Auditory Room Size Perception for Real Rooms

Claudiu B. Pop (1) and Densil Cabrera (2)

(1) Faculty of Architecture, University of Sydney, NSW 2006, Australia
(2) Faculty of Architecture, University of Sydney, NSW 2006, Australia

ABSTRACT

While research on hearing and perception of room acoustics has produced a great deal of information concerning the qualitative auditory sensations imparted on listeners in rooms (reverberance, clarity, etc), little is known about the quantitative information that listeners obtain concerning the rooms encountered in their everyday lives. Information about room size, floor construction, room shape, and many other aspects is transmitted acoustically to listeners in their environments, but it is the sense of vision that is erroneously assumed to provide all of this information to humans. A common listening task which has been left behind in research, however, is the ability of perceiving the size of rooms. In this study, subjective experiments with blindfolded people were conducted to obtain room size ratings. Anechoic speech was reproduced over loudspeakers in three small rooms and the relationships between the room acoustical parameters and the room size judgements were investigated. The results were compared with results from previous studies where modelled and measured rooms were used for subjective testing.

INTRODUCTION

The apparent size of a room, as it is experienced through sound, is a sparsely investigated topic – with only a handful of studies in the literature. However, this topic is becoming increasingly relevant, as computer-based room simulations find application in extended and virtual reality technologies, as well as in architectural acoustics modeling and design. The purpose of this study was to obtain real room data in relation to room acoustical parameters and to provide a potential verification of audio simulation systems.

Background

Humans gather a massive amount of information about their immediate surroundings from the sound reaching their ears. This information is a major supplement to visual and other sensory data in supporting normal personal activity. This process is usually carried out without conscious effort, and most people are unaware of the information received this way.

The information carried by sound includes, for example, the identity of objects in the environment (both sounding and silent), their location, movement and state, properties of the environment (e.g. materials properties, size of room). While there is much knowledge available from previous studies about the acoustical cues for sound source direction, distance and movement, the ability of the listeners to estimate room size from its auditory response is largely an unstudied field. Recent work by Mershon et al. (1989), Hameed et al. (2004), Sandvad (1999) in this area has revealed that reverberation time can be an important parameter used by the human auditory system to estimate room size. Cabrera et al. (2005) found that in situations where the actual size of the room is held constant, and reverberation time and source-receiver distance are varied, early-to-late energy ratio (clarity index) may be a good predictor of room size.

However, much speculation also exists on the potential for the early reflection pattern of the room to convey room size information.

The field of room acoustics is currently dominated by techniques associated with room impulse responses. These can be thought of as the fully detailed pattern of echoes and reverberation associated with sound travelling from a particular source position to a particular receiver position in a room.

The impulse response, when properly recorded, contains all the information about the acoustics of a steady-state room for a given set of source and receiver locations. Consider a single sound impulse emitted from a source in a room as shown in Figure 1.

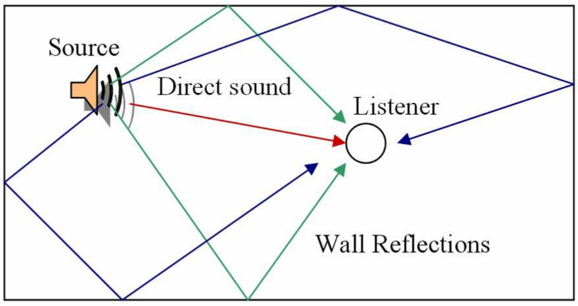

Figure 1. Direct and reflected sound paths in a room.

In the case of an omnidirectional source, spherical waves will propagate in all directions from the source, and the sound first heard in the listener's position will be the direct sound, which has travelled the shortest distance.

It will be closely followed by the early reflections, which are parts of the wave reflected by the room boundaries. These reflections will also be weaker than the direct sound because of the longer distance travelled and the attenuation due to surface absorption. Eventually these reflections will continue to be reflected and audible at the listener's position until all the sound energy has been absorbed. An example of a room impulse response is shown in Figure 2.

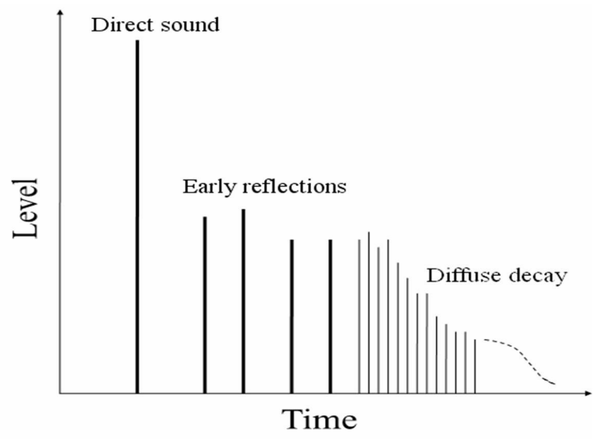

Figure 2. The impulse response in time domain.

