ENVIRONMENTAL NOISE CONTROL
The 2005 Congress and Exposition on Noise Control Engineering
07-10 August 2005 - Rio de Janeiro - Brazil

Auditory Room Size Perception for Modeled and Measured Rooms

Densil Cabrera[a], Daeup Jeong[b], Hyun Jeong Kwak[c], Ji-Young Kim[d]

[a,c]Faculty of Architecture, University of Sydney, NSW 2006, Australia
[b,d]Department of Architectural Engineering, Chonbuk National University, Duckjin-dong, Duckjin-gu, Jeonju, Chollabuk-do, Korea, 561-756

[a, c]densil@arch.usyd.edu.au; [b, d]daeupj@chonbuk.ac.kr

Abstract Although there are many studies investigating auditory spatial impression in rooms, there are few that directly investigate the seemingly more basic question of auditory room size perception. In this study, subjective experiments using the method of paired comparisons were conducted to obtain room size ratings using binaurally presented stimuli. In the first experiment, binaural impulse responses from a computer-modelled room were used to auralize an anechoic speech sample. Room volume, source-receiver distance and reverberation time were investigated as parameters. The second experiment used binaural recordings of speech made in a real room of fixed size (for the same anechoic speech sample as Experiment I), with source-receiver distance and reverberation time as experiment parameters. The final experiment used binaural impulse responses of a concert auditorium convolved with anechoic music – so that both the room volume and reverberation time were constant. Results show that reverberation time strongly affects room size perception (much more so than the physical room volume). In a room of fixed volume but variable absorption and source-receiver distance, clarity index can be a good predictor of perceived room size. A comparison of Experiments I and II shows little or no difference between results for auralizations of computer modelled rooms and binaural reproductions of a real room. Results from the second and third experiments were compared with results from previous studies (of auditory distance perception, speech quality and spatial impression) which used identical stimuli. Auditory room size perception is not closely related to auditory distance perception (Experiment II), and is related to auditory intimacy (music stimuli, Experiment III) and speech quality (speech stimuli, Experiment II).

1. INTRODUCTION

It seems reasonable that, through a lifetime of experience, people can learn the characteristic auditory features of variously sized rooms. Small rooms tend to have shorter reverberation times, higher diffuse field sound levels, denser and more intense early reflection patterns, and more distinct resonant frequencies than large rooms. Furthermore, large source-receiver distances are only possible in large rooms, so auditory distance perception and room size perception could be related. Sandvad [1] found that subjects could usually correctly identify photographs of rooms that corresponded to binaurally reproduced sound fields representing those rooms. In subsequent experiments, Sandvad found that some listeners used the direct to reverberant energy ratio as cue for room size estimates, while others used the reverberation time. McGrath et al. [2] found that both sighted (but blindfolded) and blind subjects are able to distinguish small and large rooms from the sound of their own speech and other incidental sounds (in actual rooms). Blind subjects evaluated the room acoustical environment more quickly and accurately than sighted subjects.

This study starts with the assumption that the relationship between auditorily perceived and actual room size is complex, with factors other than actual room size having a substantial influence on the perception of room size. This assumption is supported by previous studies, which have found that reverberation time, reverberation level, source-receiver distance, and perhaps even background noise level affect perceived room size, even in rooms of fixed size [3, 4].

2. EXPERIMENT I: COMPUTER-MODELLED ROOMS

2.1. Approach

This study investigates the perception of room size as it is conveyed in room acoustical simulations, recordings and room impulse response convolutions via non-individualized binaural headphone techniques. The first experiment used auralizations of computer-modelled rooms (modelled using CATT-Acoustic). The room form was taken from a real reverberation chamber, scaled to three volumes (31 m^3, 249 m^3, 1997 m^3), each given three reverberation times (0.5 s, 1 s and 2 s). Three source-receiver positions were selected for each room volume, and binaural room impulse responses were generated. There was no normalization or level matching of these impulse responses – instead they represented the range of levels that would have occurred in the modeled acoustic conditions (meaning that the impulse response energy for a distant receiver positions in an absorptive room was much less than that for a close position in a room having little sound absorption). The binaural impulse responses were optimized for the headphones used in the experiment (Sennheiser HD600).

Stimuli for the experiment were generated initially by recording a phrase of speech in an anechoic room. The phrase, “I’m speaking from over here,” was recorded from a male speaker, at ‘medium’ vocal projection. The free field microphone distance was 0.25 m on axis with the speaker (using a wind shield). This anechoic recording was edited so as to be repeated once. The recording was convolved with binaural impulse responses.

