Audio Engineering Society

Convention Paper 10378

Presented at the 148th Convention, 2020 June 2-5, Online

A new approach to predicting listener's preference based on acoustical parameters

Peter Russell Critchell[1], Dr Ludovico Ausiello[1]

[1] *Southampton Solent University, East Park Terrace, Southampton, SO14 0YN, Hampshire, United Kingdom*

Correspondence should be addressed to Author (petercritchell@outlook.com)

ABSTRACT

Since its conception, the study of room acoustics has explored the links between acoustical parameters and subjective preference. While there have been attempts to combine such metrics, e.g. Frick's combination of six acoustical parameters to predict '*acoustic quality*', no reliable method for prediction of listeners' preference has been univocally ascertained[1] or included in any ISO standard[2]. In this study an alternative perspective is presented - to derive a simple descriptor, '*Preference Rating'* (PR), through meta-analysis of metric-preference relationships, within the context of Rock and Pop venues. A statistical approach has been taken to determine the relative importance of a chosen set of factors in the form of mathematical weights.

Results of this pilot study indicate that preference may be predicted by using eight acoustical parameters: Reverberation Time (RT), Bass Ratio (BR), Tonality (TN), Definition (D50), Early Decay Time (EDT), Bonello Distribution (MD), Background Noise and Surface Diffusivity Index (SDI). Quantitative data and subjective evaluation data describing 20 venues (provided by Dr. Adelman-Larsen[3]) were used to validate this new approach and showed strong correlation in 85% of the scenarios. This suggests that the rationale behind the presented method is meaningful and can be used to set a base upon which further testing and development can be conducted to improve the reliability of such empirical approach.

1 Introduction

There are several acoustical parameters defined by ISO 3382[2] to determine the objective performance of a room, nonetheless listener preference is ultimately the true metric by which venues are compared by audience and users. However, the variables that change from venue to venue are physical, thus objective metrics and physical dimensions must influence preference.

Sabine provided us with the first empirical equation to predict the Reverberation Time (RT) for a room[4], and from it Beranek[5] defined ideal reverberation times for concert halls so that listeners could enjoy the most 'preferable' venues for symphonic music. Similarly, other correlations bond a single objective acoustical parameter or measurement to one aspect of a person's preference, so it is therefore plausible that summation of these objective factors, with a weighting could predict a person's overall preference.

This concept has been explored in Beranek's later work on Interaural Cross-Correlation Coefficient (IACC) which comprises many factors and highlights that halls with lower IACC values tend to perform better - but does not tell the whole story[6]. Expanding upon this, there have been attempts to further combine metric's for specific contexts, such as Fricke's amalgamation of Beranek's acoustical parameters; G_{mid}, BR, EDT, IACC, SDI and RT to provide an alternative approach to concert hall design[1]. However, the results of this study are inconclusive, determining that while some connection is present, it is not reliable enough for consistently accurate prediction of preference.

This study proposes an alternative approach designed to predict a quantitative value for the listener preference, which can be used to better analyze, compare and design venues for Rock and Pop music. The suggested algorithm is designed to be a tool that any user (designer, artist, tour manager, etc.) can approach without specific acoustic knowledge. It is what we could call a simple input-objective, output-subjective system.

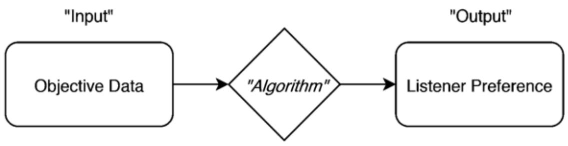

Figure 1: Preference Rating algorithm structure.

In order to create a formula, a meta-analysis of existing literature has been undertaken. This highlighted recurrent links between listener preference and objective measurements, providing an alternate viewpoint from which to assess the relative importance of the metrics involved. Eventually, eight acoustic factors have been selected, alongside a weighting to determine their relative importance.

The method suggested in this paper was also able to correctly blend different points of view in terms of listeners' preference, including aspects preferred either by musicians or by sound engineers, as it will be presented in the results and comments sections.