The impulse response in the time domain shows the direct sound, followed by easily identifiable early reflections, and then followed by a gradually denser reverberant tail that decays exponentially to silence.

The auditory system relies on the direct sound to discern the direction of the incoming acoustic signal. The early reflections contribute strongly to the sense of space and perhaps to the size of the room, while slightly later reflections have been found to contribute more to a sense of envelopment. The density of the early reflections will, for example, be greater in a small room than in a bigger room, since the reflections will arrive sooner to the listener. The amplitudes of the early reflections will also be dependent on the room size, as longer distances travelled by the waves reduce the sound level. The reverberation time is generally longer in large rooms, as it is a ratio of the volume to sound absorbing area, and the volume of the room usually grows faster than the sound absorbing area.

Previous studies using real rooms

McGrath et al. (1996) found that both blind and sighted (but blindfolded) persons were able to judge the size of the rooms and their own location in actual rooms from just the sound (own speech and other sounds).

Mershon et al. (1989) examined auditory perception of room size as a side-issue in a study on auditory distance perception. Their results found longer reverberation time and greater source-receiver distance to elicit larger room size judgements, and a non-significant tendency of the background noise to increase the apparent room size.

Previous studies using room simulations

Sandvad (1999) studied the issue of auditory perception of reverberant surroundings and performed three experiments for the study. In the first experiment the subjects were presented binaural recordings of a speech signal and pictures of the rooms where the signal was recorded. They were then asked to point to the rooms to which the acoustic responses corresponded. Since the subjects' answers were 70% correct, the study concluded that the room impulse response contains information about the room that the listeners are able to extract.

The results of the second experiment are of particular interest for the proposed study. The test subjects were asked to estimate the size of the room after listening to the recordings, finding that some listeners used the direct to reverberant energy ratio as cue for room size, while others used the reverberation time.

The third experiment was a comparison between acoustic signals obtained by simulations of a room and signals generated directly from measurements. The analysis showed that energy measures are the most important for estimating the room dimensions.

Cabrera et al. (2005) investigated the auditory room size perception by conducting subjective experiments using the method of paired comparisons. They obtained room size ratings using binaurally presented stimuli.

The study consisted of three experiments. In Experiment 1, the authors used a room acoustics software program (CATT-Acoustic) to simulate three rooms of identical shape, but large differences in room volume (1:8:64). Various distances and three reverberation times were tested. Stimuli for the experiment were generated by recording a person (male) in an anechoic room saying "I'm speaking from over here". The generation of these stimuli is described in detail in Cabrera et al. (2002). The recording was convolved with binaural impulse responses. Reverberation time and early-to-late energy ratio (known as 'clarity index') were found to be good predictors of perceived room size. They found evidence of an inverted-U response occurring as the room increased in size for a fixed reverberation time - which can be explained by the fact that reverberation level decreases, approaching inaudibility, in a very large room (compared to a small room with the same reverberation time).

In Experiment 2, binaural recordings of the same anechoic speech played from a JBL4206 loudspeaker were made in the reverberation room with a KEMAR dummy head. The recordings were made at three distances from the source, and three reverberant conditions were used, by adding sound absorbing materials in the room. It was also tested whether amplification (+/- 6dB) affected room size perception. The conclusion of the experiment was that reverberation time is the most dominant cue in the perception of room size. It could reasonably be hypothesized that amplification would be inversely related to room size perception, but that only occurred with the closest recording position.

In Experiment 3, binaural recordings of a music extract for various seats in a large auditorium (the Michael Fowler Centre in New Zealand) were used. In that case, it was found that clarity index is quite closely related to room size perception. Source-receiver distance accounted for the residual reasonably well.

Martignon et al. (2005) used binaural and stereophonic (O.R.T.F. - two cardioid microphones spaced by 17 cm and angled outwards at 110 degrees) room impulse responses, which have been recorded in five concert auditoria, to test the spatial audio quality of four reproduction systems: conventional stereophony, binaural headphones, stereo dipole, and double stereo dipole. In a subjective test, the respondents rated the room size, sound source distance and realism of the reproduction. The study found that the stereo dipole and O.R.T.F. stereophonic systems appear to work better than non-individualized binaural headphone reproduction and double stereo dipole systems. The study also showed weak or non-existent relationship between auditory room size ratings and actual room size, a finding which is at odds with previous studies.

