A New Laboratory for Evaluating Multichannel Audio Components and Systems

SEAN E. OLIVE, AES Fellow, **BRIAN CASTRO**, AES Member, **AND FLOYD E. TOOLE**, AES Fellow

R&D Group, Harman International Industries Inc., 8500 Balboa Blvd., Northridge, CA, 91329, USA

Email: Solive@harman.com

ABSTRACT

The design criteria, features and acoustic measurements of a new listening laboratory designed specifically for listening tests on multichannel loudspeakers and components are described. Among its features is a novel automated speaker shuffler that eliminates loudspeaker position effects or allows the variable to be efficiently tested. Other features include complete computer control of experimental design, control and collection of listener data, making listening tests more reliable and efficient.

1.0 INTRODUCTION

Listening tests are the final arbiter for determining whether an audio product sounds good, and they play a critical role in the research and development of new products. Designing and conducting listening tests that produce reliable and accurate data is, however, no simple task. There are many variables other than those under test that unless removed or controlled can seriously bias the results [1-9]. Two of the more difficult variables to control are the listening room [5],[7],[9] and the position(s) of the loudspeakers under test [5],[9] both of which can significantly influence the sounds that arrive at listeners' ears and listeners' perceptions of them.

Recently we had the opportunity to design and construct a new state-of-the-art listening laboratory to be used for developing and subjectively testing multichannel loudspeakers and other components. The goal from the outset was to build and equip a listening laboratory that could generate subjective measurements as accurate, efficient and free of bias as possible. To meet these goals, a large effort went into developing hardware and software that would automate the design and control of experiments, including the collection, storage and statistical analysis of listener data. Included in the design is a novel automated speaker shuffler that performs positional substitution of 9 loudspeakers so that positional biases can be eliminated or efficiently tested. By eliminating position as a variable, the speaker shuffler has reduced the length of a typical multiple loudspeaker listening test by a factor of 24:1 making product development faster and less costly. Another notable feature of the room is that the acoustics can be easily varied from almost hemi-anechoic to semi-reverberant by adding removable reflective panels to the walls and ceiling.

This paper describes the rationale, features and measurements of the new listening facility, which we call the Multichannel Listening Laboratory (MLL). Finally, the results are compared with several current international standards that recommend performance criteria for listening rooms intended for critical listening.

1.1 Listening Room Standards

Several standards recommend values for various acoustic parameters that define listening room performance. The goal of these standards is to facilitate the replication of listening evaluations in different rooms under the same test conditions. This is particularly important for radio and television broadcast corporations, audio production facilities, large audio equipment manufacturers, and international standards and research organizations, all of whom have multiple facilities in which critical judgments are made on the same program material or equipment. Ideally, if the listening rooms and test conditions in which these judgments are made are sufficiently similar, and the listeners have normal hearing are properly trained, then a consensus in opinion should be possible. If not, then there is likely something wrong with the test procedure itself.

In reviewing these various standards, a serious problem common to many is that while they define tolerances for specific acoustic parameters, they do not adequately define how the parameter is to be measured. For example, IEC is the only standard that specifies how reverberation time should be measured, even though it has been shown that RT60 can vary widely depending on the technique used. Unfortunately, this rather defeats the purpose of defining a standard in the first place! It is conceivable that one measurement method may show the room meets the standard, while another measurement method may not. Added to this is the belief, held by some authorities, that in small rooms, reverberation time is a parameter of little or no value.

A very good discussion and summary of standards as they relate to the design of multichannel listening room intended for loudspeaker listening tests are given by Jarvinen et al in [9].

The current standards that recommend listening room performance include:

1. IEC Publication 268-13: Sound System Equipment, part 13. Listening Tests on Loudspeakers (1985) [10]
2. NR-12 A, Technical Recommendation: Sound Control Rooms and Listening Rooms. 2nd Edition, The Nordic Public Broadcasting Corporation, (1992) [11]
3. ITU-R Recommendation BS.1116: Methods for Subjective Evaluation of Small Impairments in audio systems including multichannel sound systems, 2nd Edition (1997) [12]
4. ITU-R Recommendation BS.775: Multichannel stereophonic sound system with and without accompanying picture (1994) [13]
5. EBU Tech 3276 (2nd Edition, 1997) [14]
6. AES20-1996: Recommended Practice for Professional Audio – Subjective Evaluation of Loudspeakers (1996) [15]

The standards can be classified according to the intended application of the listening room and can be generally classified into two groups. The AES and IEC standards were intended for monophonic and stereophonic testing of loudspeakers in typical domestic listening rooms. Both these standards are now quite old and the recommended room sizes are too small to allow multiple comparison of multichannel systems.

The EBU, ITU and NR standards were drafted primarily by broadcasters and allow for much larger control rooms that can accommodate several listeners at a time. Only the AES, IEC and ITU standards include recommendations for listening test methodology.

At the design stage, we did not intentionally set out to meet any of the above standards. However, in post-hoc examination have found that our listening room meets both ITU and EBU standards in its current configuration in which we have added reflective and diffractive surfaces to both the ceiling and walls.

In the following sections we show measurements made in the MLL and compare these with various acoustical properties recommended in the above standards. These properties include dimensions, floor area, volume, proportions, reverberation time and background noise. The values measured for the MLL are compared with the recommended values in Table 1 for each standard, and shows that the MLL meets both ITU and EBU recommendations.

2.0 MULTICHANNEL LISTENING LAB (MLL)

2.1 Room Dimensions

The listening room itself consists of double-wall constructed shell built by Industrial Acoustics Corporation (IAC). The dimensions of the MLL were largely dictated by our requirements to be able to evaluate up to 3 different 5.1 or 7.1 channel systems at a time and accommodate 1-6 listeners. The room also had to be sufficiently large to accommodate our automated 9-loudspeaker shuffler that requires a space of approximately 9 m (L) x 1.5 m (W) x 1 m (D). This resulted the following dimensions:

MLL Dimensions	
Length	9.14 m
Width	6.58 m
Height	2.59 m
Floor Area	60.20 m^2
Volume	155.92 m^3

As shown in Table 1, the MLL satisfies the recommended volume and floor area values specified in ITU, EBU, and N12 standards. The MLL's volume exceeds the IEC and AES recommended limits of 110 m^3 and 120 m^3 respectively because the standards were intended for small domestic stereo listening rooms.