Paired comparisons tests are good at making robust distinctions between subtly different stimuli, but are also inefficient in terms of the number of stimulus presentations required for even medium, and especially large, stimulus sets. Hence, although this experiment used 27 stimuli, it was not practical to test every pair combination of the 27 stimuli. Instead, four sub-experiments were devised. Experiment Ia used only the 2 m receiver distance, with three room volumes and three reverberation times. Experiment Ib used only the mid room volume, with three receiver distances and three reverberation times. Experiment Ic used only the 1 second reverberation time, with three room volumes and three receiver distances. The remaining eight stimuli were tested in Experiment Id, along with the stimulus common to all sub-experiments (mid room volume, 2 m distance, 1 second reverberation time).

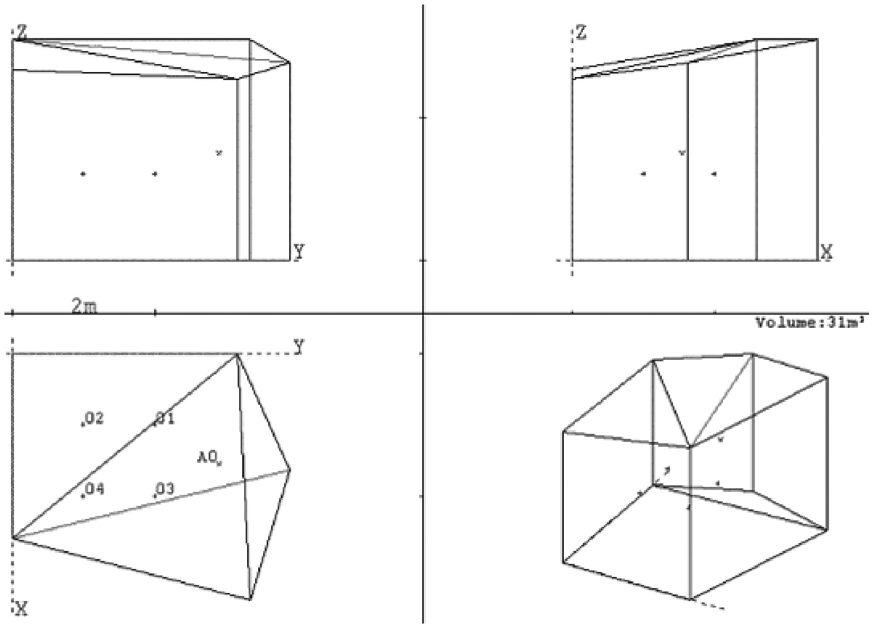

Figure 1: Form of the computer-modeled room used in Experiment I.

Sixteen subjects participated, each subject assessing all 144 pairs in random order (the four sub experiments were mixed together for stimulus presentation). Intra-pair order was counterbalanced between subjects. For each pair, subjects answered the question, "Which room is bigger?" The experiment was conducted in a quiet (anechoic) room.

2.2. Results

Results (scaled following Thurstone's Case V, accounting for saturated results and missing data as described by Togerson [5]) are shown in Figure 2. Using this scaling, 0 is the mean response (probability of stimulus selection is 0.5), and units are standard deviations of the assumed normal distribution (1 unit corresponds to a selection probability of 0.84, 1.6 units to p=0.95, and 2 units to p=0.98). Reverberation time has the strongest effect on room size perception. Actual room size and source-receiver distance also tend to positively affect perceived room size. In this experiment, the best acoustical correlate of perceived room size is simply the reverberation time (r=0.93). The correlation for C_{50} is r=-0.79, and for C_{80} is r=-0.84. When the three room volumes are considered separately, C_{80} performs somewhat better than reverberation time (C_{80} yields r= -0.99, -0.98 and -0.97 for the small, medium and large rooms respectively, whereas reverberation time yields r = 0.96, 0.98 and 0.94).

