Technische Universität Berlin

Stefan Weinzierl, Steffen Lepa, Martin Thiering

The Language of Rooms: From Perception to Cognition to Aesthetic Judgment

Open Access via institutional repository of Technische Universität Berlin

Document type
Book chapter | Accepted version
(i. e. final author-created version that incorporates referee comments and is the version accepted for publication; also known as: Author's Accepted Manuscript (AAM), Final Draft, Postprint)

This version is available at
https://doi.org/10.14279/depositonce-15239

Citation details
Weinzierl, S.; Lepa, S.; Thiering, M. (2020) The Language of Rooms: From Perception to Cognition to Aesthetic Judgment. In: Blauert, J., Braasch J. (eds) The Technology of Binaural Understanding. pp. 435–454 (Modern Acoustics and Signal Processing). Springer, Cham. https://doi.org/10.1007/978-3-030-00386-9_15.

Terms of use
This work is protected by copyright and/or related rights. You are free to use this work in any way permitted by the copyright and related rights legislation that applies to your usage. For other uses, you must obtain permission from the rights-holder(s).

Stefan Weinzierl, Steffen Lepa, Martin Thiering

Fachgebiet Audiokommunikation, Technische Universität Berlin, Berlin, Germany

Corresponding author: S. Weinzierl
E-mail: stefan.weinzierl@tu-berlin.de

Abstract Rooms are not perceptual objects themselves; they can only be perceived through their effect on the presented signal, the sound source, and the human receiver. An overview of different approaches to identify the qualities and the dimensions of “room acoustical impression” will be provided, that have resulted in psychological measuring instruments for room acoustical evaluation from the audience perspective. It will be outlined how the psychoacoustic aspects of room acoustical perception are embedded in a socio-cultural practice that leads to an aesthetic judgment on the quality of performance venues for music and speech.

1 Language and Perception

The aim of this contribution is to highlight the relationship between the characteristics of performance venues for music and speech and the language which is used to describe them. On the one hand, an overview of different approaches to using language as a “measuring instrument” for the qualities of these spaces will be provided. On the other hand, there is an interest in what conclusions can be drawn from the language used with respect to the characteristics of these spaces, the listeners using this language, and the perceptual and cognitive processes involved.

These relationships will be looked at through the lens of a theoretical frame describing the relationship between cultural artifacts (performance spaces), their perception, and their linguistic encoding. This frame model, combining elements of perceptual psychology and cognitive linguistics, assumes a perceptual front end, where an external acoustical signal is transformed into neural activity by auditory sensory organs. For hearing, this process takes place in the inner ear, where sound pressure transmitted by the outer and middle ear is transduced into neural signals. In a perceptual back end, these signals are integrated above different sensory modalities and activate *concepts*, i.e. mental representations corresponding to abstract classes of objects, which "tie our past experiences to our present interactions with the world" (Murphy 2004, p. 1) and allow us, for example, to classify an audiovisual experience in a certain social setting as a "concert" and a spatial environment of a specific size, design, and acoustical properties as a "concert hall". The result of this comparison of sensory information with preconfigured categories (concepts) is called *percept*.

The size and the structure of the concept repertoire as well as the matching process with the sensory input depends on many personal and situational factors, including the knowledge, the experience, the expectations, the motivation, and the attention of the listener. Accordingly, the percept, such as "a successful concert in an acoustically appropriate environment", depends as much on these situational factors and the conceptual repertoire of the individual listener as it depends on the sensory input at this specific moment.

The relevance of language in this process has two important aspects. First, language is—not the only, but the most important—"metrological" access to human perception. From paired comparisons, similarity judgments, sorting tasks, multidimensional scaling, semantic scales, to vocabulary profiling and related qualitative and quantitative analyses: most studies of the auditory properties of performance venues and sound description in general (Susini et al. 2011) have relied on language-related tasks borrowed from the repertoire of methods of experimental psychology and quantitative and qualitative social research. Second, the language used can itself provide information about the speaker's conceptual representation of the world (Evans and Green 2007, p. 5), such as about the spatial environments where listeners' perceive music or speech. The taxonomic organization and the privileged level of categorization that is used in everyday language about spatial concepts are the results of

Fig. 1 Musical events as a cultural, social, visual and acoustic experience. The language to describe performance venues for music and speech reflects each of these domains. The image shows the concept for a new concert hall, to be opened in Munich (©Cukrowicz Nachbaur Architekten)

preceding experiences, and it also shapes the lens through which new experiences are observed. The preferred vocabulary about performance venues does not only reflect the knowledge and the professional experience, e.g., of expert versus lay listeners; it is also assumed to have a direct impact on their instantaneous perception and the respective mental models—for an introduction on mental models see Johnson-Laird (1983), for a description on frame-theory see Minsky (1977). This kind of "linguistic relativism", which has been evidenced for the languages of different ethnic groups by many empirical observations in cognitive linguistics (Dabrowska and Divjak 2020; Dancygier 2017; Everett 2013; Levinson 2020; Thiering 2018), also occurs on a small scale in lay versus expert language linguistics, or when the languages of groups with different kinds of expertise are compared, such as music listeners versus musicians.

