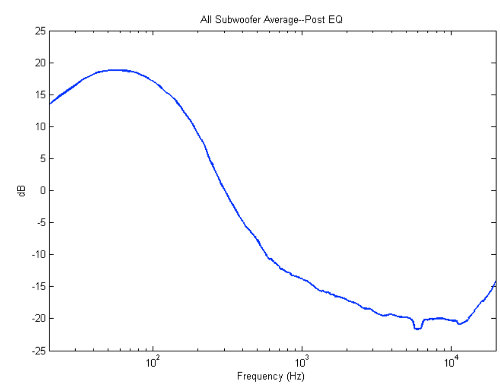


**Fig. 23 Equalized average subwoofer response, dB (arbitrary 0) vs. Frequency in Hz.**

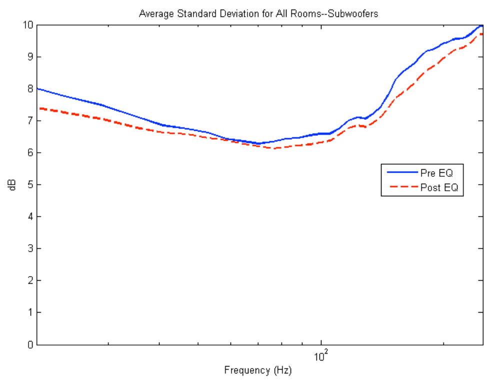


**Fig. 24 Average standard deviation from the average for the unequalized (solid line) and equalized (dashed line) conditions. dB re Frequency in Hz.**

## 6. CONCLUSIONS

Some conclusions have been drawn throughout the text, and will be repeated very briefly here, while others are newly drawn. All apply to the subject of the survey: U.S. home listening room for multichannel sound.

The room volume assumption of 3,000 cu. ft. previously made was a good one, to a first order.

The average reverberation time[9] is about 0.4 s in the mid-range and is flatter with frequency than expected. While any one room can deviate rather widely from this data, the standard deviation was small. For the average room size this reverberation time is longer by 30% than one standard for professional rooms, albeit 25% shorter than living rooms in a large-scale survey of, on the average, smaller rooms built of heavier construction, in Spain. Furnishings seem to be a main contributor to reverberation time.

The average background noise reported by others was supplemented with one studio measurement that showed a caution: the studio was noisier than the homes in which the program material will be heard, and this is thought not to be uncommon.

The three cases of rooms reported in detail and selected at random were greatly helped with equalization, such that it is likely that dialogue intelligibility and timbre were greatly improved, at all seats. At the cost of people "liking bass" and thus opting for system frequency response emphasizing the bass, other long-term improvements are to be had with a more neutral sound system. Such a bass function is better left to a program tone control or equalization function that is switchable, so that correct response and corresponding intelligibility is available. Also, since professional monitoring is usually at higher levels than home playback (by about 8 dB for primary home theater use on the average in one internet survey we conducted), the loudness effect applies, and additional measures should be taken for that. [40]

Both steady-state and direct arrival (5 ms) responses have been given for large-scale averages of 275 rooms. Averaging the data over so many rooms loses detail but shows that the average system has an emphasis in the mid-bass in the unequalized condition, and a mid-range dip.

Equalizing to a target curve with a high-frequency rolloff and a deliberate mid-range dip (selecting for those target curves in the database) results in the average response being indistinguishable from the desired target.

The first 5 ms of data post equalization shows a somewhat brighter high-frequency response than the steady-state response, as might be expected as it contains more direct sound.

The standard deviation of the response is lessened with equalization, although not as much as was expected by theory.

[9] But classical diffuse-field reverberation is actually dominated by modal ring-down time in many bands in small rooms—the two are indistinguishable in these lumped measurements, although the modal time lengthening can be identified in waterfall plots.

## 7. FURTHER WORK

The tip of the iceberg has just been scratched in this work. Among other things, the salient question is how to perform data reduction so that broad conclusions can be drawn without losing the inherent detail in each room's characteristics. Even this work shows very smooth curves from underlying data that is not nearly so smooth: it is the deviations from the average that are more interesting in many ways than the average itself, which is known to "hit" a target curve as that is the performance of the equalization system.

Why is the loudspeaker/room system 680 much worse than 583 pre-equalization at uniformity of coverage? Presumably it is the performance of the loudspeaker model in use, and this could be examined by correlating anechoic chamber measurements with these in-room ones. But what is the role of the room in this? Are the apparently periodic reflections leading to the comb filter behavior in the midrange, surpassing any room mode considerations it appears, the product of the room and loudspeaker, and in what proportion?

How do we data reduce waterfall displays to something meaningful, both psychoacoustically and across rooms?

What is the role of directivity and its change with frequency when factoring early reflections and reverberation? What is its proper role in mono, two-channel stereo, and multichannel reproduction?

