Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Room acoustical parameter values at the listener's ears – can preferred concert hall acoustics be predicted and explained?

Magne Skålevik

Brekke & Strand, Oslo, Norway
and
www.akutek.info

PACS: 43.55, 43.55.Gx

ABSTRACT

Room acoustical parameters for concert halls are basically designed to describe significant listening aspects of a room. Correlation between subjective ranking of concert halls and their measured parameter values, averaged over the seating area, have been found. However, results from simulations and measurements indicate that due to spatial variations, very few listeners will actually be in a position where the set of five parameter-averages can be experienced. Therefore, this author have pursued the possibility of explaining subjective ranking of concert halls by objective conditions at listeners' ears, as reported in this paper. It is concluded that Beranek's rank-ordering of nine halls can be explained by objective acoustical conditions at the ears of listeners seated in the better 2/3 to 3/4 of each hall. Explanation degree up to $R^2 = 0.94$ is found with a set of five parameters. Predictability was improved when excluding one of the parameters. Some of the other conclusions are: The ranges of parameter values associated with good listening quality turn out to be strikingly large in terms of noticeable differences. Since it is crucial to be able to predict subjective quality during concert hall planning, the search for significant parameters and optimal combinations of these should continue in further work. More halls should be included in an extended study. Linear regression should be handeled with care.

INTRODUCTION

During the 20th century, the number of room acoustical parameters used to describe the acoustic qualities of concert halls, as perceived by listeners, exploded. One of the problems with the large amount of parameters is that they are more or less inter-correlated, i.e. they do not form an orthogonal or linear independent set of variables. By the end of the century, one arrived at some international consensus regarding the limited selection of 5 listener aspects and their corresponding ohysical quantities, as expressed by the standard ISO 3382.

In the search for the more significant parameters and their preferres values, there is being used two major, but quite different, methods. One is the method of controlled listening tests, asking respondents for subjective preference to variations in physical quantities. Another metod is to study the values of the parameters in halls of varying reputation in terms of acoustics, in particular the most highly rated halls. Following the latter approach, Beranek (2003) presented a rank-ordering of 58 concert halls according to their acoustical quality, based on interviews of conductors, music critics and well-travelled music aficionados [1]. He pointed at a set of six significant parameters as being possible orthogonal elements to be combined in a single rating number for the halls.

In a different study, reported in this paper, this author has compared objective listening quality based on a set of 5 parameters measured in 116 source-receiver combinations in 10 different halls measured by Gade[2], with subjective ranking of the halls based on Beranek's rank-ordering.

BACKGROUND

This section explains the significance of studying parameters in each measured point instead of the hall average.

Hall averages values vs values at listeners' ears

It has been common to describe the acoustical qualities of a hall by its average parameter value, e.g. the average reverberation time (RT) measured with different source-receiver positions. While the hall average could be an adequate representation of a global parameter like the RT, this is not evident for the parameters in general since most of them are spatially dependent. The parameter for sound strength, G, tend to change by at least 1dB per 10 meters as source receiver distance changes, even in concert halls with preferred reverberance. The dryer the hall, the more does G change in dB per meter. In dryer halls the rate of change is even more. Closer to stage, where direct sound dominates over reverberant sound, both sound strength G and clarity C will increase dramatically. In many halls, G measured over the whole seating area may vary in the range of 0 to 10dB. In terms of just noticeable differences (JND), the latter corresponds to a variation of 10 JND. Similar noticeable variations in parameters over the seating area in concert halls can be seen in general. Therefore, it is to be expected that parameter values at listeners' ears are noticeably different from the hall average.

Skålevik (2008) [4] reported results from a computer simulation study indicating that in the case of Musikvereinsaal in Vienna, only 9% of the listeners experience acoustic conditions that can be described by the 5 hall averages of parameters corresponding to the set of 5 subjective listener aspects in ISO 3382, when reapective JND's are taken into account. This means that the remaining 91% of the listeners in Vienna experiences noticeably different conditions than the average conditions. Further work showed that the reputation and quality rating of the hall could be better explained by the 5 parameters when accepting seats that varied noticeably from hall average [5]. However, some problems remained to investigate: How much can a parameter value at a listener's ear deviate from hall average without affecting the listener's impression of the hall, and if large deviation is acceptable, can we still use parameters to explain the rank-ordering of halls? The latter problem in particular is pursued in this paper.

A DIGRESSION TO LINEAR REGRESSION – A STRAIGHT LINE INTO A PITFALL?

This section emphasises the importance of awareness when related using linear regression and studying linear trends when dealing with listener aspects and parameters that have optimum values rather than the-more-the-better.

A thought experiment, or three

Though linear regression is a very effective tool when trying to explain variance in data, or to point out dependency between variables, it should be handeled with care. Variables in general exhibit linear tendency when the variation in the data under study is sufficiently small.