Aim and application

The purpose of this study was to obtain real room data in relation to room acoustical parameters and to provide a potential verification of audio simulation systems.

Possible applications of room size perception modelling/simulation include architectural acoustics models, architectural acoustics design, automotive audio and contexts where the auditory environment is manipulated (eg music production, computer games, auditory display, extended reality, virtual reality).

METHOD

The experiment described in this paper involved the subjective estimation of room size, based on the sound of reproduced speech within the room.

Subjective test procedure

The listeners were asked to rate the auditory perceived size of the rooms where they listened to a speech recording.

A blindfold was used to block vision throughout the experiment, as the intention was to obtain the room ratings based only on the reproduced voice. The listeners were helped by the test operator to move between the rooms that were used in the tests. They were also asked to wear earmuffs while guided to the specific rooms and while being seated on the chair.

Once inside each room, a speech signal was played from a loudspeaker. After listening to the speech, the respondents were guided outside the room and taken into the reference room, where they were asked the question “How big was the room that you have just been in?” They were suggested to use a scale from 0 to 100 in rating the rooms, where a possible rating for the reference room would be approximately 10. After answering this question, the listeners were then escorted to the next room, where again the speech signal was played for them.

The complete tour took approximately one hour, including a 10 minutes break after the first 25 minutes of the test.

17 listeners participated in the experiment, with age ranging from 20 to 35. Five respondents were female and twelve were male. They all reported normal hearing.

Characteristics of the test signal

Stimuli were generated initially by recording a person (male) in an anechoic room saying ”I’m speaking from over here”. The recording was made at a distance of 0.25 m from the speaker using a Brüel & Kjær Type 4190 free field microphone, using a windshield. These stimuli were also used in previous experiments (Cabrera et al. 2002).

The audio files were played from the computer located in the control room using Pro Tools audio software.

The recorded speech signal was reproduced by means of Yamaha MSP5A monitor loudspeakers located in the evaluated rooms. These loudspeakers were chosen due to the fact that their size is about the same size as a human head, thus obtaining a possible more realistic playback system.

The set-up was calibrated before presenting the sound files to the listeners. The level of the speech signal was measured in the anechoic room at 1 m distance from the loudspeaker membrane. The sound volume was adjusted to match the sound pressure level for the speech in the original recording (56 dbA @1 m). Once the first loudspeaker was calibrated in the anechoic room using the speech signal, pink noise was played through the loudspeaker and the sound pressure level was measured at the front face of the loudspeaker. This level (93.9 dbA) was used to calibrate all the loudspeakers in the rooms used for the listening tests. After the calibration, the volume knobs of the loudspeakers were covered with tape to prevent adjusting the volume accidentally.

Room selection and description

The listening tests were performed in three rooms located on Level 1 at the Faculty of Architecture, The University of Sydney, Australia. A fourth room was used as reference. These rooms were chosen considering the practical matters regarding the fact that the listeners were blindfolded and had to be guided by one of the test operators through the corridor and the rooms.

The first room chosen for the experiment was a small room (15 m^3) as shown in Figure 3.

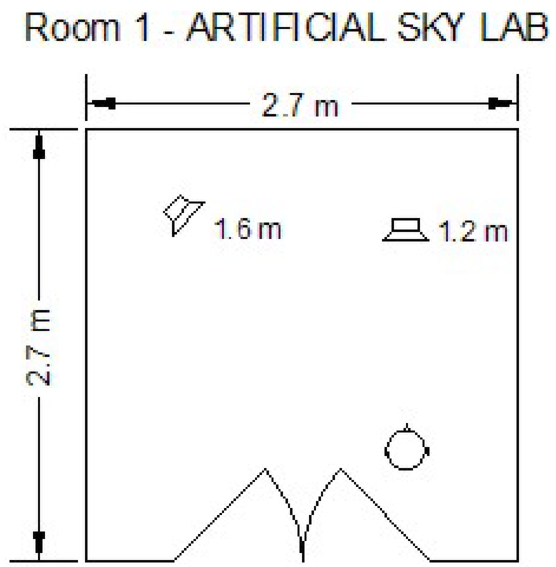

Figure 3. Listening test set-up in room 1 (first and second loudspeaker).

The walls of the room were made of mirrors and the ceiling was made of semi opaque plastic. The floor was covered with a linoleum layer.

Two loudspeaker positions and two listener positions (one location with two different orientations) were used in this room.