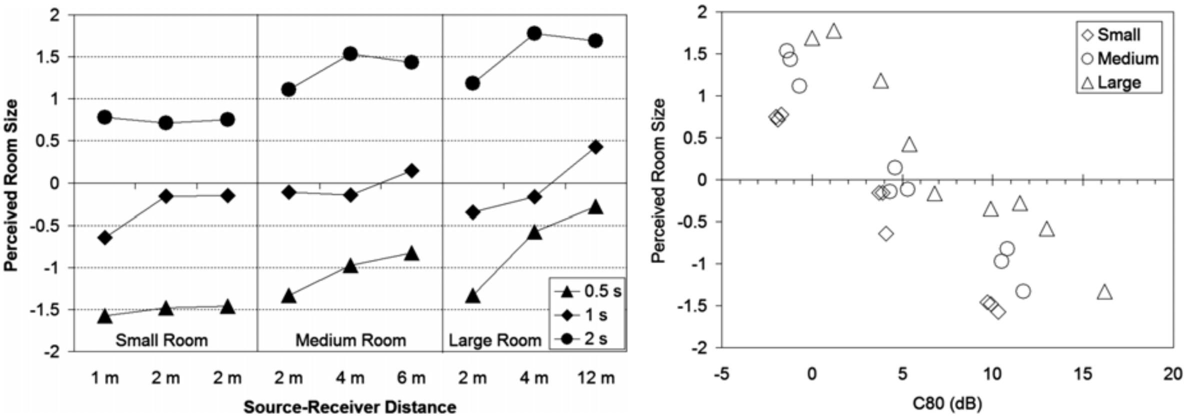

Figure 2: Results of Experiment I, showing subjective ratings in terms of source-receiver distance, reverberation time, and room volume on the left chart. The relationship between C_{80} and perceived room size is shown on the right, with data for the three room volumes separated.

3. EXPERIMENT II: ONE ROOM WITH VARIABLE REVERBERATION

3.1. Approach

This experiment is similar to Experiment I, except that a real room was used, with anechoic speech recordings played from a loudspeaker (JBL4206) and recorded by a KEMAR dummy head (Figure 3). The generation of these stimuli was described previously [6], so details are not included in the present paper. The room was a rectangular reverberation room, 130 m^3, and its reverberation time was controlled through the introduction of sound absorbing material – yielding mid-frequency reverberation times of 0.6 s, 2.1 s and 5.1 s. Source-receiver distances of 0.9 m, 2.7 m and 5.1 m were used. Note that the actual room size did not change in this experiment. The subjective test was conducted with binaural recordings made in this manner, reproduced over headphones in an anechoic room. Sixteen subjects rated the 36 stimulus pairs.

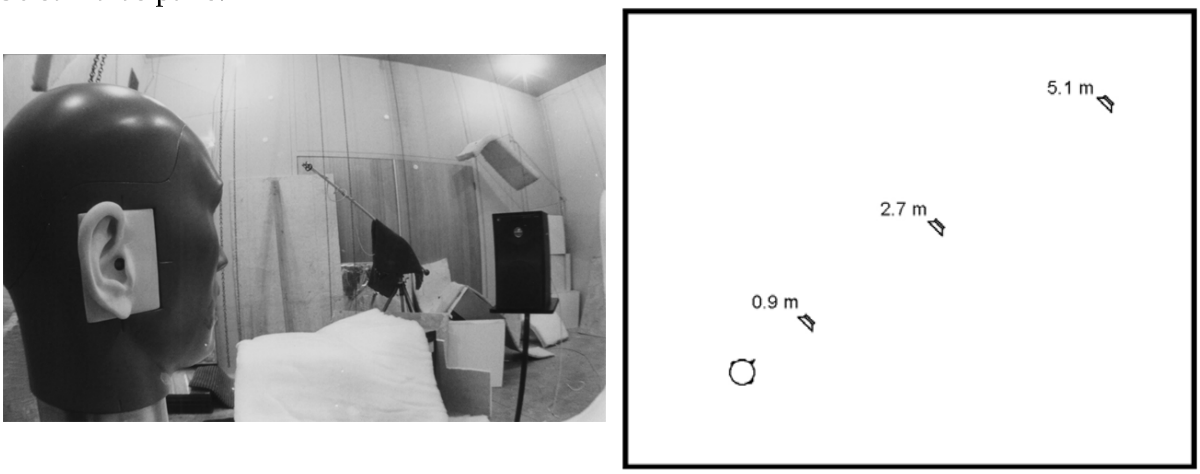

Figure 3: Photograph and plan of the room in which Experiment II stimuli were generated.

3.2. Results

Like Experiment I, results show a strong effect for reverberation time, and a weaker effect for source-receiver distance (with one slightly contrary result). As might be expected, the effect of distance is more evident for the 0.9 m to 2.7 m distance change (a three-fold increase), than for the 2.7 m to 5.1 m distance change (less than a two-fold increase).