Some of the pitfalls of linear regression can be demonstrated by the following thought experiment: Researcher A wants to investigate the behaviour of the quite recently suggested-parameter Blending Time (BT), any similarities with Reverberation Time purely accidental. He runs listening tests, exposing a group of respondents to the same music from the same musicians, but in different seats in different concert halls, obtaining a BT variation in the range of 1.4s to 2.1s measured at listeners' ears. Each test run results in a single-number from each respondent, on a scale given by researcher A and dicussed with the respondents beforehand, and plotted in Figure 1. Researcher A concludes that the data is good, especially since the wide range of BT obtained in the test covers the whole range of BT found in concert halls, from the top-ranked ones to the bottom ranked ones. The convincing result from the test is that 92% of the variation can be explained by a straight line, indicating that Preference can be predicted from Preference~1.33*BT-1.66. For example, at BT=2.0, preference is equal to 1.0, on a preference scale arbitrarily chosen by researcher A, however allowing higher values.

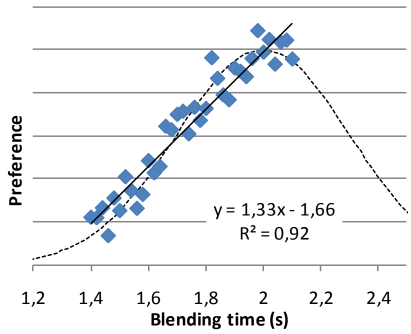


Figure 1 Plot of Preference vs Blending Time, and a linear trend found by researcher A.

Researcher A presents the inspiring results at a world-wide acoustical conference in Australia, unaware of the next speaker, namely researcher B. Equally unaware, Mr.B has run listenig tests similar to those run by Mr. A, only a different selection of halls, in seats where BT ranges from 1.8s to 2.2s. The best fit line to the data is a horizontal, providing zero explanation to the variance of the data, Figure 2. Since the halls involved are reported to have very different overall acoustical qualities, Mr.B must conclude that BT is not a critical parameter, adding no further explanation to the perceived differences of the halls. Needless to say, there was an interesting discussion during lunch break afterwards.

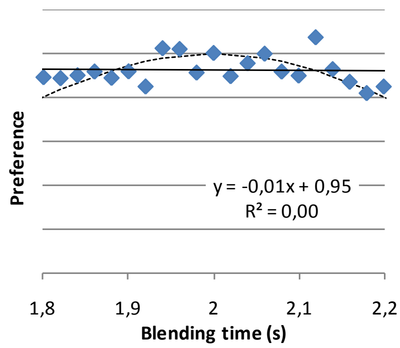


Figure 2. Plot of Preference vs Blending Time, found by researcher B.

As if this was not enough acoustical excitement for a day, the first speaker after lunch was researcher C. She had been leading an extensive investigation of how listeners' Preference varied as this recently suggested parameter, Blending Time, varied in the very wide range of 1.2s to 2.5s. The listening tests were equivalent to those run by Mr.A and Mr.B. In contrast to the two former, Mrs C reported that a best fit line would explain merely 25% of the variance in the data obtained. However she had found that Preference vs BT best could be explained by a characteristic bell-shaped Gauss-curve centered at BT=2.0s, and with a standard deviation of 0.3s. The data scattering around the Gauss-curve can be explained by ± 5% just noticeable difference when respondents detected Blending Time, and by up to ± 10% randomness in respondents' preference that could occur even if BT did not change.

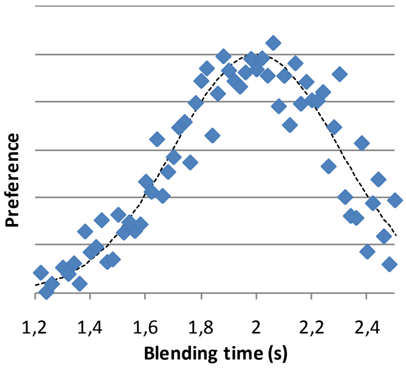


Figure 3 Plot of Preference vs BT, found by researcher C to best fit a Gauss-curve with μ=2.0s and σ =0.3s

Collecting the thougths

Summing up the results of the three investigations by researchers A, B and C, there are demonstrated several pitfalls to be aware of when following the path of linear regression. Note that without knowledge about the underlying Gauss-curve revealed by C, the conclusions by A and B each seem thrustworthy enough, but only until they share their results. The importance of international congresses is demonstrated.

- Even if result A is valid inside the investigated range of variation, the error may be large at the extremes (e.g. for BT=2.0-2.1 in Figure 1), and the trend line is not valid outside the range without furter proof. The range of data values is quite wide, leaving a thrustworthy impression, and the rising trend curve of high correlation indicates falsely that the more is the better. For the range of 1.8s to 2.1s, the result is incompatible with B.

