Proceedings of The Institute of Acoustics

SUBJECTIVE SURVEY OF BRITISH CONCERT HALLS

Michael Barron

Martin Centre, Dept. of Architecture, Cambridge CB2 2EB.

INTRODUCTION

To complement the extensive acoustic measurement programme in 40 auditoria in Britain, a subjective survey is being undertaken in more than half of these auditoria. The procedure has been to arrange tickets for volunteers at public performances. The listeners exchange seats during the interval and complete a questionnaire during each half of the performance. The listeners are nearly all specialists in acoustics, the majority being professional acoustic consultants. It was felt necessary to use experts in order to get independent responses over the different scales on the questionnaire. Though this sample may seem biassed, they are not unanimous in their judgements, as will be discussed below.

Two questionnaires have been developed: for speech and music. The speech questionnaire has been used in three theatres and the intelligibility judgements correlate well with some objective measures ($r=0.8$). This paper will concern itself with the more nebulous area of music appreciation and limit the discussions to the results in symphony concert halls. Surveys have been conducted in ten halls containing at least 1500 seats and with volumes in excess of $12{,}000m^3$.

THE QUESTIONNAIRE

The questionnaire is based on scales of the sematic differential type using adjectival descriptors. Its success is based on the human ability to estimate acoustic perceptions using descriptors associated with apparently unrelated senses, eg acoustical warmth. The use of this technique for music listening dates back to the 1960's. The two major studies by Hawkes [1] and at Berlin, described in [2], used 16 and 19 scales respectively. The results of this were subjected to factor analysis which isolated the principal dimensions in which subjects were responding. The present questionnaire can be considered second generation. It is limited to scales which are reckoned to be mutually independent. The questionnaire is illustrated in Figure 1. The final scale of Overall Acoustic Impression is of course expected to be dependent, but the subjective variable(s) on which it is dependent is of particular interest. It was decided that verbal descriptions should offer greater consistency with this scale. The labels were selected from a task performed by twenty subjects. The development of this questionnaire is described in greater detail in reference [3]. A certain degree of training is considered necessary to complete this questionnaire so the questionnaires for the first two performances for each listner have not been used.

VALIDATION OF THE TECHNIQUE

The questionnaire technique can be validated in two ways: whether subjects are scoring on different scales independently but with Overall Impression as a dependent variable and whether significant differences are observed on each scale at the various locations sampled. The independence of judgements on

SUBJECTIVE SURVEY OF BRITISH CONCERT HALLS

Hall _____ Initials _____ Date _____

Seat No. _____ Seat Position _____ L/R

Orchestra _____ Orchestra Size _____

Composer _____ Opus _____

CLARITY — Muddy ——— Clear

REVERBERANCE — Dead ——— Live

ENVELOPMENT — Expansive ——— Constricted

INTIMACY — Remote ——— Intimate

LOUDNESS — Loud ——— Quiet

BALANCE :

	Weak ——— Loud
Treble re. mid-frequencies	
Bass re. mid-frequencies	
Singers/Soloists re. orchestra	

BACKGROUND NOISE :

Inaudible	Acceptable	Tolerable	Intolerable

OVERALL IMPRESSION :

Very poor	Poor	Mediocre	Reasonable	Good	Very good	Excellent

N.B. : scale extremes refer to experience limits

Other comments (continue overleaf if necessary) :

Figure 1. Questionnaire used for music performances.

scales can be assessed by generating a correlation matrix for the individual scales. If all the scales are independent, no significant correlations will exist between them, but we expect significant correlations between certain scales and the Overall Impression scale. We do find the highest correlations with the Overall Impression scale with both 'Envelopment' and 'Intimacy', with a less significant correlation with 'Reverberance' as well. All these correlations are positive as we would expect. There are also correlations between the following pairs of scales with a correlation coefficient $r>0.4$: Reverberance - Envelopment, Envelopment - Intimacy and Intimacy-Loudness. There appears little redundance here since no two scales behave similarly relative to other scales so, for instance, Reverberance is not correlated with Intimacy. The correlation coefficients in the matrix thus appear to validate the design of the questionnaire.

The presence of individual differences in the scales is assessed by subjecting the data to a one-way analysis of variance. If this is done on a hall by hall basis, the variance by hall is highly significant on all scales. The most significant scales (in descending order) are Bass Balance, Reverberance, Intimacy, and Overall Impression; the least significant are Envelopment and Treble Balance. As far as the least significant scales are concerned, there seems to be some difficulty or divergence of opinion surrounding judgement of Envelopment. It is possible that Treble balance may not vary much in these halls. The same analysis conducted on a position within halls basis likewise gives significant results on all scales (except for a marginal failure with the Loudness scale). The judgement of subjects thus appear to be consistent on all scales.