Certain aspects are blind to us because of the large-scale nature of this database. Exact room plans, exact microphone placement, sound intensity amplitude and direction at the microphone, are all beyond the scope of what can be captured at this point in time. On the other hand, the work undertaken has been helpful to us in understanding useful installations, not computer models. Work such as that done in [35] can be revised with better-known parameters, for instance.

## 8. ACKNOWLEDGEMENTS

All the home theater installers participating in the location work in more than a thousand home theaters are each thanked for contributions to the database that underlies this work. Without them the scale of this survey would be extremely limited to those we could perform ourselves.

This work was carried out largely by two people who at the time of the work were interns at Audyssey Laboratories, over a period of more than two years. Nathan Dahlin began the initial database analysis that was continued by Ryan Green. Nate provided vital fundamentals required for the progress of the work.

Professor Chris Kyriakakis and Jeff Clark provided valuable analysis help and collegial discussions.

## 9. REFERENCES

[1] S. Bharitkar and C. Kyriakakis, "Visualization of Multiple Listener Room Acoustic Equalization with Sammon Map," IEEE Trans. on Audio, Speech, and Language Processing, vol. 15(2), pp. 542--551, Feb. 2007.

[2] S. Bharitkar, P. Hilmes, and C. Kyriakakis, "Robustness of Spatial Average Equalization: A Statistical Reverberation Model Approach," in J. Acoust. Soc. Amer., 116(6), pp. 3491-3497, Dec. 2004.

[3] S. Bharitkar, C. Kyriakakis, and T. Holman, "Time-alignment of Multiway Speakers with Group-delay Equalization," 2008 Proc. Audio Eng. Soc. Conv., Amsterdam, Netherlands, May 2008.

[4] S. Bharitkar, C. Kyriakakis, and T. Holman, "Variable-octave Complex Smoothing for Loudspeaker-room Response Equalization," Proc. IEEE Intl. Conf. Consumer Elect., Las Vegas, NV, Jan. 2008.

[5] S. Bharitkar, "Variable-octave Complex Smoothing," 2007 Proc. Audio Eng. Soc. Conv., New York, NY, Oct. 2007.

[6] S. Bharitkar and C. Kyriakakis, "Loudspeaker and Room Response Modeling with Psychoacoustic Warping, Linear Prediction, and Parametric Filters," in Proc. 121st AES Conv., San Francisco, Oct. 2006.

[7] S.Bharitkar, Y. Zhang, and C. Kyriakakis, "Room Acoustic Response Modeling and Equalization with Linear Predictive Coding

and Parametric Filters for Speech and Audio Enhancement," in Proc. 40th IEEE Asilomar Conf. on Signals, Systems, & Computers, Pacific Grove (CA), Oct. 2006.

[8] S. Bharitkar and C. Kyriakakis, "Cascaded FIR Filters for Multiple Listener Low Frequency Room Acoustic Equalization," in 2006 IEEE Intl. Conf. Acoust., Speech, \& Signal Proc. (ICASSP-2006), Toulouse, France, May 2006.

[9] S. Bharitkar and C. Kyriakakis, "Comparison Between Time Delay Based and Nonuniform Phase Based Equalization for Multichannel Loudspeaker-Room Responses," Proc. of 119th Audio Eng. Soc. Conv., Oct. 7-12, 2005. New York, NY.

[10] S. Bharitkar and C. Kyriakakis, "Objective Function for Automatic Multiposition Equalization and Bass Management Filter Selection," Proc. of 119th Audio Eng. Soc. Conv., Oct. 7-12, 2005. New York (NY).

[11] S. Bharitkar and C. Kyriakakis, "Optimization of Crossover Frequency and Crossover Region Response for Multichannel Acoustic Applications," {Proc. of 13th European Signal Proc. Conf} (EUSIPCO 2005). Antalya, Turkey, Sept. 4-9, 2005

[12] S. Bharitkar and C. Kyriakakis, "Optimization of the Bass Management Filter Parameters for Multichannel Audio Applications," Proc. 2005 39th IEEE Asilomar Conference on Signals, Systems, & Computers}. Pacific Grove (CA). Nov. 2005

[13] S. Bharitkar, "Phase Equalization for Multi-channel Loudspeaker-Room Responses," {\em Proc. 117th Audio Eng. Soc. Conv}. San Francisco, Oct. 2004

[14] S. Bharitkar and C. Kyriakakis, "Multirate Signal Processing for Multiple Listener Low Frequency Room Acoustic Equalization," Proc. 2004 38th IEEE Asilomar Conference on Signals, Systems, & Computers. Pacific Grove (CA). Nov. 2004.

[15] S. Bharitkar, "An Alternative Design for Multichannel And Multiple Listener Room Acoustic Equalization," Proc. 2004 38th IEEE Asilomar Conference on Signals, Systems, & Computers}. Pacific Grove (CA). Nov. 2004.