- From research B, we see that around its optimal value, a parameter can vary relatively much (±2 jnds) without affecting preference, even if overall quality varies. This does not justify judging the parameter to be insignificant, as proved by the results from research C. Analog to this, the fact that the temperature may be close to 18 degrees Celcius in most concert halls, good and bad, does not mean that temperature does not matter to the concert experience. Neither if values of 14 and 22 are found in both good and bad halls. Try turning off the heat in mid-winter. As long as a parameter is sufficiently close to its optimum value, we are happy. If not, the hall will certainly get a lower rank.

- Research C demonstrates clearly the importance of studying a wide range of parameter values, trying to search for possible optimum values, rejecting any false indications that the-more-is-the-better, like the one from research A.

Multiple regression – multiple pitfalls?

A multiple regression result could lead to a single number prediction for acoustical ranking of a concert hall, e.g. on the following form, based on the 5 consensus aspects mentioned above:

24 – 8.3*EDT – 1.5*G + 0.1*C – 2.3*LF + 1.0*G,late

However, this approach could lead to some unsetteling results. The formula introduces the idea of trading one aspect for another. For example, if comparing two seats, a 0.3s shorter EDT would compensate for 2.5dB weeker G,late, if the other three parameters are equal. Besides, stronger G would result in better ranking (smaller number), if the other parameters remain unchanged. However, we know that in practice, certain combinations are more likely than others. Toward the stage, G and G,late will rise, C will rise dramatdramatically, while LF will decrease.

The formula does not reflect the fact that all of the 5 parameters involved have optimum values, like G can become too strong, and EDT can become too long.

In the solving of the main problem of this paper, the multiple regression method is not being used. Instead, preferred parameter value ranges from top ranked halls are being used as a criterion for listening quality, as will be explained below.

SEAT RANKING METHOD

In the context of this paper, the acoustical quality of the seat and the acoustical conditions at the ear of the listener sitting in the chair is considered one and the same. Since measuring at every seat in a hall would be very resource consuming, a smaller representative selection of seats in each hall will form a databasis of statistical analysis.

The procedure for ranking concert hall seats after their acoustical quality as chosen in this study is quite straightforward, and can be described by the following steps.

1. Make a proper selection of measurement positions that each represents a group of seats in the hall

2. From measured impulse responses, calculate the single number values of the five parameters EDT, G, C, LF and G_{late} , se Annex. Note the use of mid-frequency G_{late} to describe Envelopment[1], with JND=1dB, in difference to ISO 3382-1.

3. For each of the five parameters, from the values found in the reference hall(s), in this case 12 positions in Musikvereinsaal Vienna and 10 positions in Concertgebouw Amsterdam, exclude the higest and the lowest value, and let the range of the remaining values define the quality criteria

4. Reward each measured seat, in every hall, with one point for each parameter that satisfies the criteria defined above, resulting in a total point reward P ranging from zero points to five points for each measured seat group

5. Calculate the rank number 6-P for each seat group

6. For each hall, calculate the **nominal seat rank** of each hall. This shall be a single rank-number that statistically represents its population of seats, e.g. by the average seat-rank, or the X-percentile, i.e. the lowest rank of the seats remaining when the worst X% of the seats in the hall are neglected. X in the range of ¼ to 1/3 is suggested. Note that this choice of statistic descriptor is equivalent to assuming that Beranek’s ranking is based on experienced concert goers who will generally avoid the poorest X% of seats in every hall.

[1] The choice of G_{late} instead of the late lateral energy level (LG80) to describe Envelopment is due to evidence that late energy from all directions contributes more or less to perceived envelopment, Beranek (2008)[7].

7. Compare the objective rank-number calculated above with some subjective rank-ordering, in this case the Beranek rank ordering

8. Calculate the Pearsons R^2 correlation between objective and subjective ranking as a descriptor of explanability

9. By trial and error, test the effect of removing extreme values in the quality criteria basis in 3, or the percentile in 6, then repeating the next steps to see whether this provides more objective explanation to the subjective ranking

INPUT DATA

Occupied halls vs empty halls

It is assumed that the Beranek rank ordering is related to acoustical conditions in occupied halls, even if many of the interviewed sources are conductors having major experience from the halls in their empty condition as well.

Since the required amount of measurement data available is from empty halls only, several methods have been tried in order to provide data for the occupied conditions. In practice, it would be very valuable if prediction methods could provide ability to explain ranking of halls, since predictability is crucial when planning new halls or when halls are refurbished, corrected for acoustical flaws, or changed for non-acoustical reasons. Six input data sets provided in different ways have been tested out, and their explanability in terms of (R^2) has been studied.

Measurements of the five parameters in empty halls are the values reported by Gade[2].