2.2 Room Proportions

The most problematic performance issue in small listening rooms is non-uniform low frequency reproduction caused by standing waves that produce large pressure peaks and nulls in the lower 3-4 octaves of the audio range. The distribution and frequencies at which these peaks and notches occur are directly related to the geometry of the room. If the ratio of the room dimensions is carefully chosen, a more uniform response is possible. Walker from the BBC [16] has created a room geometry criterion that has been adopted by both the EBU and ITU standards. The "Walker" criterion defines the limits of the ratios for length (*l*), width (*w*) and height (*h*) as:

$$1.1\frac{w}{h} \leq \frac{l}{h} \leq 4.5\frac{w}{h} - 4 \qquad \textbf{(1)}$$

As shown in Table 1, the ratio of dimensions for the MLL meet the "Walker" criterion and therefore satisfies the EBU and ITU standards. The relatively large size of the MLL also benefit uniform frequency response in the lower octaves since the first order width and length modes are below 25 Hz.

2.3 Background Noise

Accurate and repeatable subjective measurements require a listening room with low background noise so those listeners are able to reliably judge the quality of low-level signals. Perception of timbre, nonlinear distortion, loudness and spatial qualities are all influenced by the presence and masking effects of background noise.

Minimizing background noise in the MLL was carefully considered during the design and construction. The IAC double-wall shell itself is located in a large room that has limited access to both people and noisy equipment. No part of the shell touches the structural walls of the building except the floor, which is mechanically floated.

The inner walls and ceiling of the double-wall IAC shell are made of heavy gauge steel panels separated 10 cm and filled with fiberglass. The inner surfaces are perforated with 2.34-mm openings to provide substantial sound absorption inside the room. The inner walls are entirely floated and separated from the outer wall of the shell by a 10 cm space to minimize mechanical and acoustic transmission of noise.

The room has its own dedicated HVAC system with ventilation silencers and acoustically lined ducts that create a comfortable and quiet environment. For experiments that require extremely low background noise the room can be cooled and the HVAC can be completely shut off during the test. The room requires minimal lighting during the test itself (i.e. 1 Halogen light) which means that noise from lights is not an issue. All audio equipment, other than the required amplifiers, is located outside the room, and this also helps to minimize electrical noise as well.

In an effort to simulate the construction of floors found in many homes, a carpeted “squeak-free” plywood floor was laid on 5 cm x 15 cm wooden joists separated 41 cm apart. The joists are mounted on 6.4 mm neoprene pads for isolation from the concrete floor beneath. The rationale for constructing this floor is to allow transmission of low bass from the loudspeaker through the floor to the listeners’ feet, since the perception of bass depends on what is felt, as well as what is heard. The front and middle sections of the floor can be removed to allow easy access of audio, video and data cables that run underneath the floor to access panels both inside and outside the room.

In reviewing the various listening room standards there is a wide range of recommended levels for background noise. The most stringent requirements are specified by the EBU and ITU standards, which call for minimum level of NR10, not exceeding NR15. These rather demanding requirements are likely justified in broadcast environments where listeners are frequently required to evaluate small signal linearity, for example in relation to CODECS.

At the other extreme, the AES and IEC standards both have rather liberal recommended background noise limit of 35 dBA measured using a slow time constant. The AES standard has an additional limit of 50 dB C-weighted for low frequency noise. The less stringent requirements are likely justified on the basis that they are aimed at loudspeaker evaluations in typical domestic environments where background noise levels are typically higher.

Figure 1 shows the background noise measured in the MLL with the air conditioner turned both on and off. Also plotted are the NR curves 0 through 15. The MLL noise curves each represent an average of four measurements take at 4 different locations around the listening area. The time over which each measurement was averaged was 64 s. The measurement was taken using a Bruel & Kjaer 4179, 1 inch microphone, a Bruel & Kjaer preamp Type 2660, and a Bruel & Kjaer real-time analyzer. The low noise microphone and preamp allow accurate measurement of sound pressure levels below the threshold of hearing, which is necessary at higher frequencies for measuring rooms below NR20. Figure 1 shows that with the air conditioning turned off, the MLL meets NR5, thus meeting the requirements of the EBU and ITU specification. With the air conditioning turned on the noise increases to NR15.

2.4 Reverberation Time

The reflected sounds and reverberation time in a room have been shown to have an important influence on the perception of loudness, timbre and spatial qualities and speech intelligibility in both live and reproduced sound. While this is a complex phenomena, the acoustic community sees fit to summarize it all in a T_{60} measurement.

Both the EBU and the ITU standards specify values for the average reverberation time in the room. ITU and EBU recommend the value (within a tolerance of ± 0.05 s) be determined using the following equation:

$$T_m = 0.25\left(\frac{V}{V_{ref}}\right)^{1/3} s \qquad \textbf{(2)}$$

where T_m is the average reverberation time between 200 Hz to 4 kHz, V is the volume of the room, and V_{ref} is the reference volume of 100 ^{3} m. The EBU also put limits on the range of values specifying that the value should lie between $0.2 < T_m < 0.4$ s.

The IEC standard specifies a T_m of 0.3 – 0.6 seconds which is very similar to the AES standard that recommends 0.45 s (± 0.05 s). The N 12-A standard specifies T_m be measured in 1/3 octaves between 200 Hz to 2.5 kHz and be determined as a function of the floor area using the following equation:

$$T_m = 0.35\left(\frac{S}{S_{ref}}\right) \pm 0.05\,\mathrm{s} \qquad \textbf{(3)}$$

where S is the floor area of the room and S_{ref} is the reference area of 60 ^{2} m.

In addition to specifying the average reverberation time, most of the standards recommend that T_m be relatively independent of frequency within a certain bandwidth and tolerance. For ITU and EBU standards, the T_m value for each octave band between 200 Hz - 3.5 kHz should vary no more than ± 0.05 s from the calculated optimum value. Below 200 Hz, T_m is allowed to increase monotonically with frequency to 0.3 s above the optimum value. Above 3.5 kHz, the tolerance is increased to ± 0.1 s from the optimal value.

By substituting the volume of the MLL (155.92 ^{3} m) into equation (2), we calculate that T_m should be 0.29 s to meet ITU and EBU standards. According to N 12-A, the T_m for the MLL should be 0.35 s.

The T_{m} of the MLL was measured using a MLSSA system from DRA laboratory. The microphone was a Bruel & Kjaer 4134 microphone. The sound source consisted of four JBL Synthesis satellite loudspeakers crossed at 80 Hz over to a JBL Synthesis Two subwoofer located in the corner of the room. Each of the four satellites was located approximately 2 m apart and aimed at a different corner in an attempt to create a diffuse sound field. The measurement shown in Figure 2 represents a spatial average of four microphone locations. The average T_m value for the MLL is about 0.23 s, which is slightly below the calculated ITU and EBU optimal value of 0.29 s. However, the curve falls within the minimum recommended value, and is quite uniform with frequency, only rising slightly below 125 Hz.