A previous study by Adelman-Larsen et al. measured various acoustical parameters in 20 music venues in Denmark[7] and linked them with the preference of a panel formed by both musicians and sound engineers. The same data set describing these 20 venues was used in this present paper at hand for validation.

Section 2 will describe the background of this research, while section 3 will introduce the proposed algorithm. The data used to validate the method will be described in section 4, and in section 5 and 6 the analysis of the results, further developments and comments will be discussed.

2 Background

In order to derive an algorithm able to predict a listener's preference, firstly the context must be defined. For this study, the algorithm attempts to predict the performance of Rock and Pop music halls and venues specifically. A meta-analysis[8] was conducted to determine the relationships that are present between performance spaces and listener preference.

While individual studies of any nature are expected to have some degree of error, a meta-analysis can overcome this boundary by eliminating outlying studies that disagree with the overall pattern presented by a wide array of investigations[8]. In this context, meta-analysis was used to discover any conclusion that draws a link between objective data and listener preference. From this, the metrics to be included within the algorithm have been determined, as well as to what degree they influence preference.

In our meta-analysis we firstly collected all the possible acoustical parameters discussed in literature, and then we tried to associate them with a certain acoustic context. For example, all the metrics related to Reverberation time were collected together in a class, or, in a similar way, several other parameters were grouped in a category related to intelligibility. We then formalized this use of broad sub-categories, by using the capitalized term Acoustic Context, which could then be used to contain different nuances (in terms of metrics) of the same physical property of a room. Eventually the last part of the meta-analysis consisted in selecting the smallest set of acoustical parameters from all the Acoustic Contexts, in order to ensure the Preference Rating algorithm would not over-estimate or under-estimate any acoustical aspect or physical attribute of the scenario at hand.

Acoustic Contexts and the acoustical parameters are grouped in Table 1, in which each row presents one acoustical parameter and how frequently it was found within the analyzed database, which is based on the J-AES, J-IOA and J-ASA catalogue, for a total of 136 published articles. Further details of the meta-analysis process and the full bibliography are contained in: "*Could a metric, based on objective acoustic criteria, accurately predict a listener's preference of a venue?"*, by P. Critchell[9] and can be retrieved via correspondence with the authors.

Acoustic Context:	**Variable:**	**No.:**
Shape / Layout:	***Room Modes***	25
	Early Reflections	16
	ITDG	2
RT:	***RT*** *(AVG)*	28
	Bass Ratio	17
	Treble Ratio	3
	'Even' Decay	9
EDT:	***EDT***	5
Intelligibility:	***Definition***	9
	Clarity	8
	STI	4
Materials:	***Diffusion (SDI)***	7
External Fact.:	***Background Noise***	3
	Total Papers:	**136**

Table 1: Relative frequency of acoustical parameters within the meta-analysis database. In **bold** the chosen parameters included in the final algorithm.

Some metrics, such as Time-Alignment (or phase) and Loudness have not been included because they are influenced by the loudspeaker system used (and its tuning) within the space. As these factors are not a fixed aspect of the room they have been omitted.

Generally, to select an acoustical parameter to represent its Acoustic Context, a condition of 5 minimum mentions in the body of literature must have been reached. This was decided to eliminate outlying or uncommon results, streamlining our model by including only the most frequently associated variables. However, if only one variable existed within an Acoustic Context, it was included disregarding how frequently it was mentioned. This ensures that each physical aspect of the room could be correctly represented within the algorithm. For example, background noise was included even if retrieved only three times in our meta-analysis.

Where more than one variable was identified within an Acoustic Context, the variable(s) that was most suitable was chosen. For example, when assessing the "Shape/Layout" Acoustic Context, Bonello Distribution[10,11] was preferred to Early Reflections and the latter was dropped although mentioned a high number of times. The choice was made because the modal response can be easily analyzed from the room dimensions. This simplifies the algorithm for prediction, requiring fewer prior calculations. Furthermore, Early Reflections was omitted as in many music venues the 'optimum' listening location is a single, fixed point. On the contrary, Rock and Pop venues are usually full of standing and moving guests, so analyzing this quality from the perspective of a single location might not be relevant to the listeners' preference.