Data sets 1 thru 4 have LF values as measured in empty hall, assuming that increased absorption in the seating area will affect the nominator and the denominator of the early lateral fraction in a similar manner, thus largely leaving the fraction unchanged. In data set #1 and #2, EDT, G, C and G_{late}, are calculated from volume (V), global reverberation time (RT) and source-receiver distance (r) by Barron's Revised Theory, i.e. the TVr-predictor[12]. In #1 RTs are measured values from Beranek [8], and in #2, RTs are as predicted by Odeon 10. In #3 and #4, EDT, G, C and G_{late} are based on measurements from empty halls by Gade, only corrected by the noticeable changes corresponding to changes in global RT from empty hall to occupied hall. In #3, the RTs are from Beranek (as in #1), and in #4, the RTs are as predicted by Odeon 10 (as in #2).

In data set #5, all five parameters are based on values measured in empty hall, but corrected by the noticeable changes from emty to occupied conditions as predicted by Odeon 10. Finally, data set #6 are ordinary predictions by simulations in computer models of occupied halls.

The six occupied hall data sets are obtained as given in Table 1.

Ten halls

For this study the selection of halls are simply the 10 halls that are found in both Gade's 11 halls survey, and in Beranek's ranking of 58 halls, see

Table **2**. The 10 halls vary significantly in volume and shape, see models in Annex. For example, the width of the halls varies in the range from 20 meters to 55 meters, the splay (angle between side walls) from 0 to 70 degrees, and the floor-rake from 5 to 20 degrees. The total number of measurements (and corresponding source-receiver positions) from the 10 halls are 116.

The rank order with the rank number in the leftmost column follows directly from Beranek's rank-ordering of 58 halls.

Table 1. Six input data sets, providing data for occupied concert hall conditions in six different ways; Five parameters; 116 source-receiver positions, r.

#	EDT, G, C and G_{late}	LF
1	Calculated from volume, global RT and r by Barron's Revised Theory; Measured RTs from Gade, Beranek and Barron	Measured, empty hall
2	Calculated from volume, global RT and r by Barron revised theory; predicted RT from Odeon	Measured, empty hall
3	Measured by Gade in empty hall and corrected by <u>changes</u> corresponding to change in global RT from empty to occupied hall, predicted by Revised Rheory; Measured RTs from Gade, Beranek and Barron	Measured, empty hall
4	Measured by Gade in empty hall and corrected by <u>changes</u> corresponding to change in global RT from empty to occupied hall, predicted by Revised Theory; Predicted RTs from Odeon	Measured, empty hall
5	Measured by Gade in empty hall and corrected by <u>changes</u> from empty to occupied conditions predicted by Odeon	As EDT, G, C and G_{late}
6	Odeon simulation	Odeon simulation

Table 2. The ten halls

Rank	Concert hall	Volume m^3	RT occ (s)	Beranek Ranking
1	Musikverein, Vienna	15000	2,0	1
2	Concertgebouw, Amsterdam	19000	2,0	5
3	St David, Cardiff	22000	2,0	10
4	Gasteig, Munich	30000	1,9	19-39
4	Konserthus, Gøteborg	12000	1,6	19-39
6	Festspielhaus, Salzburg	15500	1,5	40
7	Liederhalle, Stuttgart	16000	1,6	41
8	Usher, Edinburg	16000	1,3	44
9	Royal Festival Hall, London	22000	1,5	46
10	Barbican, London	18000	1,7	56

RESULTS

Results presented are according to the procedure described in the ranking method paragraph above. The nominal seat rank of the hall is defined equal to the rank of the lowest ranked seat, i.e. the calculated X-percentile, when X% of the least good seats in a hall are excluded. Then 1-X is the percentage of seats that satisfies the nominal seat rank of the hall. The nominal seat rank of the hall is then an objective rank of the hall, measuring the listening quality in the hall. Maximum correlation R^2 between objective ranking of seats and subjective ranking of halls are obtained iteratively by varying the percentage of lower-ranked seats that are neglected. The optimisation procedure is carried out individually for each of the six input data sets.

Early results in an iteration process

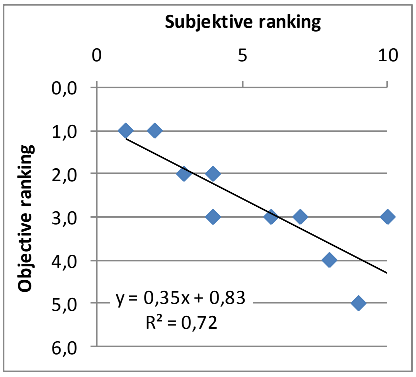


Figure 4: Input data set #2 provided R^2=0.72 when excluding 45% of least ranked seats in each hall

Examples of initial results are presented by plot in Figure 4 and numercal data in Table 3. Figure 4 is a plot of subjective hall ranking (Beranek) versus objective hall ranking, with seat ranking obtained from input data set #2.

Table 3. Seat rank from input data sets #2 and #3 when given percentage of low-ranked seats are neglected; Beranek subjective rank-order in leftmost column.