2.5 Control of Early Reflections

With the advent of 5.1 and 7.1 multichannel and 3D audio playback systems, there is a trend among professional and home theater listening room designs towards lower reverberation times and the control of early reflections. There are sound scientific reasons for doing this, since strong early reflections are known to influence the perceived spatial and timbral qualities of reproduced sound [7], [17]. In the new generation of multichannel recordings and video disks, the additional center and surround channels allow the producer and recording artist to create much more realistic and spatially-enriched environments than ever before. There is less need to use the room's boundaries and the loudspeakers' directional characteristics to compensate for the obvious spatial deficiencies inherent to stereo.

The EBU standard recommends that all reflections within the first 15 ms after the arrival of sound be no greater than 10 dB in level relative to the direct sound from each sound source. With multichannel setups the early sound field is rather complex given that there are between 5-7 loudspeakers and several boundaries. For example with 5 loudspeakers and 6 boundaries there are 30 first order reflections and 150 second order reflections. Measuring and separating out these reflections is no trivial task. The reflections from the floor are particularly problematic to treat since in most facilities, the floor surfaces must be hard and reflective to facilitate the movement of people and equipment. Nonetheless, several organizations [18], [19] are building such rooms that meet this reflection-free part of the specification with the exception of the floor bounce.

In the MLL room, the only significant first order reflections are from the floor, and these are attenuated at higher frequencies by the carpet. At listener-loudspeaker distances greater than 2 m any reflection with a path length greater than 6.34 m will be attenuated 10 dB by spreading loss [18]. This effectively eliminates all second order reflections since their path length exceeds this value. For front channel sources, first order reflections from the side walls will also be sufficiently delayed beyond the 15 ms time gap. The main culprits are reflections from the front and back walls, and the ceiling. Fortunately these surfaces can be made absorptive by simply removing the reflective panels so that the absorptive surface is exposed. To reduce flutter echoes from reflective surfaces and to increase reverberation, 120 RPG Skylines, an omnidirectional primitive root number theory 2D diffusor, are placed on the reflective panels located on the walls, as well as on the ceiling and areas behind the loudspeaker as shown in Figures 4 and 5. These light-weight diffusors are easily removed or relocated, and help reduce any other specular reflections that may arrive after the direct sound.

2.6 Automated Speaker Mover

The position of a loudspeaker in a room has a significant impact on its perceived sound quality. Changing its position affects the way it couples to the standing wave modes of the room, and alters the physical characteristics of broadband reflections that arrive at the listener. In listening tests that involve multiple comparisons among loudspeakers the positional effects on listeners' ratings can be larger than the differences between the loudspeakers under test [8]. Unless these positional effects are controlled, the results may be contaminated by a nuisance variable.

For multiple comparison loudspeaker tests, asking human beings to sit behind a double-blind screen and quickly and smoothly substitute the positions of 2-9 loudspeakers (some weighing upwards to 100 kg) on command presents an obvious logistical problem. Clearly the problem of positional substitution calls for an automated solution. This realization led to the development of our own custom-built speaker shuffler. Prior to having a speaker shuffler, the positional effects in loudspeaker tests had been balanced by testing each loudspeaker in each position. Any position-related bias would be equally distributed or balanced across each loudspeaker. More scientifically rigorous designs go even further and test all possible loudspeaker-position permutations so that any possible context effects between loudspeaker and position are also balanced.

The disadvantage of not having a speaker mover is that an additional number of trials are required to balance the variable position. This relationship in illustrated in Figures 3(a)-(b), which compare the number of trials required to balance the variable position in multiple comparison tests, with and without a speaker mover. The number of trials is calculated using the following equation:

$$\text{Number of Trials} = \text{N}_{\text{Speaker Positions}}! \times \text{N}_{\text{Programs}} \times \text{N}_{\text{Repeats}} \qquad \textbf{(4)}$$

Where $N_{Speaker\ Positions}$ equals the number of speaker positions in the test, $N_{Programs}$ equals the number of program selections being used and $N_{Repeats}$ is the number of repeats. In Figure 3 we, the experimental design shows no repeats, that is $N_{Repeat} = 1$.

The graphs clearly shows that an automated speaker mover can drastically reduce the length of the experiment because the variable $N_{Speaker\ Positions}$ always equals 1, regardless of how many loudspeakers are compared. In comparing the two graphs we see that there is a 2:1 advantage for paired comparisons, a 6:1 advantage for triple comparisons, and a 24:1 advantage for comparisons among four loudspeakers. When you multiply these ratios by the number of programs and repeats used in the experimental design, the number of trials quickly escalates. For multiple comparisons between four loudspeakers using 4 programs with no repeats, a total of 96 trials are required without a speaker mover. Having a speaker mover reduces the experiment to 4 trials. This enormous difference provided the justification to design and build a custom speaker shuffler, since over the long-term, it could afford considerable savings in person-listening hours and product development time.

A custom-built floor at the front of the room allows us to perform positional substitution of up to 9 different loudspeakers. A photograph of the speaker mover set up for an A/B stereo loudspeaker comparison is shown in Figure 4. Figure 5 shows a photograph of the speaker mover set up for a single comparison of a 5.1 loudspeaker system. For the purposes of the photograph the front, side and rear listening curtains have been retraced out of the way. Each loudspeaker is attached to one of nine pallets that move in 1-inch increments over a range of 4 feet forwards and backwards while the entire array moves 4 feet to the left and right of the listener. The movement of the floor can be controlled manually from a programmable logic controller (PLC), or from a computer that is linked serially to the PLC via RS232. This allows all positions of loudspeakers to be programmed, stored and recalled quickly. The movement of the floor is extremely quiet, repeatable to within 1 inch, and fast. Transit time between positions is no greater than 3 s, and most positional changes are under 2 s. The transit speed is also programmable and can be decreased or increased if desired. As a safety measure, a light fence is installed in front of the moving floor so that if anyone crosses the light beam the speaker mover automatically stops.

The speaker shuffler allows position-controlled loudspeaker comparisons in mono (up to 4 different systems), stereo (4 different systems) or three different left/center/right channel loudspeakers. At this time, positional substitution of surround and rear channel speakers must be done manually for multichannel experiments. The speakers can be placed away from the side and rear boundaries on stands, or placed on adjustable shelves that are mounted on baffles made of high-density board, that slide in a track along the perimeter of the room.