Based on Table 1, the algorithm will consist of eight variables (highlighted in bold): Reverberation Time (T30), Bass Ratio (BR), Definition (D50), Early Decay Time (EDT), Bonello Distribution, Linearity of decay from 63Hz-4kHz, Background Noise (dB) and the Surface Diffusivity Index (SDI).

3 Defining Preference Rating (PR)

Two key processes are applied to the raw acoustic data in order to derive the Preference Rating (PR). The first step defines the 'ideal' values for Rock and Pop venues to 'rate' each factor against. The values are interpolated linearly, producing a 'Score' (S_x) value for each factor. The metrics and their respective targets are directly linked to literature highlighted by the meta-analysis in the context of 'Rock' and 'Pop' venues.

In addition to the parameters presented in Table 1, an estimated value for SDI and Background Noise were used, as real-life data was not available. These values will be represented as S_{SDI} and S_{BG} respectively in further testing.

Acoustic Metric:	**Target:**	**"S_x":**
RT30:	0.8s	S_{RT}
Shape / Bonello Distribution:	*"Curves?"*	$S_{MD\text{-}C}$
	"Increasing"	S_{MD}
Bass Ratio (BR):	1.2dB	S_{BR}
Flat RT / Decay.	0.0s Deviation	S_{TN}
Definition (D50):	1	S_{DN}
EDT:	20ms	S_{EDT}

Table 2: Scoring Acoustic Factors for Preference Rating.

It's worth mentioning that in case we dealt with other music genres or other scenarios, such as schools or spaces dedicated to speech more than to music creation/fruition, we should obviously ask ourselves the questions whether different targets or other acoustical parameters should be considered.

For the numerical aspects, the raw acoustic data was linearly interpolated, except for the Bonello Distribution[10,11], which was graded in bands from 'Strictly Increasing' to 'Strictly Decreasing' providing 5 'steps'. For further development, the authors recognise that the inclusion of the room mode equation into the algorithm will provide better functionality to the users, as they will only need to ascertain the dimensions of the room. In addition to this, as a subset of S_{MD}, a penalty factor $S_{MD\text{-}C}$ is to be applied to any room with a curved wall / ceiling present, reducing the value of S_{MD} by 50%.

An example of the calculation of one "score" is provided for S_{RT} (1), (2). In our case RT is the one-third octave-band averaged reverberation time from 63Hz to 4kHz.

Scoring Reverberation Time (S_{RT}):

a) if RT > 0.8s:

$$S_{RT} = \left(0 + (RT - 0.8) \times \left(\frac{1}{1.25 - 0.8}\right)\right) \quad (1)$$

b) if RT < 0.8s:

$$S_{RT} = \left(1 + (RT - 0.8) \times \left(\frac{1}{0.4 - 0.8}\right)\right) \quad (2)$$

If the integrated value S_{RT} is negative, then $S_{RT} = 0$.

Once all the "scores" are calculated, the second step of the algorithm consists in a weighted average of the interpolated data, again using meta-analysis as a source to assign to each parameter its corresponding relative weight. The outputs of this process were called Relative Importance Weightings (RIW's). The weightings are derived from the number of mentions each acoustic metric received in the meta-analysis.

Also RIW's are intrinsically linked to context – in our case Rock/Pop venues. If, for example, further meta-analyses were conducted from the perspective of classical music and concert hall acoustics, the weightings and target values may shift accordingly.

Similarly, if the performance of a hall for speech was the focus of the algorithm, the context would then suggest greater weighting on Intelligibility related factors, and the omission of metrics such as BR for the inclusion of Strength Factor (G)[12].

Table 2 details each RIW, distilled through meta-analysis in the context of Rock and Pop music.

Acoustic Metric:	**RIW (%):**
RT_{AVG}:	26.67%
Bonello Distribution:	23.81%
Bass Ratio (BR):	16.19%
Flat RT / Decay.	8.57%
Definition (D50):	8.57%
EDT:	6.67%
(Add.) Diffusion:	6.67%
(Add.) B.G. Noise:	2.86%

Table 3: Relative Importance Weighting's (RIW) for Weighted Averaging within Preference Rating.