	Input data set	**#2**	**#3**
1	Musikverein, Vienna	1	1
5	Concertgebouw, Amsterdam	1	2
10	St Davids Hall, Cardiff	2	3
19-39	Gasteig, Munich	3	3
19-39	Konserthus, Gøteborg	2	2
40	Festspielhaus, Salzburg	3	4
41	Liederhalle, Stuttgart	3	4
44	Usher Hall, Edinburg	4	5
46	Royal Festival Hall, London	5	5
56	Barbican, London	3	3
	Seats neglected in each hall	45 %	27 %
	Correlation R^2	**0,72**	**0,60**

The maximum correlation between objective and subjective ranking in input data set #2, R^2 =0.72, occurred for X=45%, i.e. when neglecting 45% of the lowest ranked seats in each hall. However, X much larger than ¼ or 1/3 intuitively seems to contradict the assumption that the subjective ranking of halls must be based on the experience from a large number of listeners having been seated in most parts of the hall, maybe exept under overhangs, close to walls, and behind columns etc. Whenever R^2 reaches its maximum for X>25%, then X should be chosen in order to make (1-X)*R^2 large to maintain relevance of the result.

Table 4 presents the correlation R^2 between objective rank of the hall, based on acoustical quality at listeners ears, and Beranek's subjective rank of the hall, for input data sets #1 thru #6; R^2 can be interpreted as the part of explanation of Beranek's ranking that relates to acoustical quality at listeners' ears when X percent of the poorest seats are neglected. Then 1-X must be the percentage of seats contributing to the subjective ranking of the hall, and (1-X)*R^2 indicates the amount of relevant explanation of the subjective hall rank that is caused by acoustical conditions at listeners' ears.

In input data set #1, up to R^2 =0.75 was found when neglecting 38% of the lowest ranked seats in each hall. Optimising (1-X)*R^2 led to R^2 =0.73 with 36% seats neglected. This was the highest degree of relevant explanation that could be provided by any of the six input data set.

Table 4. Ten halls. Correlation R^2 between objective rank of the hall and Beranek's subjective rank of the hall; 1-X is the percentage of seats satisfying the objective rank of the hall; Input data sets #1 thru #6. See text.

10 halls	#1	#2	#3	#4	#5	#6
R^2	0,73	0,69	0,60	0,65	0,51	0,44
1-X	64 %	64 %	73 %	74 %	62 %	77 %

However, from the plot in Figure 4 we see that Barbican, London deviates in particular from the regression line, with its bottom subjective rank (10) contradicting a medium objective rank of 3. The reason for this is not clear. A possible explanation may be the poor low frequency response in the hall, an aspect that is not taken into account by the single number frequency averaging (500Hz and 1000Hz) applied in this study, according to ISO 3382. It is therefore natural to investigate the resulting correlation with the hall excluded from the correlation analysis.

Excluding one hall – nine halls left

After excluding the hard-to-explain hall commented above, the calculation process is repeated, and the result on the same form as in Table 4 is presented in Table 5.

Table 5. Nine halls. Correlation R^2 between objective rank of the hall and Beranek's subjective rank of the hall; 1-X is the percentage of seats satisfying the objective rank of the hall; Input data sets #1 thru #6.

9 halls	#1	#2	#3	#4	#5	#6
R^2	0,76	0,89	0,90	0,94	0,82	0,61
1-X	88 %	64 %	74 %	74 %	60 %	73 %

Table 5 shows that excluding the hard-to-explain hall results in far better explanation for all six input data sets, and #4 in particular. Objective seat quality ranking calculated from input data set #4 is plotted against subjective hall-ranking in Figure 5 and presented numerically in Table 6.

Data set #4 provides parameter data corresponding to occupied hall condition by correcting measurements by Gade with the same source-receiver positions, as described in Table 1. The maximum correlation of R^2 = 0.94 occurs when the seat ranking is based on the better 74% of the seats. Close to this result is the one from input data set #3, with #2 not far behind.

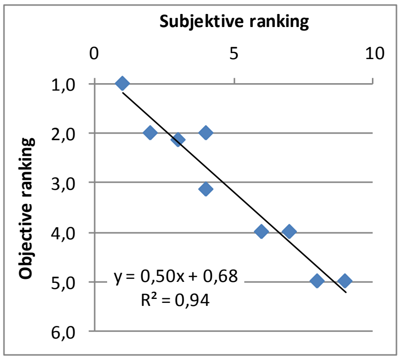


Figure 5. Nine halls: Objective (seat) rank of hall plotted against subjective rank of hall, input data set #4. Rank values different from whole numbers are due to percentile computation.

Qualifying parameter criteria

An interesting by-product of the process that led to the results above, are the qualifying criterias applied to each parameter in the calculation of the objective seat rank. These are defined by the value range from 12 source-receiver positions in Musikvereinsaal and 10 in Concertgebouw after eliminating the highest and the lowest value. Since the six data sets are generally different in all halls, it follows that the criteria ranges are generally different in each of the six data sets, as can be seen in Table 7.

The equal LF-values in #1 thru #4 are all the same measurements by Gade, according to definition of the data sets.