The moving floor gives us an efficient means to eliminate the effects of loudspeaker position, or it can do the reverse, and allow us to test the interaction effects between loudspeaker and position. By statistically-averaging a loudspeaker's performance over a number of different positions we can assess its off-axis performance, and a number of other parameters that are position dependent. All of this becomes essential as we aim to design loudspeakers that are 'room friendly' and develop digital room equalization systems.

Finally, the speaker mover also allows us to efficiently randomize between each trial, how the loudspeaker is identified to the listener (e.g. "A,B,C..). This ensures that listeners' judgments in each trial are statistically independent between program selections. Without a speaker mover, experimenters normally do not move the loudspeakers behind the screen until a complete block of programs has been rated. These are not independent judgments since the listener knows they are rating the same loudspeaker(s) within each block. The extent to which this biases the results has not yet been reported.

2.7 Blind versus Sighted Listening Tests

It is generally accepted among scientists that psychometric experiments must be performed double blind. For audio tests, this means the identities of the components under test cannot be made known to the listener, and the experimenter cannot not directly control or administer the actual test.

In 1996 Toole and Olive in [2] conducted some blind versus sighted loudspeaker tests that showed both experienced and inexperienced listeners' judgments were significantly influenced by factors such as price, brand name, size and cosmetics. In fact, the effect of these biases in the sighted tests were larger than any other significant factors found in the blind tests, including loudspeaker, position and program interactions. These experiments clearly show that an accurate and unbiased measurement of sound quality requires that the tests be done blind.

To remove these biases from listening tests in the MLL an acoustically transparent curtain that is visually opaque is placed between the products and the listeners so that they do not know the identities of the products under test. All other associated equipment in the signal path is also out-of-sight and locked in an equipment rack, since the performance and paranoia of some listeners can be affected by simply having knowledge that a certain brand of interconnect or CD player is in the signal path.

The front screen consists of a black open knit polyester knit cloth chosen for its acoustic transparency and used as grille clothe in many of our loudspeakers. The material is attached to a large automated curtain roller so it can be easily lifted down and up with an infrared remote control. Weights are attached to a seam in the bottom so the cloth retains its tautness when in use. Retractable curtains made of the same material surround the listeners to hide the identities of loudspeakers located at the sides and rear of the listening room. Figures 4 and 5 show the front, side and rear curtains fully retracted when not in use, and Figure 8 shows the curtains in place during an actual listening test.

2.8 Video Playback

Video and audio are increasingly becoming recorded, processed and distributed together. There is a growing interest among researchers in studying how the perceived quality of one affects the perception of the other. Although much research still needs to be done, evidence suggests there are bimodal interactions between the two that influence listeners' expectations and judgments of the quality of the audio, and vice versa. Keeping this in mind, we were careful in selecting a video playback system within our budget that had sufficient quality, so that it would not negatively impact listeners' opinions of the sound quality.

We selected a three gun front projection CRT made by Audio Video Source for its above-average picture quality and the additional advantage that is has no fan. The picture is projected on a 100 inch Stewart Microperf screen that is retractable so it can be removed for audio-only listening tests. The acoustical effect of the screen is another factor that is not completely understood, and will be a subject of investigation.

2.9 Automated Control, Collection and Analysis of Data

In designing the MLL, we wanted to automate as much as possible the design and running of experiments including the collection, storage and analysis of data, in order to reduce the time and costs of performing listening tests. Automation of experiments has the additional benefit of making listening tests more reproducible, largely because it reduces the risk of human errors and biases introduced by the experimenter. Considerable ongoing effort in software development is helping us to fulfill these goals.

Automation begins at the experimental design stage where all important experimental parameters and details are defined by the experimenter as a "*.exp" file that is stored in a database that resides on the Windows NT server.

The experiment file contains the following information:

- The name of the experiment and a brief description
- Detailed information related to the experimental design and protocol including definition of scales and randomization of variables. Protocol choices include single or multiple comparisons, ABX, ABC(with hidden reference) and different threshold measurement protocols.
- Instructions to the listeners
- Equipment control information and operational parameters required by the audio switcher for level matching, switching and overall output level.
- The file names or track information for each program selection. This information is sent to the appropriate signal source device.
- Information related to the position and movement of loudspeakers
- A list of trials which the software randomly selects

The Windows NT server controls the running of the experiment including control of all associated equipment in the signal path. A block diagram of the equipment and signal path for the MLL is shown in Appendix 1. The lines that connect each block as well as the signal paths are color coded and typed according to whether the signals are audio (either analog or digital), video, infrared or RF control, computer data, MIDI control or sent over PCI or serial buss. The signal sources are the blocks on the top left of Appendix 1. They currently include DVD and Laser Disk player, an 8-channel PCM digital recorder, and an 8-channel PC-based hard disk recorder (Lexicon Studio) and its associated A/D and D/A I/O cards. The audio and video outputs of the DVD and LD players are sent to the Lexicon DC-1 which provides AC-3 and DTS decoding when required. The analog outputs are sent to the Spirit 328 digital mixer which provides signal switching and level matching (within 0.03 dB) for up to 16 analog or digital inputs. The 8 channel sources are sent digitally to the Spirit mixer and remain digital up to the power amplifier before they are converted by the Studer D/A's.

All operational parameters of the Spirit mixer can be viewed, stored and recalled from the NT Server via MIDI control.

The input of listener data, feedback and status information is done using laptops connected to the NT Server through a LAN. For single listener experiments, the listener can control switching of the stimulae remotely from their laptop. A photograph of a listener entering data on the laptop connected to the NT Server is shown in Figure 6. For multiple listener experiments, the NT Server controls the switching either manually or through software automation. During the experiment, all changes in listener response data can be viewed in real-time on the NT Server which performs running statistical averages and graphs of the results.

Remote access to the NT Server and control of the equipment from inside the listening room is also possible through a wireless RF mouse, keyboard and a flat panel display, all of which are connected to the Server. This might be required during set up or for informal listening sessions or product demonstrations. The flat panel display also shows status information to the listener(s) indicating what stimulus (i.e. A, B, C…) is currently selected, and any other necessary information.

Finally, all experimental data and information related to listeners (date and time, name, seat position, age) is stored in a relational data base which can be formatted and imported into various statistical packages we use for analysis of results.

Not shown in the block diagram is a video camera used for monitoring subjects and to detect and hopefully deter possible cheating. Also not shown is a two-way intercom that allows communication between the subject(s) and the experimenter.

3.0 CONTROL ROOM AND LISTENER TRAINING LAB

Outside the MLL is a lab area dedicated for audio and test equipment used during the set up, running and monitoring of listening experiments. Here a space is also dedicated for the training of listeners, which is done over headphones at computer audio workstations.