The second room used for the listening test had a larger volume (124 m^3) than the first one. Figure 4 shows a sketch of the room.

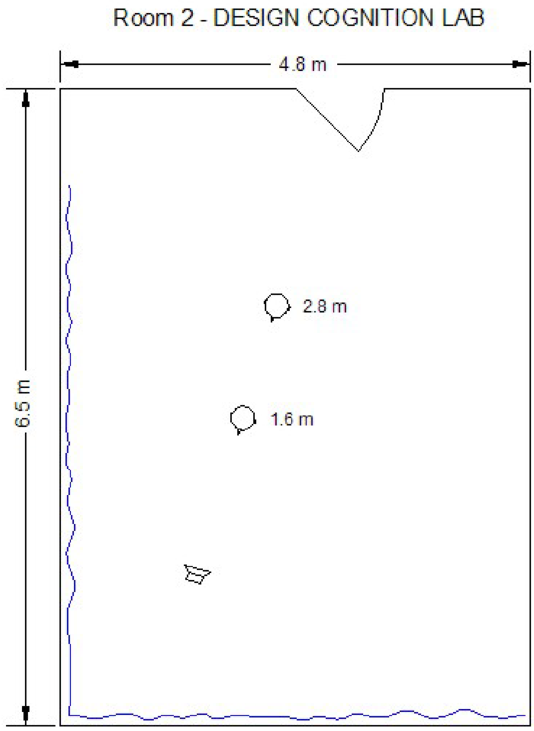

Figure 4. Listening test set-up in room 2 (first loudspeaker).

Two of the room walls were made of concrete, while the other two were plasterboard walls. The floor was carpeted and the room was unfurnished. On the two concrete walls there was mounted a curtain, which could be drawn in and out making possible the change in the room acoustics.

Two loudspeaker positions and two distances for each loudspeaker to the listener were used in the listening test in this room. Figure 5 shows the location of the second loudspeaker.

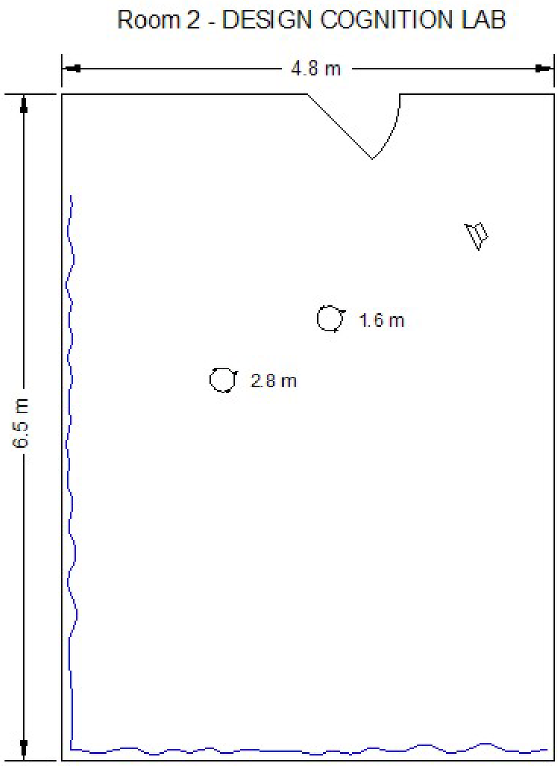

Figure 5. Listening test set-up in room 2 (second loudspeaker).

For each loudspeaker-listener distance and positioning the listening test was performed with and without the curtains drawn. A number of eight stimuli were thus generated in this room.

The third room had the largest volume (188 m^3) and had a non rectangular shape as shown in Figure 6.

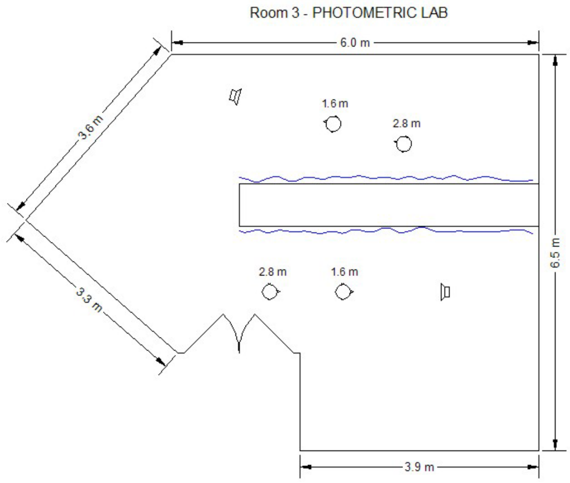

Figure 6. Listening test set-up in room 3 (first and second loudspeaker).