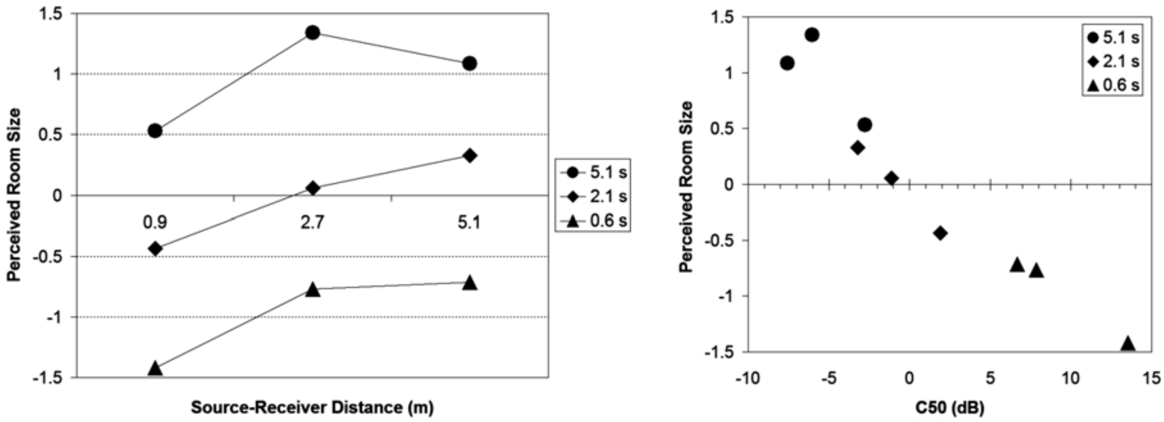

Figure 4: Results of Experiment II, in terms of source-receiver distance and reverberation time shown in the left chart. The right chart shows the relationship between C_{50} and perceived room size.

Most of the acoustical parameters were represented in three frequency bands. Low frequency is the combination of 125 Hz and 250 Hz octave band values; mid frequency combines 500 Hz and 1 kHz octave band values; and high frequency combines 2 kHz and 4 kHz octave band values. A high correlation was found between room size judgments and various measures of early to late energy ratio in the impulse responses (including C_{50}, C_{80}, D_{50} and T_s). The mid frequency clarity index C_{50} yielded the best correlation, with a coefficient of r=-0.97 (p<0.0001). This relationship is shown in Figure 4.

One negative hypothesis that might be made regarding auditory room size perception is that it could be closely related to auditory distance perception – because large distances are only possible in large rooms. With the strong effect of reverberation time changes in this experiment, there is little relationship between actual distance and perceived room size (r=0.31, p=0.44). There is more of a relationship between perceived distance (data from a previous experiment using these stimuli [6]) and perceived room size, but this also fails to reach significance (r=0.51, p=0.17). This lack of significance is partly due to the fact that only nine cases are considered in this experiment – nevertheless, the results suggest that auditory distance perception and auditory room size perception are different in this context. Nevertheless, there is a correlation of r=-0.90 (p=0.0004) between perceived room size and the speech quality results of Cabrera and Gilfillan [6] for the same stimuli, which is unsurprising considering the close relationship between perceived room size and C_{50}.

4. EXPERIMENT III: CONCERT AUDITORIUM

4.1. Approach

Since much of the interest in auditory spatial impression is with auditorium acoustics, this part of the study used binaural measurements made in a concert auditorium, examining how perceived room size varies within this auditorium (the actual room size and the room's spatially averaged reverberation time did not vary at all). The hypothesis underlying this experiment is that perceived room size can vary in complex room, such as an auditorium, even though the room volume is constant and reverberation time is approximately constant. Knowledge of such variation in perceived room size could be useful in understanding more generally studied aspects of auditory spatial impression. However, if judgments of perceived room size are indistinguishable from other judgments of spatial impression (including perceived distance of the sound source), then the concept of perceived room size in this context becomes redundant, and perhaps only of incidental interest.

The auditorium used in this experiment is the Michael Fowler Centre, in Wellington, New Zealand. This is an interesting auditorium to study because it was designed with a particular regard to optimizing auditory spatial impression [7]. It has a volume of 25,000 m^3, and seats 2,500. In spite of its large capacity, it achieves visual intimacy partly through its elliptic form, a gallery level that surrounds the stage, and large overhanging acoustic reflecting panels. Its maximum internal length, width and height are 43 m, 36 m and 22 m respectively.