Bech in [20] has shown that 6 trained listeners can provide data that is as statistically reliable as data gathered from 18 untrained listeners. Clearly, considerable cost-savings in time and money can be realized if listeners are trained before they participate in formal listening experiments. At Harman, listeners with normal hearing undergo a listener training program, which self-administered through a computer and custom software developed in-house [21]. The software teaches listeners to identify and rate using different scales, frequency response irregularities according to the center frequency, amplitude and Q of the distortion. The graphical user interface of the training software is shown in Figure 8.

The training focuses on frequency-related problems since these are the common and most serious audible problems found in most loudspeaker-related listening tests, which many untrained listeners find difficult to describe. The training solves this problem by teaching listeners to describe these phenomena in technical terms that design engineers can understand and use to correct any problematic audible artifacts in product designs.

The training software has proved to be a valuable tool for teaching listeners how to describe and scale the various dimensions of sound quality in meaningful terms, and allows their performance to be quantified in terms that allow us to discriminate good listeners from bad ones. An additional, indirect, benefit accrued from training is that we have learned which program selections are most revealing of typical frequency-related artifacts introduced during the training exercises, and we now use these in our product evaluations.

4.0 CONCLUSIONS

In summary, we have described a new facility designed to test multichannel components efficiently and as bias-free as possible. The facility includes acoustically transparent listening screens that hide the identities of all multichannel loudspeakers and equipment within the audio path. Particular attention has been taken to address the two of the most problematic variables in listening tests: the listening room and the position(s) of the loudspeaker. Through the use of a computer automated speaker shuffler, we have greatly reduced the amount of time and effort required to set up and test multiple comparisons between loudspeakers by reducing the factor position to a one-dimension or level variable. Typical loudspeaker evaluations should be reduced in length by a factor of 24:1.

The listening room itself is capable of testing up to three different 5.1 or 7.1 channel systems and accommodate 1-6 listeners at a time. The measurements we have shown in this paper indicate its performance in its current form meets the very highest standards set out by the ITU and EBU recommendations, in terms of volume, geometry, reverberation time, and the control of early reflections. The acoustics of the room can be easily altered from hemi-anechoic to more typical domestic room conditions by adding reflective panels to the room’s boundaries.

Finally, the experimental design, set up and control are computer-automated so that experiments can be easily repeated, and are less prone to human error. The more time-consuming and mundane tasks such as collection and analysis of data have also been computer-automated, so that experiment report writing becomes a simple cut-and-paste operation.

5.0 ACKNOWLEDGEMENTS

The authors would like to thank Tom Roberts of Bruel & Kjaer for his assistance and loan of the equipment used to make the background noise measurements shown in this paper.

6.0 REFERENCES

[1] F.E. Toole, “Listening Tests - Identifying and Controlling the Variables”, Proceedings of the 8th International Conference, Audio Eng., Soc. (1990 May).

[2] F.E. Toole and S.E. Olive, "Hearing is Believing vs. Believing is Hearing: Blind vs. Sighted Listening Tests and Other Interesting Things", 97th Convention, Audio Eng. Soc., Preprint No. 3894 (1994 Nov.)

[3] F.E. Toole, “Listening Tests, Turning Opinion Into Fact”, J. Audio Eng. Soc., vol. 30, pp. 431-445 (1982 June).

[4] F.E. Toole, “Subjective Measurements of Loudspeaker Sound Quality and Listener Performance”, J. Audio Eng. Soc., vol. 33, pp. 2-32 (1985 January/February).

[5] Soren Bech, “ Perception of Timbre of Reproduced Sound in Small Rooms: Influence of Room and Loudspeaker Position J AES, *Vol. 42, Number 12 pp. 999 (1994).*

[6] S.E. Olive, P. Schuck, J. Ryan, S. Sally, M. Bonneville, “The Variability of Loudspeaker Sound Quality Among Four Domestic-Sized Rooms”, presented at the 99th AES Convention, preprint 4092 K-1 (1995 October).

[7] F.E. Toole, “Loudspeakers and Rooms for Stereophonic Sound Reproduction”, Proceedings of the 8th International Conference, Audio Eng., Soc. (1990 May).

[8] S.E. Olive, P. Schuck, S. Sally, M. Bonneville, “The Effects of Loudspeaker Placement on Listener Preference Ratings”, J. Audio Eng. Soc., Vol. 42, pp. 651-669 (1994 September).

[9] Antti Jarvinen, Lauri Savioja, Henrik Moller, Veijo Ikonen, Anssi Ruusuvuori, “Design of a Reference Listening Room – A Case Study”, AES 103rd Convention, New York, Preprint 4559, September 26-29, 1997.

[10] IEC Publication 268-13: Sound System Equipment, part 13. Listening Tests on Loudspeakers (1985)

[11 NR-12 A, Technical Recommendation: Sound Control Rooms and Listening Rooms. 2nd Edition, The Nordic Public Broadcasting Corporation, 1992.

[12] ITU-R Recommendation BS.1116: Methods for Subjective Evaluation of Small Impairments in audio systems including multichannel sound systems, 2nd Edition (1997)

[13] ITU-R Recommendation BS.775: Multichannel stereophonic sound system with and without accompanying picture (1994).

[14] EBU Tech 3276 (2nd Edition, 1997).

[15] AES20-1996: Recommended Practice for Professional Audio – Subjective Evaluation of Loudspeakers (1996).

[16] Walker, R. “Optimum Dimension Ratios For Small Rooms”. 100th AES Convention. Preprint 4191 (Copenhagen, Denmark, 1996).

[17] S.E. Olive and F.E. Toole, “The Detection of Reflections in Typical Rooms”, J. Audio Eng., Soc., vol. 37, pp. 539-553 (1989 July/August).

[18] R.Walker,” A controlled-reflection listening room for multichannel sound”, AES 104th Convention Amsterdam, The Netherlands, Preprint #4645, May 16-19, 1998

[19] E. Arató Borsi, T. Póth, and A. Fürjes," New Reference Listening Room for Two-Channel and Multichannel Stereophonic" AES 104th Convention Amsterdam, The Netherlands, Preprint #4732, May 16-19, 1998.

[20]Soren Bech,"Selection and Training of Subjects for Listening Tests on Sound-Reproducing Equipment" Vol. 40, Number 7 pp. 590 (1992).

[21] S. E. Olive, "A Method for Training of Listeners and Selecting Program Material for Listening Tests", 97th Convention, Audio Eng. Soc., Preprint No. 3893 (1994 November).