The room had concrete walls, except one, which was a plasterboard wall. While all the other rooms were empty during the listening test, this room was furnished with some photometric equipment. The room was separated in two sections by a 1 m high bench as shown in Figure 7.

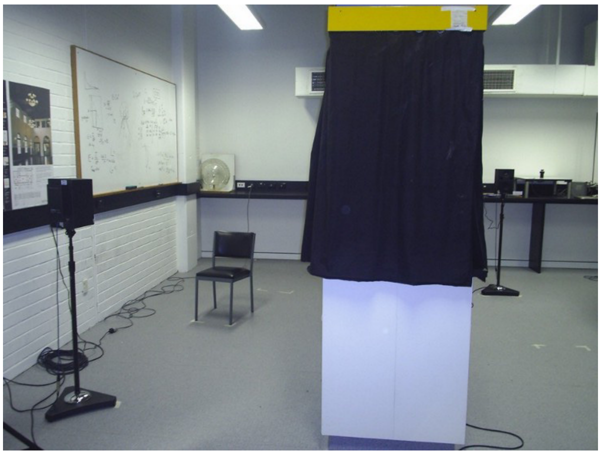

Figure 7. The bench in the photometric lab (Room 3)

Above the bench there was a 1.2 m high curtain separating the room in two sections. The curtain was left in the original position during the experiment.

Two loudspeaker positions were used in the listening test in this room and two distances from each loudspeaker to the listener.

The reference room was a vestibule with a volume of 19 m^3 and connected different rooms as shown in Figure 8.

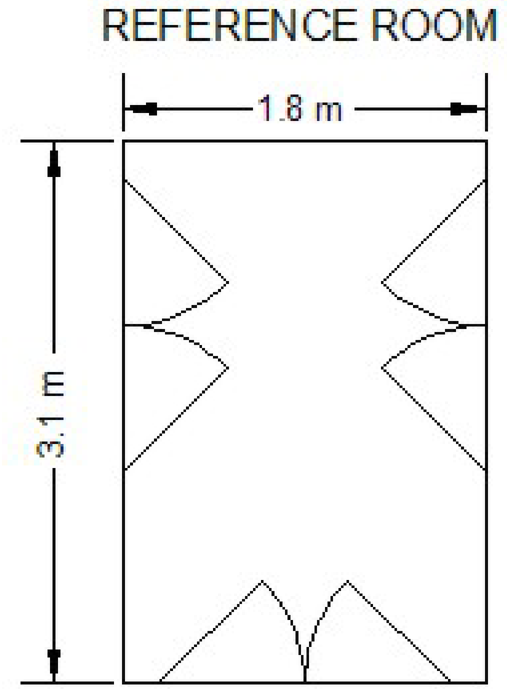

Figure 8. The reference room.

The room was not furnished and all the walls were made of concrete. The concrete floor was covered with a linoleum layer. The wooden doors represented a significant percentage of the total area of the room.

Acoustical characteristics of the situations

The room impulse responses were measured for the three rooms for all the source and receiver positions. The binaural impulse responses were recorded with a Brüel & Kjær Head and Torso simulator (HATS), as shown in Figure 9.

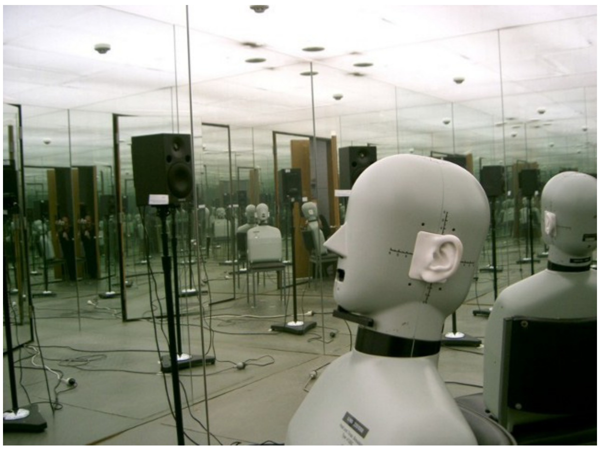

Figure 9. HATS impulse response measurements in the artificial sky lab (room 1).

The impulse responses were also measured using a Brüel & Kjær omnidirectional microphone. The measurements have been made using the same loudspeakers used for the listening test (Yamaha MSP5A) and the test signal was an exponential swept sine wave played from a portable computer.