Figure 5: Views of the Michael Fowler Centre from seats 26H3 (left) and W08 (right).

Like Experiment II, stimuli used in this experiment were initially prepared in a previous study – so details of the stimulus preparation are given elsewhere [8]. A key difference between this experiment and Experiments I and II is that music stimuli were used (binaural impulse responses were convolved with an anechoic recording of a fragment of Mozart's *Marriage of Figaro* overture). The fifteen stimuli were tested in a paired comparisons test (every pair was tested), with fourteen subjects, all music students aged in their mid-twenties. Unlike Experiments I and II this experiment merely observes effects found in a given auditorium, rather than systematically controlling specific room parameters through experiment design.

4.2. Results

Results (Figure 6) show little variation in room size perception for most receiver positions, but an exceptionally small rating for the closest seat in the stalls (H32), and large ratings for three other positions (W08, 23D6 and 26H3). In the stalls, generally the perceived room size increases slightly between front and rear receiver positions. The seat W08 is deep under the gallery (Figures 5 and 6), and receives less significantly less mid and high frequency sound than other seats at a similar distance from the source. The result for 23D6 might be explained by the dummy head being outside the main directivity coverage of the loudspeaker source at high frequencies (since a studio monitor was used as the source). This is suggested by the relatively weak high frequency sound pressure level measurements for this stimulus. Sound pressure level in the high frequency range also appears to be influential in the rating difference between 26H3 and nearby 26C9, since most other acoustical parameters are similar for these two seats.

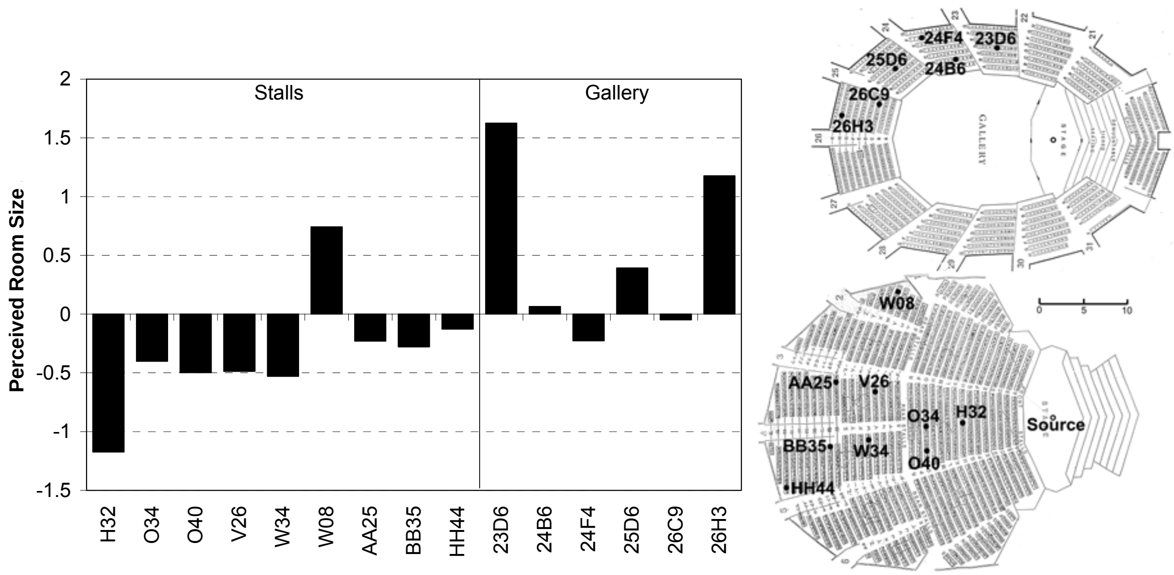

Figure 6: Results of Experiment III, in terms of receiver seat in the auditorium.

While there appears to be some relationship between source-receiver distance and perceived room size, the result for 23D6 in particular undermines the correlation (the correlation is insignificant, at r=0.11, p=0.69). Figure 7 shows that when perceived auditory distance (from [8]) is substituted for physical distance, the relationship with perceived room size strengthens greatly (r=0.83, p<0.0001). In the previous work by Cabrera *et al.* [8], auditory intimacy was rated on an integer scale from 0 to 10. The results were closely related to auditory distance estimates (r=-0.88, p<0.0001), and they appear to be very closely related to the perceived room size results of the present study (Figure 7, right). The result for seat 23D6 undermines what would otherwise be a correlation of r=-0.92 (p<0.0001), yielding a correlation of r=-0.81 (p<0.0001).