TABLE 1

Parameter	Harman MLL	ITU	EBU	N12-A	IEC	AES
Volume (m^3)	155.92				60-110 (80)	50-120
Floor area (m^2)	60.20	20-70	> 40	60 ± 10		> 20
Height *h* (m)	2.59				2.3 - 3.0 rec. 2.8	> 2.1
Length *l* (m)	9.14				≥ 6 rec. 6.7	
Width *w* (m)	6.58				≥ 4 rec. 4.2	
(1.1 w / h)	*2.80*					
(l / h)	*3.53*					
(4.5w / h - 4)	*7.44*					
T_m (s)	0.23	0.29 ± 0.05	0.29	0.35	0.3 -0.6 0.4 ± 0.05	0.45 ± 0.15
T_{63} Hz Max (s)	.34	Tm(s)	0.2 - 0.4	0.35	0.8	
Noise Level	NR 5	NR10; abs. max NR 15	NR10; abs. max NR 15	NR 10 or L_{pA} < 15 dB	L_{pA} < 35 dB	L_{pA} < 35 dB and L_{pC} < 50 dB

Table 1: Dimensions and Acoustic Parameters of Harman MLL versus Recommendations of Various Standards

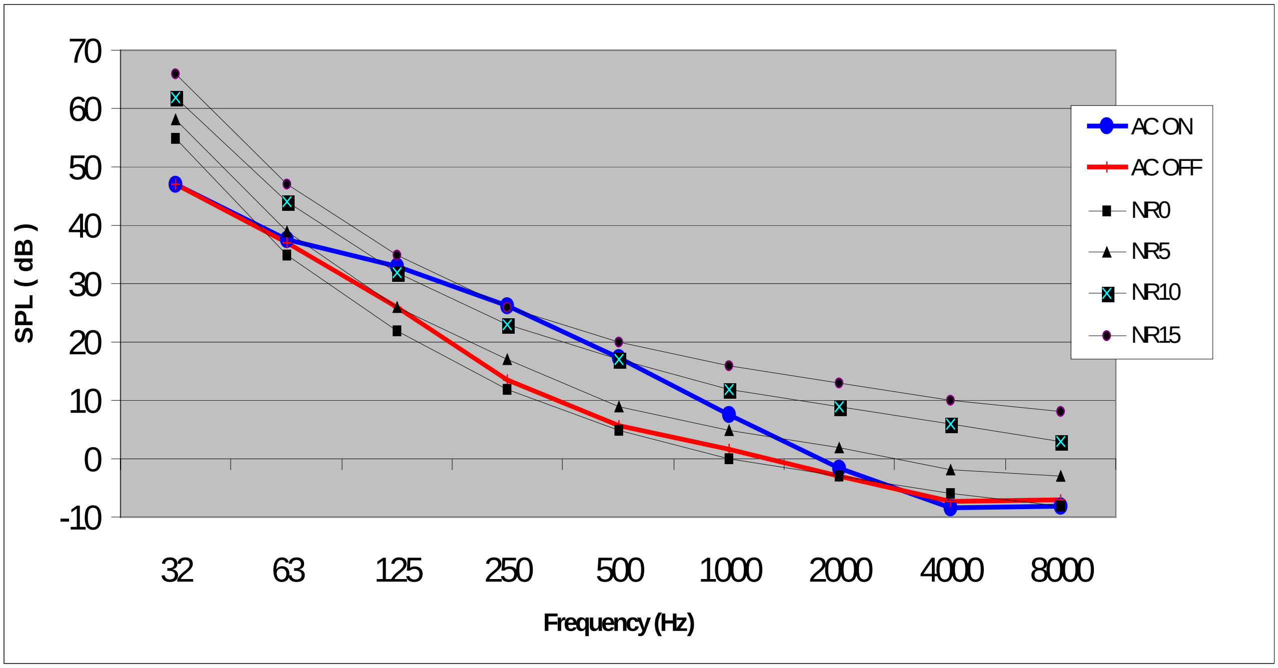

Figure 1 A spatially-averaged measurement showing the background noise in the MLL with the air conditioning off (dotted) and turned on (dashed) compared to the NR curves: 0,5,10 and 15.

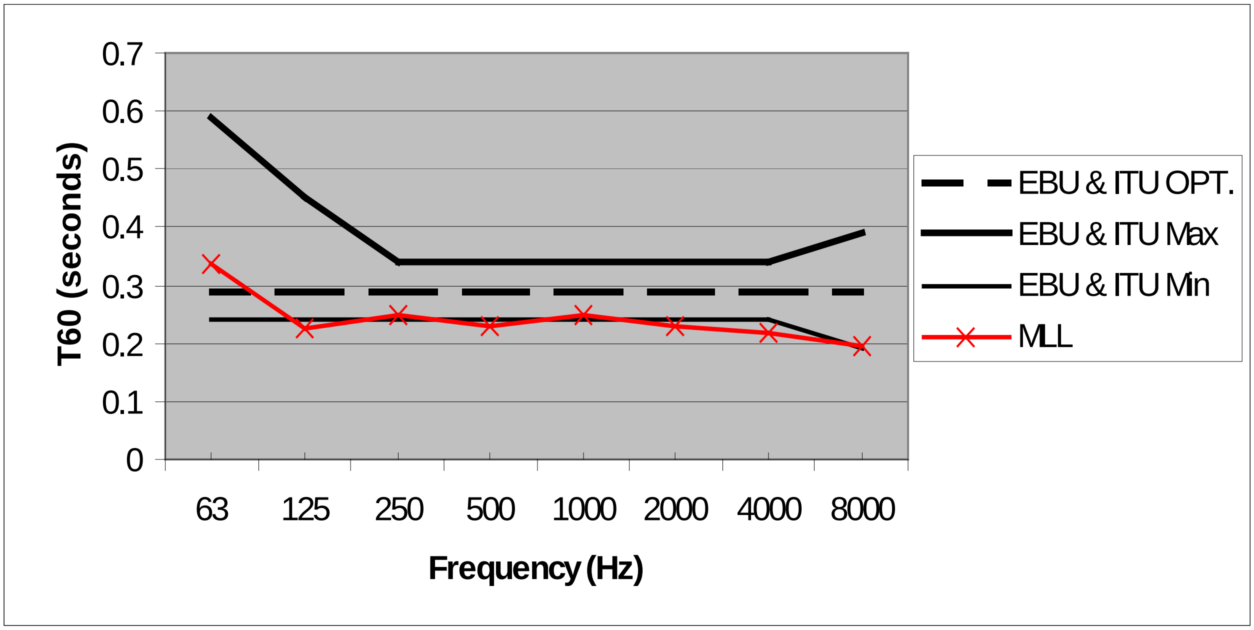

Figure 2 The T_m (RT60) values measured in the MLL compared to the optimal, maximum and minimum values recommended by the EBU and ITU standards.