Table 6. Nine halls: Objective rank result from data #4.

Subjective rank	Nine halls	Objective rank
1	Musikverein, Vienna	1
2	Concertgebouw, Amsterdam	2
3	St Davids Hall, Cardiff	2
4	Gasteig Philharmonie, Munich	3
4	Konserthus, Gøteborg	2
6	Festspielhaus, Salzburg	4
7	Liederhalle, Stuttgart	4
8	Usher Hall, Edinburg	5
9	Royal Festival Hall, London	5
	Seats neglected in each hall	26 %
	Correlation R^2	**0,94**

Table 7. Qualifying criteria applied in the ranking of seat quality for the five parameters and the six input data sets, defined by value range from 12 measurements in Musikvereinsaal and 10 measurements in Concertgebouw after eliminating the highest and the lowest value.

		EDT	G	C	LF	GL
#1	max	2,0	8	4	0,24	3
	min	1,4	2	-1	0,10	0
#2	max	2,1	8	4	0,24	4
	min	1,5	3	-1	0,10	0
#3	max	2,3	8	2	0,24	4
	min	1,5	2	-5	0,10	0
#4	max	2,3	8	2	0,24	4
	min	1,5	3	-5	0,10	0
#5	max	2,3	7	4	0,26	3
	min	1,3	1	-5	0,06	-2
#6	max	2,3	7	5	0,31	1
	min	1,5	1	-2	0,01	-2

INTERPRETATION, COMMENTS, DISCUSSION

Explainability

Results presented above indicate that three of the six data sets describing listening conditions in nine occupied concert halls explain 89-94% of the subjective ranking of the halls, according to Beranek's rank-ordering. This high degree of explanation 89%, 90% and 94%, is based on a reasonably large part of the seating area, 64%, 74% and 74% respectively. Intuitively it appears plausible that the respondents usually will avoid the presumably least good seats, and rather chose a seat among the better 64-74% of seats in every hall.

The highest degree of explanation in this study, R^2 = 0.94, occurs with input data set #4, based on a seating percentage of 74%. It is interesting to note that the data is based on parameters measured by Gade in empty hall, then corrected by the noticeable differences from empty to occupied conditions by means of Barron Revised Theory, with global RTs in empty and occupied condition taken from average RTs simulated in Odeon 10 for the actual measurement positions. The only difference between these data and those in #3, is that in #3 the global RTs are measurements available in the literature. Note that this difference results in a difference in correlation of merely 0.04, based on the same 74% of seats. Common for these two is that correlation is high and that they are both based on measurements corrected by means of Revised Theory.

Predictability of subjective quality

Predictability is crucial in concert hall design, either in context of a new hall or changes in an existing hall. Therefore it is important to see what the results in this study may imply to possibilities of predicting the subjective quality of a new hall, or difference in subjective quality due to ammendments, during the design phase. It is a bit disappointing that the only data set based on prediction “from scratch”, #6, provides the lowest degree (R^2 = 0.61) of explanation of the subjective ranking of nine existing hall. This is interpreted to a predictability of 61% with this prediction method, which may be too unreliable to the client.

Input data set #2 would have been another example of prediction from scratch if it had not been for the LF parameter, which by today can be predicted only with great uncertainty with ray tracing methods. On the other hand, LF proves not to be among the most critical parameters in this study. Thus the effect of excluding LF data was investigated. Results in Table 8 give the following indication of possible predictability with the current methods:

- #6 Concert hall design "from scratch": Subjective quality of new hall and differences in subjective quality from amendments of existing hall can be predicted with 65% certainty by simulating EDT, G, C and G_{late} in at least 12 representative (well-distributed) positions with Odeon 10

- #2 Concert hall design "from scratch": Subjective quality can be predicted with 82% certainty from EDT, G, C and G_{late} calculated by Revised Theory with the input of volume, source-receiver distances from at least 12 representative (well-distributed) positions, and global RT predicted by Odeon 10

- #4 Change to existing concert halls: Difference in subjective quality after rebuilding can be predicted with 90% certainty by measuring EDT, G, C and G_{late} in at least 12 representative source-receiver positions before change, correcting for differenses computed by Revised Theory, using global RTs in the before-condition and global RT in the after-condition as predicted by Odeon 10

Table 8. Nine halls, similar to Table 5, but LF data excluded.

9 halls ex LF	#1	#2	#3	#4	#5	#6
R^2	0,71	0,82	0,80	0,90	0,80	0,65
1-X	82 %	70 %	67 %	70 %	68 %	72 %

Six representations of listening qualities

Table 7 reveils differences, as must be expected, in the six data sets that provide physical representations of listening quality in six different ways. One should not make the mistake that the differences means that some data sets must be right and some must be wrong, while some are more right than others.

Here, the six data sets and their different values and criterias should be considered six different representations of listening quality. Measurements would be just a seventh representation. In particular should never data from one set be compared with criteria from a different data set.