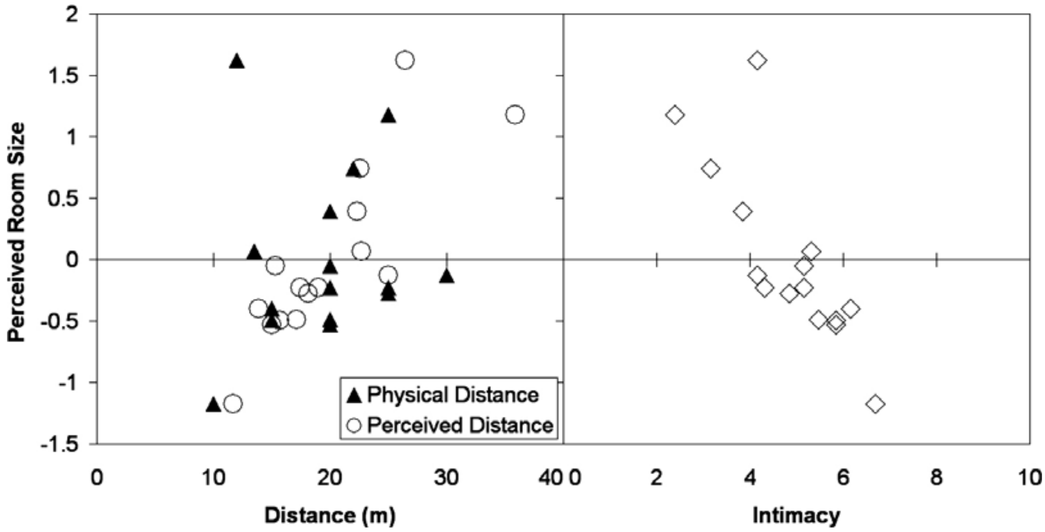

Figure 7: Results of Experiment III, in terms of physical source-receiver distance, perceived source receiver distance (from Cabrera et al. [8]) and auditory intimacy [8].

Stimulus sound pressure level (measured at the ears of a dummy head), especially in the treble range, predicts perceived room size moderately well (r=-0.79, p=0.0002 – Figure 8). In this context, there is quite a close relationship between clarity index at high frequencies and stimulus sound pressure level, so a correlation also exists between C_{80} at high frequencies and room size perception, with the major outlier being 26H3 (Figure 8). If this outlier is excluded, then the correlation for C_{80} in the treble range is r=-0.89 (p<0.0001).

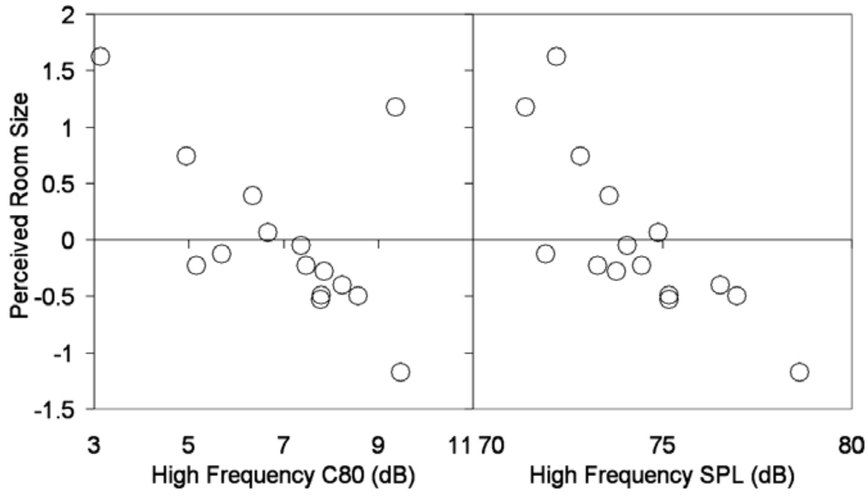

Figure 8: Results of Experiment III, in terms of C_{80} and stimulus sound pressure level at high frequencies (combined 2 kHz and 4 kHz octave band measurements) .