[16] S. Bharitkar and C. Kyriakakis, "The Influence of Reverberation on Multi-channel Equalization: An Experimental Comparison Between Methods," Proc. 37$^{th}$ IEEE Asilomar Conference on Signals, Systems, & Computers}, Pacific Grove, CA. Nov. 2003. (Best Student Paper Award

[17] S. Bharitkar and C. Kyriakakis, "A Comparison Between Multi-channel Audio Equalization Filters Using Warping," Proc. 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2003)}, New Paltz, NY. Oct. 2003

[18] S. Bharitkar and C. Kyriakakis, "Sensitivity of multichannel room equalization to listener position," Proc. IEEE Conf. on Multimedia & Expo} (ICME), Baltimore. July 2003

[19] S. Bharitkar and C. Kyriakakis, "Pattern recognition based perceptual equalization of room responses," Proc. IEEE 36th Asilomar Conference on Signals, Systems, & Computers}, Pacific Grove, Nov. 2002.

[20] S. Bharitkar, P. Hilmes, and C. Kyriakakis, "Robustness of spatial averaging equalization: A statistical approach," Proc. IEEE 36$^{th}$ Asilomar Conf. on Signals, Systems, and Computers}, Pacific Grove, Oct. 2002.

[21] S. Bharitkar, P. Hilmes, and C. Kyriakakis, "Robustness of magnitude response average equalization using modal method," Proc. Audio Eng. Soc. 113th Conv., Los Angeles, Sept. 2002

[22] S. Bharitkar and C. Kyriakakis, "Visualization of multiple listener room response equalization using sammon mapping," Proc. IEEE Conf. on Multimedia & Expo} (ICME), Lausanne, Switzerland. Aug. 2002

[23] S. Bharitkar and C. Kyriakakis, "New Factors in Room Equalization Using a Fuzzy Logic

Approach," in Proc. Audio Eng. Soc. 111th Convention}, New York, Dec. 2001.

[24] B&K Microphone Handbook, vol. 1, retrieved on 9/1/10 from http://www.bksv.com/doc/be1447.pdf (requires login).

[25] N. Kyouno, Y Ooshuga, N. Yahshima, and S. Himori, "Digital Filter Application for an In-Phase and Linear-Phase Loudspeaker System: Realization of a Two-Way Loudspeaker System," AES preprint 3240 presented at the 91st AES Convention, Oct., 1991.

[26] Brian C. J. Moore and Brian R. Glasberg, "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.,* 74 (3), Sept. 1983, 750–753..

[27] T. Holman, "Motion-Picture Theater Sound System Performance: New Studies of the B-Chain," *SMPTE Journal,* March, 1994.

[28] T. Holman, "Cinema Sound Quality Redux," Conference Proceedings of the 2005 SMPTE New York Conference.

[29] "Goodbye to TV Ghosts," IEEE Spectrum, July 1992, 50–52.

[30] C. Diaz and A. Pedrero "The reverberation time of furnished rooms in dwellings," Applied Acoustics, vol. 66, no. 8 (Aug., 2005), pps. 945–956.

[31] Paralleled in ITU by ITU-R Recommendation BS.1116-1 (1997): Methods for the subjective assessment of small impairments in audio systems including multichannel sound.

[32] M. A Burgess and W. A. Uteley, "Reverberation Time in British Living Rooms," Applied Acoustics, vol. 18, 1985, pps. 369–380.

[33] L. Fielder and E. Cohen, "Determining Noise Criteria for Recording Environments," JAES vol. 40, no. 5, May, 1992, pps. 384–402.

[34] W. Kuhl and R. Plantz, "The Significance of the Diffuse Sound Radiated from Loudspeakers for the Subjective 'Hearing Event'", Acustica, vol. 40, 1978, pps. 182–191. Translated from the original German by Karen Thayer.

[35] T. Holman, "Scaling the Experience," 12th International Conference of the AES Proceedings, 1993.

[36] T. Holman and N. Zacharov, "Comments on 'Subjective Appraisal of Loudspeaker Directivity for Multichannel Reproduction' and Author's Reply," JAES vol. 48, no. 4, pps. 314–321.

[37] BBC Research Report 1964 L-049/4.

[38] G. Leembruggen, et. al., "Further Investigations into Improving STI's Recognition of the Effects of Poor Frequency Response on Subjective Intelligibility," AES Preprint 8051, 128th Convention, May 2010.

[39] Due to David Lubman, retrieved 9/4/2010 from http://www.natscience.com/Uwe/Forum.aspx/acoustics/320/Schroeder-frequency-counterintuitive

[40] T. Holman and F. Kampmann, "Loudness Compensation: Use and Abuse," JAES, vol. 26, no. 7/8, Aug. 1978, 526–536.

[41] J. S. Bradley, "Sound Absorption of Gypsum Board Cavity Walls," JAES vol. 45, no. 4, 1997 April, 253–259.

[42] A. Ikeda et. al., "The acoustical effect of audio equipment and furniture," Acoust. Sci. & Tech., vol. 26, no. 2, 2005, 233–236.