Models vs Measurements

We tend to think that measurements are truth while the results we get from computations are guesswork, implicitly more or less false. All six data sets are a mix of measurement and computations. Thus it is tempting to judge their reliability after their content of measurements. However, since the goal is to predict the subjective response of listeners, and not to predict what will physically be measured, the data can only be judged after their ability to do just that – predict subjective respons. This does not mean that it is not important to be able to predict close to measurements, it is simply not the objective of this study.

When investigating apparent connections between variables, like the one between subjective quality and objective quality, science has to deal with uncertainties. One way to proceed is to see the connection as a chain with links that can be investigated one by one, e.g. Model – Physicality – Measurement – Hearing – Perception – Experience which introduces five links and a sum of five uncertainties. A second, very different, approach is to treat the links between Model and Experience as a so-called black box, inside which we do not really care what happens. Then the chain has reduced to Model – Experience involing just one (though complex) link with its single (big or small) uncertainty. The study reported in this paper follows the second approach. If a model can be found that correlates well enough with experience, then the uncertainty is acceptable, and the aim of the study is reached: Preference for concert halls can be explained by parameters.

Detailed measurements in occupied halls would be valuable in order to increase the understanding of listening experience in concert halls, and the same goes for details in our hearing, our perception, sociological aspects, and so on. However, increased understanding is not equivalent to improved explanation in the scientific sense. The black box can provide an explanation even though the processes inside is not understood. Concert halls are huge investments and their quality should please musicians and audiencin at least a hundred years. The crucial matter is to be able to predict the listening experience in the completed hall, during the design phase before measurements are available.

Accepted range of parameter values

An interesting by-product from this study is the range of parameter values from the reference halls in Vienna and Amsterdam, chosen as the qualifying criteria for each listening aspect. Upper and lower limits for those parameter values assumed to be acceptable to listeners are presented in Table 7. In terms of span in just noticeable differences, the ranges are presented in Table 9.

One of the motivations for this study was the fact that hall averages are experienced by a small minority of listeners, and it was not clear how large noticeable differences can be accepted. It is quite surprising that the acceptable ranges are so wide, since this intuitively seems to have the consequence that every hall would have a large amount of seats with top listening quality. In spite of this, we see that many seats will fail to satisfy more than one or two of the parameter criteria at once, leaving them with seat rank 2 or 3.

Table 9. Span of parameter-values associated with good listening quality in terms of just noticeable differences (JNDs, see Annex); Five parameters, six input data sets. Data set with highest explanation degree (0.94) in bold.

jnds	EDT	G	C	LF	GL
#1	7	5	5	3	3
#2	7	5	5	3	3
#3	8	6	7	3	4
#4	**8**	**6**	**7**	**3**	**4**
#5	11	6	9	4	5
#6	8	6	7	6	4

FURTHER WORK

Further work should include testing of other parameter combinations, in aim for higher predictability. In particular, a potential improvement is seen in reducing uncertainty of LF-predictions.

It is natural to extend the study to include more of the 58 halls in Beranek's rank-ordering.

Since the ranges of parameter values associated with good listening quality are quite large, we might see chairs "strange parameter combinations. For example: A seat with medium values in all parameters getting top rank is intuitively acceptable, but it is a bit unsettling that a chair with low reverberance, low strength, little apparent source width and little envelopment should get the same top rank. Some combinations within the qualifying ranges are not very likely to find, e.g. high G, low G_{late} and low C, leading to the problem of dependency between parameters.

Further work will include a continued search for possible golden combinations [5] in the preference halls, in order to see if these can provide quality criteria that improve the explanation of subjective listening quality. One possibility is to introduce other criterias in front of hall than in the rest of the hall, or even a further division: front-middle-back.

The Loudness-parameter L suggested by Barron [11] is expected to predict subjective loudness of sound more accurate that G strength. However, brief testing from L, of curiosity, did not lead to improved explanation in the context of this study. Since L allows G to drop at seats far from stage, apparently the large, week halls were judged better with L than with G. This should be tested more thoroughly in furter work.

CONCLUSION

It is concluded that Beranek's rank-ordering of nine halls can be explained by objective acoustical conditions at the ears of listeners seated in the better 2/3 to 3/4 of each hall. Explanation degree up to R^2 = 0.94 is found with a set of five parameters.

Predictability of listening quality, relevant for prediction methods available in concert hall planning, is improved when excluding the LF-parameter. The set of four remaining parameters provides predictability up to R^2 = 0.90 by Revised Theory with hall volume, source-receiver distances, and reverberation time calculated with prediction tool Odeon 10. The reason for excluding LF is that methods for predicting LF with low uncertainty are not available – not that LF and apparent source width are insignificant to listening quality.

The ranges of parameter values associated with good listening quality differ from one parameter to another, but are quite large, from 3 to 8 just noticeable differences (jnd).