SUBJECTIVE DETERMINANTS OF OVERALL IMPRESSION

From the correlation matrix discussed above, based on 183 questionnaires, the Overall Impression was found to be correlated with 'Envelopment', 'Intimacy', and 'Reverberance'. However, do subjects place the same weight on each of these subjective attributes in establishing an overall judgement? To establish this, correlation matrices have been derived for individual subjects. The 12 subjects who have completed 8 or more questionnaires have been used. For these 12 subjects, 2 have their highest correlation between Overall Impression and Reverberance, likewise 4 with Envelopment, 3 with Intimacy and 3 with Loudness. This suggests either that subjects prefer different subjective attributes or that there is an element of redundancy in the scales (eg subjects might be scoring on all scales as if they were one, known as the 'halo effect'). We can discount the latter because responses of Reverberance, for instance, are wholly uncorrelated with Intimacy and Loudness. So subjects appear to be making overall judgements on different criteria. How many different groups then is it necessary to consider?

In the previous section it was noted that there were correlations between the following pairs of subjective responses: Reverberence-Envelopment, Envelopment-Intimacy and Intimacy-Loudness. If we look at correlations for individual subjects between Overall Impression and other scales, we find that 11 out of 12 subjects have more than one significant correlation with Overall Impression. There are 3 subjects for instance who have significant correlations between Overall Impression and both Intimacy and Loudness. In the case of the 7 subjects who have correlations between Overall Impression and Envelopment, 4 also have correlations with Reverberence as well and 2 with Intimacy and one with

SUBJECTIVE SURVEY OF BRITISH CONCERT HALLS

Loudness. Significantly no-one has a correlation between Overall Impression and both Reverberance and Intimacy. These observations suggest that subjects might be subdivided into those that prefer 'Reverberance' and those that prefer 'Intimacy'. Of the 12 subjects, 5 are found to belong to the 'Reverberance' group and 7 to the 'Intimacy' group. The correlation coefficients for the two groups between 'Overall Impression' and 'Reverberance', 'Intimacy' and 'Envelopment' are given in Figure 2.

	Subjects preferring 'Reverberance'	Subjects preferring 'Intimacy'
Correlation coeff. of 'Reverberance' with 'Overall Impression'	0.73	0.22
Correlation coeff. of 'Intimacy' with 'Overall Impression'	0.26	0.68
Correlation coeff. of 'Envelopment' with 'Overall Impression'	0.63	0.50

Figure 2. Correlation coefficients (r) between 'Overall Impression' and other scales for two groups of subject.

The apparent indifference of those preferring 'Reverberance' to Intimate sound and v.v. is persuasive for this two group model.

It is appropriate at this point to mention that in the Berlin study [2] they found a subdivision between listeners who preferred Clarity and those that preferred Loudness. The responses to Loudness and Intimacy are probably equivalent, but none of the 12 listeners in this survey has chosen Clarity as the predominant criterion. It may be significant that the Berlin study included students training to be recording engineers. Listeners who place comprehension of the musical detail above all room effects are likely to prefer Clarity. There may therefore be at least three groups of listeners in the concert-going public.

COMPARISON BETWEEN SUBJECTIVE AND OBJECTIVE RESULTS

Objective measurements at each listening position have been taken [4]. The following objective measures have been calculated:

1. Reverberation Time
2. Early Decay Time (EDT) [2]
3. Ratio of early to late energy [2]
4. Centre Time [2]
5. Total sound level
6. Early lateral energy fraction (LEF) [5]
7. Initial time delay gap [6]

In the case of measures 1. to 5. a value for bass frequencies (mean of 125 and 250 Hz octaves) and a mid-frequency value (mean of 500, 1k and 2kHz octaves) have been calculated. The mid-frequency early (80ms) and late sound level have also been calculated from 3. and 5. [4]. The corresponding subjective data has been aggregated at each listening position, giving a total of 33 results in 10 halls. This analysis comparing subjective and objective results is not yet fully complete.

The following correlations have been observed by running a stepwise multiple regression program on the three most significant subjective variables and the 14 objective measures. Since all 14 measures are considered in each case, the fact

that meaningful correlations emerge is very encouraging, with regard both to the subjective and objective measurement procedures.

In the case of the subjective variable Reverberance the regression is only with mid-frequency EDT (r=.39), but this improves to r=.47 for the mean EDT across both bass and mid-frequencies. The Berlin study [7] gave the same result.