Based on these elements, the formula is arranged according to Figure 2, where *PR* is equal to the weighted average of all Scored (*S*) variables. Equation (3) details this mathematically.

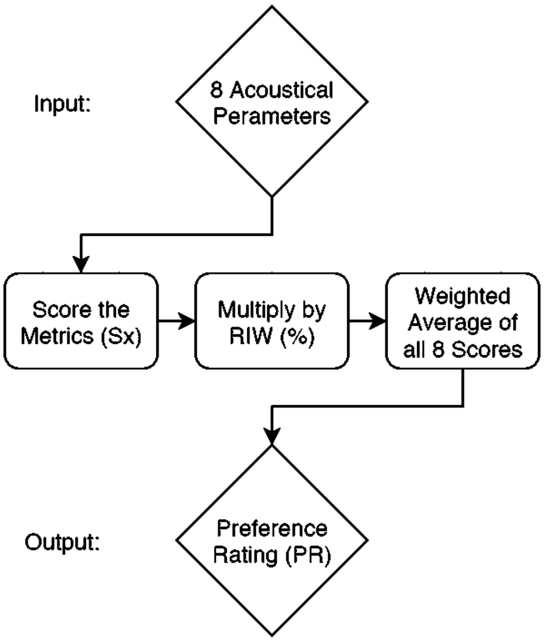

Figure 2: Preference Rating (PR) Algorithm flow and processes.

Preference Rating (PR):

$$PR = \left(\begin{array}{c} S_{RT} \times 26.67 \\ (S_{MD} \times S_{MD-C}) \times 23.81 \\ S_{BR} \times 16.19 \\ S_{TN} \times 8.57 \\ S_{DN} \times 8.57 \\ S_{EDT} \times 6.67 \\ S_{DIF} \times 6.67 \\ S_{BG} \times 2.86 \end{array} \right) / 100\% \qquad (3)$$

With the presented calculation of scores and weights, Preference Rating becomes a function returning a real value included in the interval [0, 1], and its output can be used to compare venues to each other. With enough data a 'benchmark' could be defined (i.e. *PR=0.8*) against which a designer would be able to test a theoretical room to ensure it will meet a minimum standard of preference.

4 Data and Validation

To ascertain the effectiveness of the new algorithm, it has been fed with the all the data of the acoustical parameters from Adelman-Larsen past research[7]. The data set contains subjective ratings of all rooms from sound engineers and musicians. Therefore, the effectiveness of *PR* can be evaluated based on how closely it matches the subjective data. As previously stated, values of SDI and Background Noise were not included in the data set of the 20 measured venues, and accordingly estimations for these parameters were made based on images and floor plans.

5 Results and Comments

Initial testing provided a strong correlation between the algorithm output and the subjective data. Presented in Figure 2 is the Preference Rating in full, both with and without the Additional Factors (namely the scores S_{SDI} and S_{BG} calculated from Surface Diffusivity Index and Background Noise). Result are plotted against the subjective data expressed by the listeners' panel interviewed by Adelman-Larsen[7] in his research.