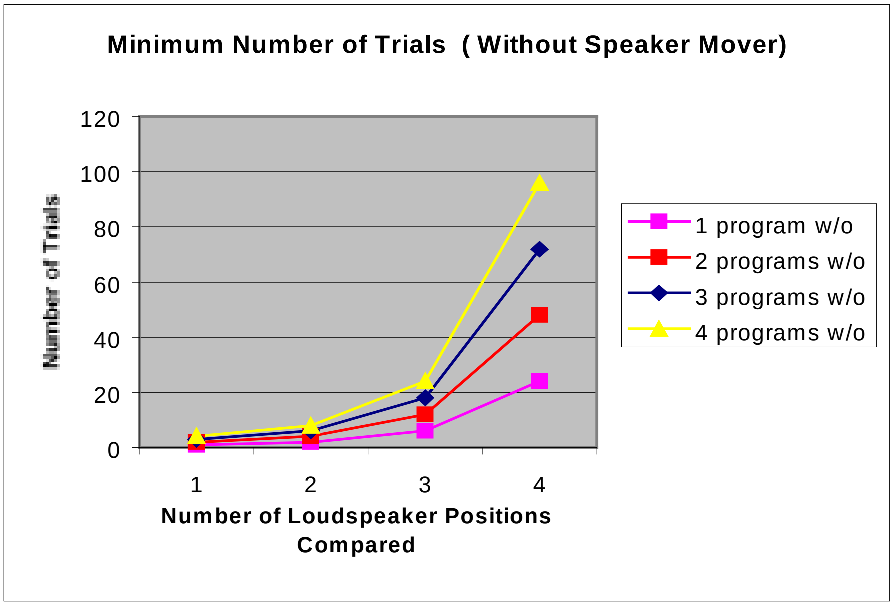

Figure 3(A) The above graph shows the number of trials required for a multiple comparison loudspeaker experiment as a function of the number loudspeaker positions compared. The lines represent experiments in which 1-4 programs are used. The design balances all position and context effects and has no repeats.

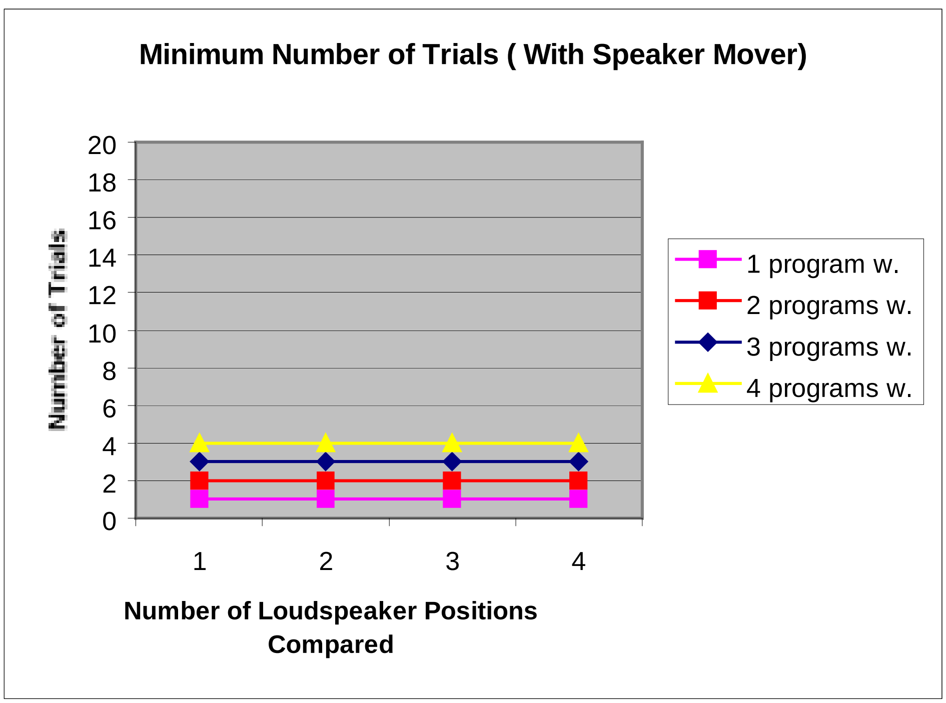

Figure 3(B) The same experiment is shown as in Figure 3(A) above except here a speaker mover is used.

Figure 4 Shown is the automated speaker shuffler of the MLL set up for A/B stereo testing of two stereo loudspeakers. Here the front listening screen is pulled up.

Figure 5 A front-left wide-angle shot of the MLL with the listening screens pulled back. The automated speaker shuffler is in the foreground setup for 5.1 playback. Note the side and rear channel speaker baffles in the background, and the audio and computer data control box on the back wall. The video projector is mounted on the ceiling with a retractable screen in front of the speaker mover.

Figure 6 A listener performing a test by entering their data on a laptop computer that is networked to the NT Server. In this test, video is displayed and both front, side and rear curtains are drawn to hide the identifies of the 5.1 loudspeaker systems under test.