The room acoustical parameters were obtained from the impulse responses. These include reverberation time (T30), early decay time (EDT), clarity index (C80), speech transmission index (STI) and inter-aural cross correlation coefficient (IACC). Strength factor (G) was not determined, but the reproduced sound pressure level (L_{eq}) of each stimulus was.

RESULTS

The reference room was perceived through both vision and audition, including the sound of the subject's own voice. The relationship between this experience and the auditory experience of the test rooms was, to some extent, a mental construct of the subjects with great potential for inter-subject variation. This led some subjects to give larger room ratings than others. This was not considered a problem, since the role of the reference room was to allow subjects to quickly grasp the concept of a ratio scale, rather than to use the reference room in the analysis. Therefore the raw ratings of each subject were divided by the mean rating for that subject, thereby factoring out the reference room as the base for the ratio scale. This scaling provided greater compatibility between subject results, which was important, considering the small size of the dataset. Mean stimulus ratings following this scaling process are shown in Figure 10.

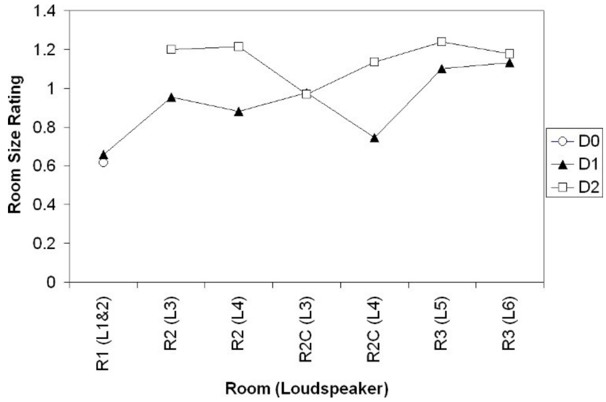

Figure 10. Mean scaled results for each stimulus, identified by loudspeaker distance (D0-2), room number (R1-3), drawn curtain condition for Room 2 (R2C) and loudspeaker number (L1-6).

Effect of room size and curtain

Results show that physical room size significantly affects auditorily perceived room size, with a positive relationship between the two. Analysis of variance (ANOVA) followed by a Scheffe test shows this to be a strong effect (f=24.7, p<0.0001), with significant differences between the mean results for each of the three rooms. However, when results for Room 2 are separated into curtains drawn and aside, a Scheffe test shows that the rating difference between Room 2 (curtains apart) and Room 3 (R2 – R3) is not significant, nor is the effect of curtains (R2 – R2C) significant for Room 2. All other rating differences remain significant (including R2C – R3). Room size ratings, with the curtain conditions shown separately, are given in Figure 11.

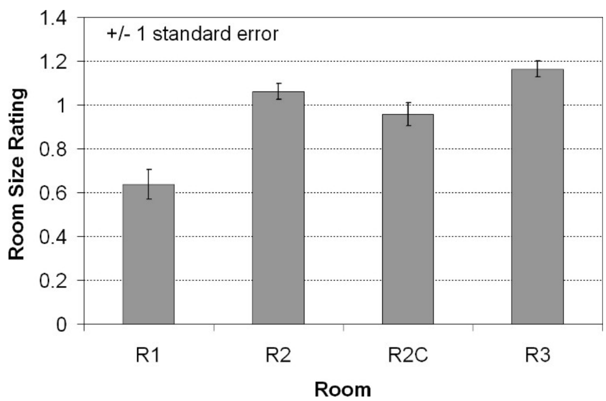

Figure 11. Mean scaled ratings of room size for the three rooms, showing Room 2 with the curtains drawn (R2C) and apart (R2).

Effect of loudspeaker distance

It is already apparent in Figure 11 that the distance between the loudspeaker and the subject positively affects the room size rating. Since Room 1 was too small to have Distance 2 in it, the analysis of the effect of distance uses data only from Rooms 2 and 3. ANOVA yields a significant effect of distance (f=16.5, p<0.0001) and a somewhat reduced effect of room (f=6.4, p=0.002) when Room 1 data are omitted (with no significant interaction). The effect of distance is illustrated in Figure 12.

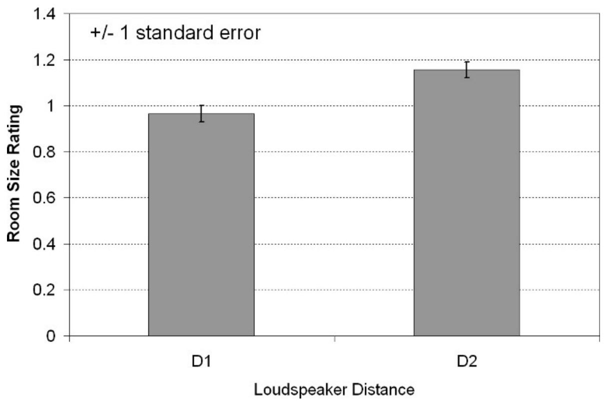

Figure 12. Mean scaled ratings of room size for the two loudspeaker distances used in R2, R2C and R3.