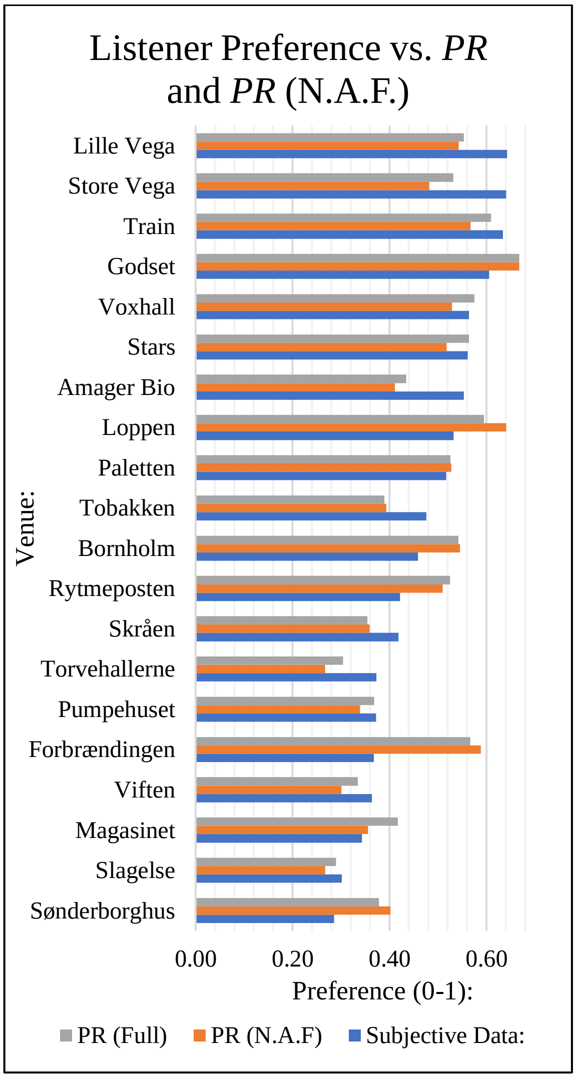

Figure 3: Preference Rating (PR) relative to Listener Preference[7]. No Additional Factors (N.A.F) testing indicates that the estimated values S_{SDI} and S_{BG} have been omitted.

Considering that the numerical results of PR for a few venues were not as accurate as desired, a series of modifications to the algorithm were attempted. In particular, altering the calculation of specific acoustical parameters, or replacing them with alternate acoustical parameters of the same Acoustic Context[9].

Rather than analyzing a plethora of plots to discern the differences between subjective data and the various forms of preference rating, the cumulative sum of the difference between the two data sets was calculated as a 'total deviation' from the subjective data, and is plotted in Figure 3. Achieving a value of zero indicates that for every venue the prediction was identical to that of the subjective testing. The larger the value of the total deviation, the farther the estimation predicted by PR lies from the subjective data. Therefore, the most consistent version of the algorithm lies closest to zero and is highlighted in *orange*.

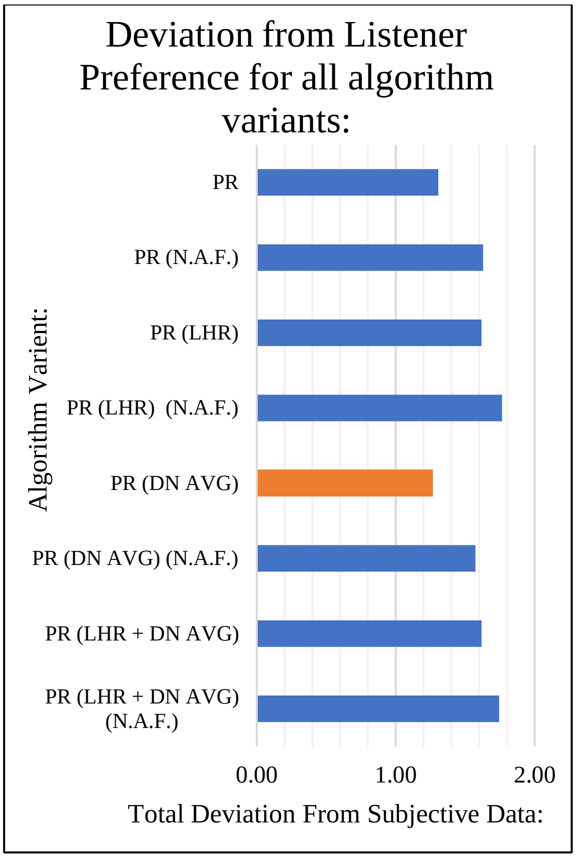

Figure 4: PR Deviation from Subjective Data. The most successful variant (indicated in *orange*) is achieved by using an average D50 in place of the ISO D50[2].

The variant of the algorithm, which reached the best performance, is called PR (DN AVG). This version uses an alternative version of the Definition parameter, D50, that is calculated by averaging it from 63Hz to 4kHz instead of the ISO standard calculation of 500Hz and 1kHz octave bands only[3] to improve the results of the algorithm.