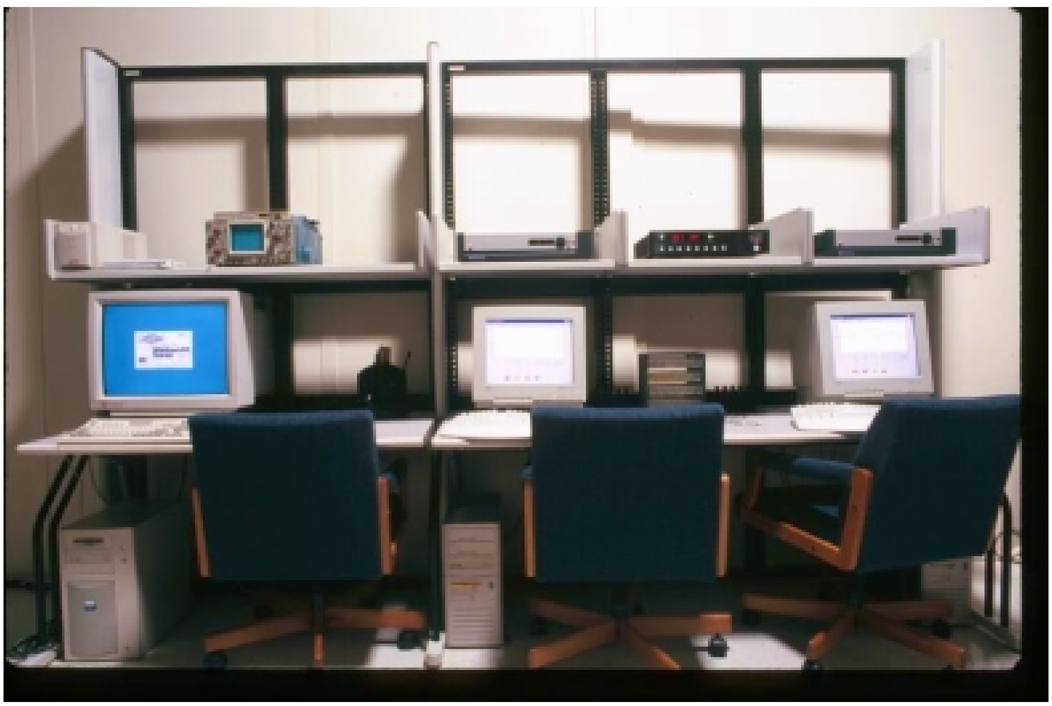

Figure 7 Shown is the control room area outside the listening room where all audio equipment,experimental control and monitoring takes place. Shown here is the NT Server on the left, and two listener training workstations on the right.

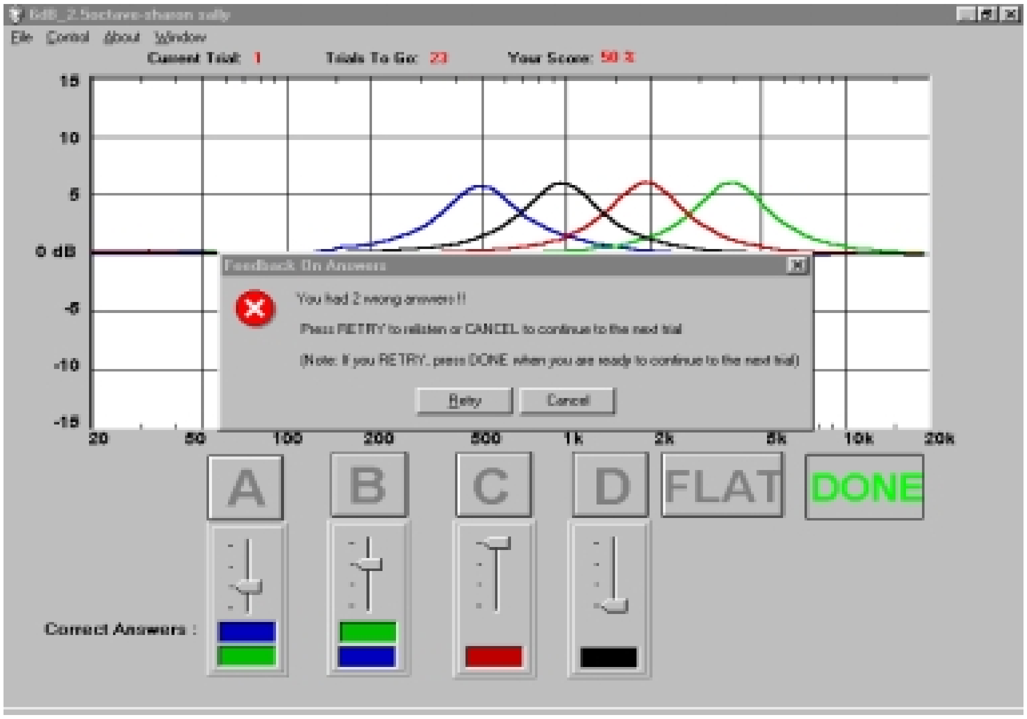

Figure 8: The GUI of the listener training software. The listeners' task is to match the 4 different equalizations indicates by their frequency response curves that are randomly assigned to Buttons A-D Feedback is given on their responses. The "FLAT" button allows listeners to audition the program without any equalization added.

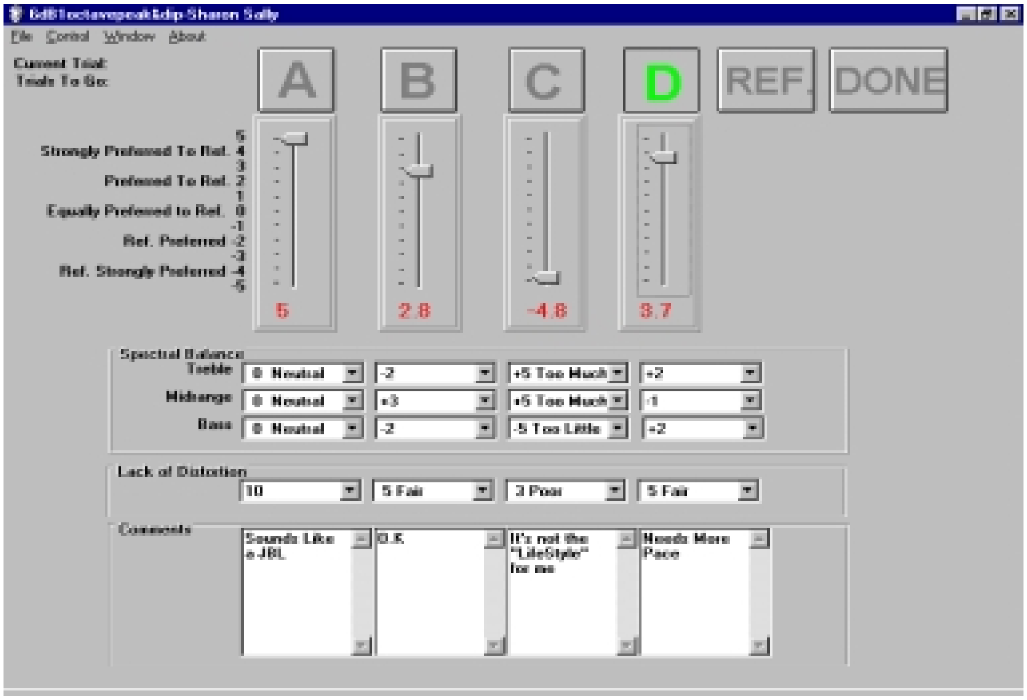

Figure 9: The GUI of the software used for a typical listening test or training exercise. Listeners enter their preference ratings for sounds A-D relative to a given reference ("REF"). Ratings are also given on spectral balance and distortion. Relevant comments are optional.

Appendix 1: Block Diagram of Harman Multichannel Listening Laboratory (MLL) showing key features, equipment and path for audio, video, data and control signals.