Acoustical Correlates of Ratings

Examination of the results in relation to measured acoustical parameters showed that the stimulus sound pressure level discriminates well between the perceived room size rating of Room 1 and the larger rooms. There is also a generally negative relationship between stimulus sound level and room size rating, as shown in Figure 13 (left). Clarity index, in the form of mid frequency C_{80}, also shows some promise in forming negative correlations to room size rating within a given room condition, as shown in Figure 13 (right).

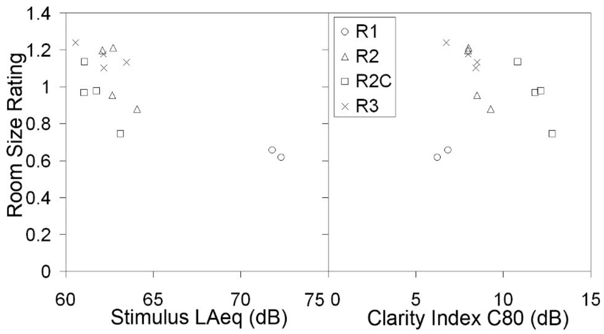

Figure 13. Relationships between room size rating and the objective parameters of stimulus sound level and clarity index.

The correlation between A-weighted stimulus level and clarity index (mid frequency C80) is low enough for them to be considered orthogonal (r^2=0.24). A multiple regression with these acoustical parameters yields a model predicting 88% of variance in the mean room size ratings, as illustrated in Figure 14.

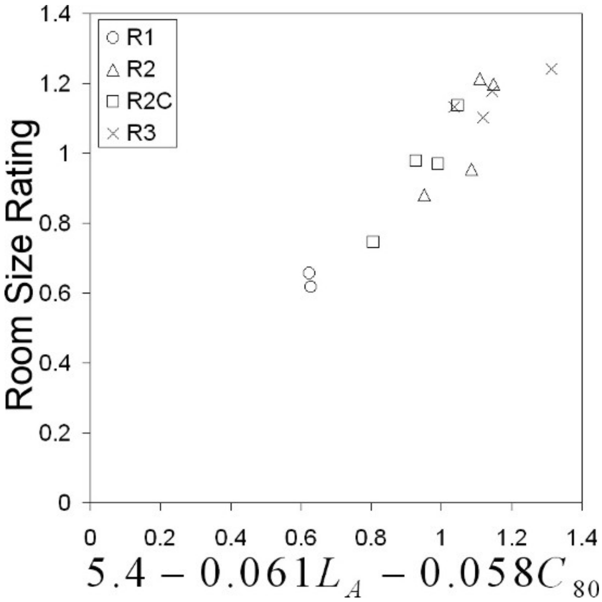

Figure 14. Model of room size rating based on A-weighted equivalent stimulus level and mid-frequency clarity index C_{80}.

Cluster analysis of subjects

Previous studies of room size perception have sometimes found that subjects can be grouped into various response patterns. As a tentative exploration of this, a complete linkage cluster analysis was conducted to investigate subject response groupings, yielding the dendrogram shown in Figure 15. The first division of the subjects yields a cluster of four and a cluster of thirteen subjects (labled Cluster 1 and 2 respectively).

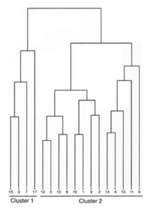

Figure 15. Cluster analysis of subjects (the subject identification numbers correspond to the order in which they did the experiment).

Since Cluster 1 has only four subjects, further analysis was concentrated on Cluster 2, which might be considered to be the more typical response pattern. As illustrated in Figure 16, Cluster 2 has a seemingly stronger relationship with C_{80} than the full subject set, apart from the Room 1 stimuli. However, this analysis is merely indicative of an approach that might be taken with a larger subject sample in future for a higher confidence analysis.