For the subjective variable Envelopment, the best single correlation is with LEF (r=.43), the measure associated with early lateral reflections. However the addition of the bass total sound level increases the coefficient to r=.60, and with further addition of the mid-frequency EDT the multiple coefficient becomes r=.68. This is a very encouraging degree of agreement but the choice of additional measures is very intriguing. The selection of the bass total sound level supports theories about the importance of both loudness and bass sound for spatial impression [5]. The final inclusion of EDT reflects the contribution of reverberant sound to the sense of space.

Finally with the subjective variable Intimacy there is a good correlation with source-receiver distance (r=-.61), a better correlation with mid-frequency total sound level (r=.67), which is a function of distance [4]. However the best correlation, and sole multiple correlation, is with the mid-frequency early sound level, r=.70. This is a novel result but one that concurs with expectations. The early sound level is the sum of the early energy fraction (expressed in dB) and the total sound level. The former relates to clarity and the latter to loudness, so it appears that clear loud sound is judged as intimate, as one might expect.

Beranek [6] postulated that Intimacy was related to the Initial Time Delay Gap, a result which this study does not support. This study does however indicate the subjective importance of acoustic intimacy. Smaller halls will have high early energy levels and therefore be more acceptable. For a large hall to be acceptable, an adequate number of early reflections are required. For given source-receiver distances, this can be assessed by qualitative analysis of squared impulse responses.

ASSESSMENTS OF INDIVIDUAL HALLS

As an example of the responses on a per-hall basis, Figure 3 shows the mean subjective ratings for the scale Clarity for the ten symphony halls. Maximum clarity is observed in the Royal Concert Hall, Nottingham and the Royal Festival Hall. Both these halls have elements which specifically direct early sound to the seating areas. At the opposite extreme, poorest clarity is observed, predictably in the Royal Albert Hall, followed by the Fairfield Hall, in which early reflections are obscured. The results in Figure 3 support the notion that subjective clarity is a function of the proportion of early energy.

For the Overall Impression judgement it is to be expected, and indeed we find, that the best judged halls score well on both the Reverberance and Intimacy scales.

CONCLUSIONS

There are several arguments for or against subjective testing in the laboratory and in live concerts. Laboratory testing allows greater precision, more control over experimental conditions, probably more subjects and absence of visual distraction but there will always be some limitations in the recording technique. Live concert testing involves no modification of the acoustic stimulus and acoustic stimulae are presented in the appropriate visual conditions. Though the

SUBJECTIVE SURVEY OF BRITISH CONCERT HALLS

Figure 3. Mean responses by hall on the Clarity scale.

KEY

A - Royal Albert Hall,London	L - Philharmonic Hall, Liverpool
B - Colston Hall, Bristol	M - Free Trade Hall, Manchester
C - Fairfield Hall, Croydon	N - Royal Concert Hall, Nottingham
D - St. David's Hall, Cardiff	P - Wessex Hall, Poole
F - Royal Festival Hall, London	R - Barbican Concert Hall, London

precision of live concert testing is generally worse, it is necessary to validate laboratory results in the true environment. Some scepticism was expressed at the outset of this subjective exercise. The results indicate reasonable agreement between subjects in their judgements on the individual subjective scales but that subjects differ in their weighting of these scales to assess their overall preference. Results support the importance of the Early Decay Time and Early Lateral Energy Fraction but introduce as a "new" measure the mid-frequency Early Energy Level.

REFERENCES

1. R.J.HAWKES and H.DOUGLAS 1971 Acustica 24, 235-250. Subjective acoustic experience in concert auditoria.
2. L.CREMER and H.A.MÜLLER (trans. T.J.SCHULTZ) 1982 Principles and applications of room acoustics. Applied Science, London.
3. M.BARRON 1982 Proceedings of the Institute of Acoustics, Conference on Auditorium Acoustics and Electro-Acoustics, Edinburgh. Listener survey of U.K. auditoria.
4. M.BARRON and L.J.LEE submitted to the Journal of the Acoustical Society of America. Energy relations in concert auditoria, I.
5. M.BARRON and A.H.MARSHALL 1981 Journal of Sound and Vibration 77,211-232. Spatial Impression due to early lateral reflections in concert halls: the derivation of a physical measure.
6. L.L.BERANEK 1962 Music,Acoustics and Architecture. John Wiley and Sons,New York.
7. P.LEHMANN 1976 Doctorate Thesis, Technische Universität, Berlin. Über die Ermittlung raumakustischer Kriterien and deren Zusammenhang mit subjektiven Beurteilungen der Hörsamkeit.