Using Low-High Ratio (*LHR*)[13] in place of *BR*[2] and any combinations of these two proved to worsen the correlation. In Figure 3 there is also depicted the difference between results when including and excluding the additional factors of SDI and Background Noise, however as stated prior, these are estimates within this study, and further investigation should be undertaken.

Plotted in full against subjective data provided by Adelman-Larsen[7], Figure 4 details a positive correlation between objective data and preference - despite the relative lack of research into the ideal acoustics of Rock/Pop music venues when compared to classical music spaces. The authors believe this research will help to narrow the gap and provide further insight into room acoustics for live music.

The correlation displayed within Figure 4 is significant, depicting 17 of the 20 venues within a 5% margin. Additionally, compared to the combined subjective data of both musicians and engineers, the Preference Rating tends towards the musician's subjective preference[9].

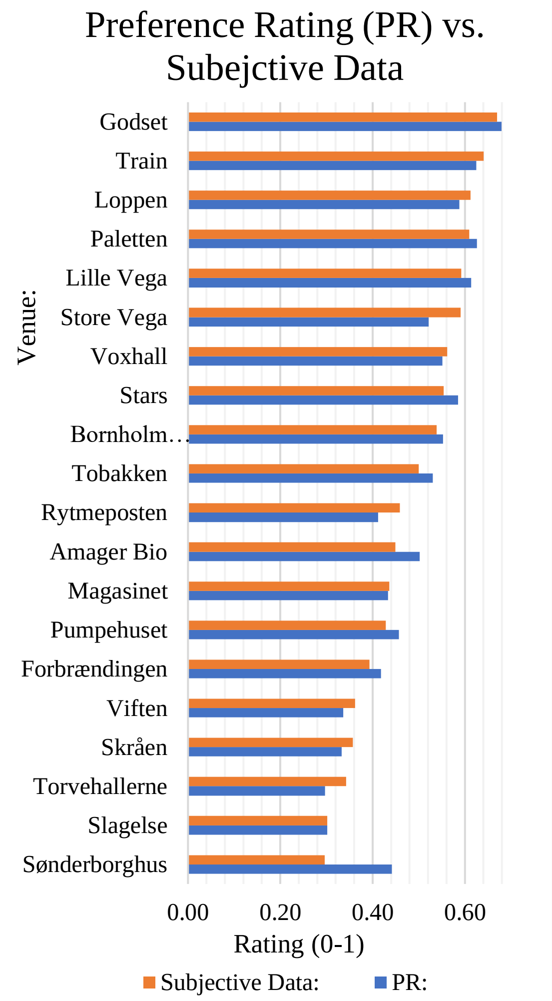

Figure 5: Preference Rating (PR) compared to Listener Preference[7].

Finally, it should be noted once again that this is not the *only* method derived for predicting acoustic quality, but this alternative approach does provide clear links between acoustic metrics and listener preference for Rock and Pop music venues. The algorithm acts as a simple input-output system in which any user, acoustician or not, would be able to gather the data of a specific hall, or (using computer aided simulation) predict the values of the acoustical parameters and the corresponding PRs, and receive an instant gauge on how well the venue is likely to perform within this context.

6 Future Development and Acknowledgment

At this stage the suggested empirical model is not yet perfect, and all different variants of the algorithm that were tested presented some outlying results (<15%). Interestingly, often the same venues were offering contradictory results in each test, suggesting that there may be other additional factors at work that are not included within this algorithm. More specifically, visual clues and non-acoustical aspects as highlighted on by Adelman-Larsen when discussing the recording of the subjective data used[3].

Considering the work by Jain Kang[14] and Barry Truax's seminal work on soundscaping[15], it is to be considered that a person's preference of a venue acoustically is most likely affected not only by sound, but also by visual stimuli to a significant degree. This factor of *'How bad does the room look?'* should be investigated as it may provide an additional weighting to the algorithm that may account for outlying venues. This has not been considered previously in attempts to create a combined metric, and it is therefore suggested that listener preference *may* not align with the most acoustically perfect room every time.