Linear regression should be handeled with care, especially when comparing parameters with subjective quality that requires an optimum range rather than the-more-the-better.

Further work should include testing of other parameter combinations, in aim for higher predictability. In particular, a potential improvement is seen in reducing uncertainty of LF-predictions.

Consequences of the rather large qualifying ranges of the parameters seen in this study should be studied carefully.

It is natural to extend the study to include more of the 58 halls in Beranek's rank-ordering.

REFERENCES

[1] L. Beranek, "Subjective rank-orderings for concert halls", Acta Acoustica u w Acoustica Vol 89 (2003)

[2] A. C. Gade, "Acoustical Survey of Eleven European Concert Halls", DTU Report No.44, 1989, ISSN 0105-3027

[3] Hiroyoshi Shiokawa and Jens Holger Rindel, "Comparisons between Computer Simulations of Room Acoustical Parameters and Those Measured in Concert Halls", Report of the Research Institute of Industrial Technology, Nihon University Number 89, 2007, ISSN 0386-1678

[4] M. Skålevik, "Room acoustic parameters and their distribution over concert hall seats", Proceedings of Auditorium Acoustics 2008, Institue of Acoustics, Vol.30. PT.3. 2008, ISSN 1478-6095, http://www.akutek.info/Papers/MS_Parameters_Distribution.pdf

[5] M. Skålevik, "Room acoustic parameters and their distribution over concert hall seats", Conference presentation, Auditorium Acoustics 2008, Oslo 2008. http://www.akutek.info/Presentations/MS_Parameters_Distribution_Pres.pdf

[6] ISO 3382, ''Acoustics — Measurement of the reverberation time of rooms with reference to other acoustic parameters'' (1997): Annex A of later revisions.

[7] L.Beranek, Concert Hall Acoustics—2008, J. Audio Eng. Soc., Vol. 56, No. 7/8, 2008 July/August http://www.leoberanek.com/pages/concerthalls2008.pdf

[8] L. Beranek: Concert Halls and Opera Houses, second edition, Springer, 2003

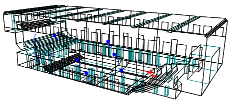
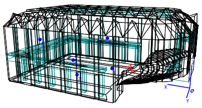
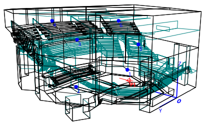
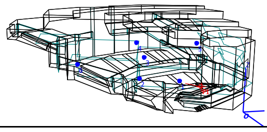
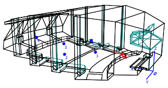
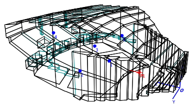
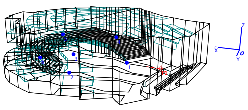
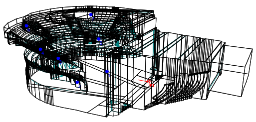
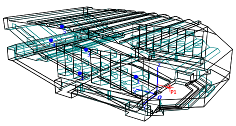
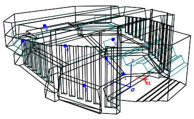
[9] M. Barron, Auditorium Acoustics and Architectural Design (E & FN Spon, London, 1993).

[10] M. Barron, Using the standard on objective measures for concert auditoria, ISO 3382, to give reliable results, Acoust. Sci. & Tech. 26, 2 (2005)

[11] M. Barron, "Objective assessment of concert hall acoustics", Proceedings of Auditorium Acoustics 2008, Institue of Acoustics, Vol.30. PT.3. 2008, ISSN 1478-6095 http://www.akutek.info/Papers/MB_Objective_Assessment.pdf

[12] M. Skålevik, "Reverberation Time, the mother of all acoustical parameters", Proceedings of ICA 2010 http://akutek.info/Papers/MS_reverberationtime_BNAM2010.pdf

ANNEX – HALL MODELS

Concert hall	Odeon Model
Musikverein, Vienna	
Concertgebouw, Amsterdam	
St David, Cardiff	
Gasteig, Munich	
Konserthus, Gøteborg	
Festspielhaus, Salzburg	
Liederhalle, Stuttgart	
Usher Hall, Edinburg	
Royal Festival Hall, London	
Barbican, London	

THE FIVE LISTENER ASPECTS AND PARAMETERS IN THIS STUDY

Note that the parameter chosen to describe Listener envelopment LEV is the late Sound Level G_{late}, in contrast to Late Lateral Sound Level LG in ISO 3382-1

Subjective listener aspect	**Acoustic quantity**	**Just Noticeable Difference (JND)**
Subjective level of sound (SOUND LEVEL)	Sound Strength G, in dB	1dB
Perceived reverberance (REVERBERANCE)	Early Decay Time, EDT, in s	5%
Perceived clarity of sound (CLARITY)	Clarity, C80, in dB	1 dB
Apparent source width, ASW	Early Lateral Energy Fraction, LF	0.05
Listener envelopment LEV	Late Sound Level, G_{late} , in dB	1 dB