In this experiment, where neither reverberation time nor actual room volume were varied, there is not a great distinction between perceived room size and auditory distance, and scarcely any distinction between auditory room size and intimacy judgments. Hence, at least in the context of this experiment, auditory room size perception appears to be somewhat redundant, except that it gives some insight into the meaning of auditory intimacy. Nevertheless, the results do not undermine the observation in Experiments I and II of a relationship between clarity index and perceived room size for rooms of fixed physical volume.

5. DISCUSSION

The series of experiments reported in this paper concur with previous studies showing reverberation time to have a strong effect on perceived room size judgments. In addition, the present study identifies a close relationship between early-to-late energy ratios (known as clarity index) and perceived room size, especially in situations where the actual room volume is held constant. This may be related to Sandvad's observations [1] that direct to reverberant ratios can affect perceived room size. Since the three different experiments yield similar results, using subjects in three countries (Australia, Korea and New Zealand), the relationship between clarity index and perceived room size seems likely to apply beyond the specific contexts of the experiments. Nevertheless, clarity index does not fully explain the results (indeed, different versions of clarity index correlate best to the results of each experiment), and more data is required to develop a model of perceived room size based on room acoustical and stimulus parameters.

Eliciting room size judgments in experiments in which the actual room size does not change might have the danger of obtaining spurious results (where subjects, unable to distinguish the size of rooms, revert to other criteria in forming their responses). However, this study shows that even when the size of the room is varied, subjective responses are influenced primarily by other independent variables, especially reverberation time, but also the source-receiver distance.

One limitation of this study is the use of non-individualized binaural recordings, presented via headphones. There are well known directional perception artifacts in this type of stimulus (such that an image that should be frontally localized is instead rotated around the median plane, possibly even to the rear), and the extent of these artifacts was studied previously in the case of the Experiment II stimuli [6]. The rotation of the auditory image around the median plane was least severe for distant source-receiver positions and long reverberation times. Martignon *et al*. [9] find that non-individualized binaural headphone reproduction performs less well than simple alternatives (especially non-individualized stereo dipole), at least for concert auditorium simulations. Further work is required to assess the performance of various reproduction methods for room size, and the extent to which the reproduction system used in the present study limits its findings. A point in favor of the present approach is that the room size judgments reported were purely relative, and so would not be affected by an absolute difference caused by the presentation method.

Acoustical characteristics such as the strength and density of the early reflection sequence, or the modal density of a room, are plausible room size cues for small rooms. This study does not use these as parameters, and the investigation of these remains for future work. Nevertheless, Sandvad [1] finds that patterns of discrete reflections may not strongly influence auditory room size judgments.

6. CONCLUSIONS

This study indicates that auditorily perceived room size may be related to actual room size, but that room acoustical characteristics can have stronger effects. In situations where actual room size is held constant, and reverberation time and source-receiver distance are varied, clarity index may be a good predictor of perceived room size.

7. ACKNOWLEDGEMENTS

The authors thank the Acoustics Research Centre, University of Auckland, for hosting parts of this work. They thank Micklin Halstead (Marshall Day, Wellington) and Neville Brown (Michael Fowler Centre) for assistance with the auditorium measurements.

8. REFERENCES

[1] J. Sandvad, Auditory perception of reverberant surroundings, *Journal of the Acoustical Society of America* **105**(2) Pt. 2, 1193 (paper 3pSP3) 1999

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

[3] D.H. Mershon, W.L. Ballenger, A.D. Little, P.L. McMurtry and J.L. Buchanan, Effects of Room Reflectance and Background Noise on Perceived Auditory Distance, *Perception* **18**, 403-416, 1989

[4] S. Hameed, J. Pakarinen, K. Valde and V. Pulkki, Psychoacoustic cues in room size perception, *Proc. 116th Audio Engineering Society Convention*, Berlin, Germany, 2004

[5] W.S. Torgerson, *Theory and Methods of Scaling*, John Wiley & Sons, New York, 1962

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

[7] A.H. Marshall and J.R. Hyde, Some preliminary acoustical considerations in the design for the proposed Wellington (New Zealand) town hall, *Journal of Sound and Vibration* **63**(2), 201-221, 1979

[8] D. Cabrera, A. Nguyen and Y.-J. Choi, Auditory versus visual spatial impression: a study of two auditoria, *Proc. Int. Conf. Auditory Display*, Sydney, Australia, 2004 (http://www.icad.org)

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