The suggested formula provides a platform for further investigation into similar areas. For example, alteration of the boundaries of the Scoring part of the algorithm could provide a system with multiple use-cases. If RT were altered for an optimum of 1.8s, then the formula could predict the preference of concert halls. Similarly, if RT were lowered, and target values of C50[3] and STI[16] were optimised for intelligibility, then the algorithm could be used to predict the preference rating of a conference room or lecture theatre for speech.

In addition, the algorithm could be implemented to produce two outputs. One could be the existing listener preference prediction, while a second one could be a variant tailored on the on-stage acoustics, thus assessing the musician's preference.

This topic has been recently investigated, among the many, by D'Antonio whose research clearly highlights that when a musician struggles while playing due to poor acoustic on stage acoustics, then his performance has a smaller chance of engaging the audience, which accordingly is less likely to enjoy it[17]. This indicates that the quality of on-stage acoustics will affect the Preference Rating accordingly for every venue.

The present research would not have been possible without the large data set of acoustical measurements and subjective ratings of venues against which the algorithm and RIW's have been validated. The authors want to acknowledge this data collection to Dr Adelman-Larsen, as referenced in *"Rock and Pop Venues: Acoustic and Architectural Design*[3]*"*, and *"Suitable reverberation times for halls for rock and pop music*[7]", and want to deeply thank him for his support and open collaboration.

References

[1] Fricke, F., 2000. Concert Hall Acoustic Desing: An Alternative Approach. *Building Acoustics,* 7(4), pp. 223-246.

[2] The International Organization for Standardization, 2012. *ISO 3382:2012 Acoustics -- Measurement of room acoustic parameters.* Geneva, ISO.

[3] Adelman-Larsen, N. W., 2014. *Rock and Pop Venues: Acoustic and Architectural Design.* 1st ed. Heidelberg: Springer.

[4] Sabine, W. C., 1922. *Collected Papers on Acousticcs.* 1st ed. New York: Harvard University Press.

[5] Beranek, L. L., 2008. Concert Hall Acoustics (Collective Papers). *Journal of the Audio Engineering Society,* 56(7/8), pp. 523-544.

[6] Beranek, L., 2004. *Concert Halls and Opera Houses: Music, Acoustics and Architecture.* 2nd ed. New York: Springer.

[7] Adelman-Larsen, N. W., Gade, A. C. & Thompson, E. R., 2010. Suitable reverberation times for halls for rock and pop music. *Journal of the Acoustical Society of America,* Volume 127, p. 247.

[8] Lipsey, M. W. & Wilson, D. B., 2001. *Applied social research methods series, Vol. 49.* 1st ed. London: Sage Publications Inc.

[9] Critchell, P., 2019. *Could a metric, based on objective acoustic criteria, accurately predict a listener's preference of a venue?* Southampton: Solent University.

[10] Bonello, O. J., 1979. Acoustical Evaluation and Control of Normal Room Modes. *Journal of the Acoustical Society of America,* Volume 66.

[11] Bonello, O. J., 1981. A New Criterion for the Distrubution of Normal Room Modes. *Journal of the Audio Engineering Society,* Volume 29.

[12] Beranek, L., 2003. *Acousics of Concert Halls.* 2nd ed. New York City: Springer.

[13] Weisser, A. & Rindel, J. H., 2006. Evaluation of Sound Quality, Boominess, and Boxiness in Small Rooms. *JAES*, 54(6), pp. 495-511.

[14] Kang, J., 2007. *Urban Sound Environment* (1st ed.). London: Taylor & Francis.

[15] Truax, B., 1984. Acoustic Communication. New York City: Ablex Publishing.

[16] The International Electrotechnical Commission., 2005. *IEC 60268-16 Objective rating of speech intelligibility by speech transmission index* (3rd ed.). London: IEC.

[17] D'Antonio, P., 1991. *Performance Acoustics: The Importance of Diffusing Surfaces and the Variable Acoustics Modular Performance Shell (3118).* New York, Reprinted by The Audio Engineering Society.