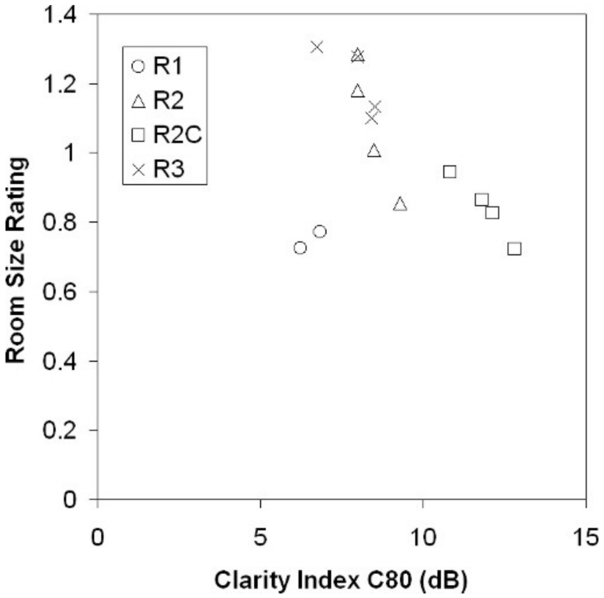

Figure 16. Relationship between C_{80} and room size rating for Cluster 2.

DISCUSSION

The apparent relationships between sound pressure level and perceived room size, and between clarity index and perceived room size, make some sense in terms of the typical characteristics of rooms of various sizes. Small rooms tend to have less sound absorption, and hence a higher diffuse field level. Furthermore, large distances are only possible in large rooms, and the direct soundfield level decreases with distance, usually following -6 dB per doubling of distance.

The effect of clarity index on perceived room size was observed previously for binaural headphone experiments using the same speech sample as the present study (Cabrera et al. 2005). However, a study of perceived room size in concert auditorium simulations, using four audio systems, found that the acoustical correlate of perceived room size can depend on the audio recording and reproduction system (Martignon et al. 2005). Nevertheless, the present study, using soundfields of real rooms, provides some support to the binaural headphone study of Cabrera et al. (2005). However, a verification experiment is planned, in which the soundfield conditions of the present experiment are simulated using various audio reproduction techniques.

The analysis in this paper is restricted to basic acoustical parameters. However, more complex acoustic phenomena such as the early reflection sequence or the fine frequency response could reasonably be hypothesised to affect room size perception. These remain to be investigated in more carefully controlled experiments using larger stimulus sets.

This experiment could be improved by having a larger number of participants, since the 17 subjects yielded only 14 data-points each. The authors plan to continue the project with more subjects to increase the reliability of the results, and the prospects of subject sub-group analysis.

CONCLUSIONS

This study supports previous studies regarding the possible influence of the clarity index on the acoustically perceived room size. The results of this study also show that there is a generally negative relationship between stimulus sound level and room size rating.

ACKNOWLEDGMENTS

The authors are thankful to the experiment volunteers. They also thank Phil Granger, Ken Stewart and Matt Storey for assistance with the experiment set up, and Simon Hayman for assistance with the analysis. This project was supported by a University of Sydney Research and Development grant.

REFERENCES

Cabrera, D and Gilfillan, D 2002, Auditory distance perception of speech in the presence of noise, *Proc. Int. Conf. Auditory Display*, Kyoto, Japan, 2002 (http://www.icad.org)

Cabrera, D, Jeong, D, Kwak, HJ, Kim, J-Y 2005 Auditory Room Size Perception for Modelled and Measured Rooms, Inter-Noise 2005, *The 2005 Congress and Exposition on Noise Control Engineering*. Rio de Janeiro, Brazil, August 2005.

Hameed, S, Pakarinen, J, Valde, K, Pulkki, V 2004 Psychoacoustic Cues in Room Size Perception, in *Proceedings of Convention paper # 6084 presented at the 116th AES Convention*, Audio Engineering Society, Berlin, Germany, May 8-11, 2004.

Martignon, A, Azzali, A, Cabrera, D, Capra, A, and Farina, A 2005, Reproduction of Auditorium spatial impression with binaural and stereophonic sound systems, *Proc. 118th Audio Engineering Society Convention*, Barcelona, Spain, 2005.

Mershon, DH, Ballenger, WL, Little, AD, McMurtry, PL, & Buchanan, JL 1989 Effects of room reflectance and background noise on perceived auditory distance. *Perception*, 18, 1989, p. 403-416.

McGrath R, Waldmann T, and Fernström M 1996, Listening to rooms and objects, *Proc. 16th Audio Engineering Society Int. Conf.*, Rovaniemi, Finland, 1996.

Sandvad, J Auditory perception of reverberant surroundings.(1999) *Journal of the Acoustical Society of America*, Vol. 105, No. 2, Pt. 2, February 1999, p. 1193 (paper 3pSP3)