2 Linguistic Inventories as a Basis for Psychological Measuring Instruments in Room Acoustics

Throughout the first half of the 20th century, the investigation of room acoustical environments was mainly focused on the effects of reverberation, with its dependence on frequency, and with its control through volume and absorption according to Sabine's formula (Sabine 1900). Psychological experiments were conducted already by Sabine himself. In 1902, he invited a number of musical experts to the then recently built New England Conservatory of Music to judge the acoustic quality of piano instruction rooms while seat cushions were successively added to the rooms in order to reduce their reverberation times. Sabine observed that the listeners judged all rooms to be acoustically optimal if the reverberation time was within a quite narrow range of tolerance, from which he concluded a common taste of "*surprising accuracy*" for the acoustical conditions of musical performance venues (Sabine 1906). This unexpected consensus on the appropriate acoustical conditions for classical concert venues, which would hardly have been observed 100 years earlier (Weinzierl 2002), can only be interpreted as the result of a cultural process which accompanied the emergence of public concert life, and which was largely completed around 1900 (Tkaczyk and Weinzierl 2019).

While Sabine asked his subjects only whether the duration of reverberation was appropriate, the British architect and acoustician Hope Bagenal carried out similar experiments with musicians as test participants, who were asked to assess the effect of room acoustic conditions on different sound qualities including "reverberation" (too long/too short), "tone" (full/bright/rich/soft), "tone" (hard/thin/dead/dull), "loudness" (sense of power, body of tone), "reinforcement of notes" (even/uneven) and "conditions", by which he asked his subjects to name specific halls which resemble the conditions reached (Bagenal 1925). Even if this scheme was not consistently adhered to by his subjects, it can be considered as the first attempt to have the acoustics of concert halls evaluated by a semantic differential, covering the room acoustical effect on spatial, dynamic and timbral dimensions, and considering the effect of typicality with respect to prototypical reference halls. The first standard with guidelines for auditorium acoustics, issued in 1926 by the American Bureau of Standards, however, was mainly focused on specifying an optimal range of reverberation times for halls of different sizes, along with recommendations of how to reach these values (Bureau of Standards 1926).

After 1950, an increasing awareness can be observed that an optimal reverberation time alone is no guarantee for a successful room-acoustical design, and that "reverberance" should not form the only criterion for the perceptual assessment of halls. The British Broadcasting Corporation (BBC) and the Acoustics Group of the Physical Society sponsored a number of experiments, where an identical repertoire (*Don Juan* by R. Strauss) was performed and recorded in four different British concert halls. Subjects listening to the different recordings were then asked to produce a ranking of these halls concerning tonal quality, definition, and overall preference. Consistent rankings, however, could not be observed, and practising musicians exhibited preferences that were different from other skilled listeners (Somerville 1953).

In a further study, a glossary of 14 acoustic terms was collected, which were, according to the authors, commonly used to describe the qualities of concert halls and recording studios—see entry for Somerville and Gilford (1957) in Table 1. A similar list of 18 attributes was proposed by Beranek (1914–2016) in his landmark book on *Music, Acoustics and Architecture*, along with relations between these perceptual qualities and physical properties of the hall, which were based on his intuition and experience (Beranek 1962).

The two lists of attributes, along with a third, originally German list described below, demonstrate the grown awareness for the multidimensional impact of room acoustical conditions on the perceived sound qualities in these halls. At the same time, the studies reflect an awareness for the need to separate the physical and the perceptual domain more clearly. Somerville and Gilford emphasized that the "*subject under investigation is purely aesthetic and therefore must begin and end with human aesthetic judgments*" (Somerville and Gilford 1957, p. 171). Nevertheless, their list is a mix of perceptual and physical items, including aspects such as "scattering" or "standing wave system" without an obvious equivalent in the perceptual domain. Beranek's features, on the other hand, are psychological throughout, at least if aspects such as "ensemble" and "dynamic range" are understood as "perceived ensemble" and "perceived dynamic range".

The way to analyze the results of questionnaire studies constructed on the basis of these terms was paved by the psychological fundamentals of the use of semantic differentials (Osgood et al. 1957) and the statistical techniques of multidimensional scaling (MDS, Torgerson 1952) and factor analysis (Spearman and Jones 1950), all of which were introduced during the 1950s. One of the first applications of these new tools was made in a study by Hawkes and Douglas (1971). Sixteen attributes inspired by Beranek's list, shown in Table 1, were used for a questionnaire applied in four different British Concert Halls (with different musical programs and performers), in the Royal Festival Hall, London, with the newly installed Assisted Resonance system in different technical settings, and in the Royal Festival Hall at 23 different positions. With interest in identifying the different dimensions of acoustic experience (Hawkes

Table 1 Attributes addressed in early investigations on the differential qualities of room-acoustical environments. The translation of the originally German attributes by Wilkens (1977) was adopted from Kahle (1995)

	Somerville and Gilford (1957)	Beranek (1962)	Wilkens (1977)
1	Balance	Intimacy	Small/large
2	Bass masking	Liveness	Pleasant/unpleasant
3	Coloration	Warmth	Unclear/clear
4	Deadness	Loudness of the direct sound	Soft/hard
5	Definition	Loudness of the reverberant sound	Brilliant/dull
6	Diffusion	Definition/Clarity	Rounded/pointed
7	Echoes	Brilliance	Vigorous/muted
8	Flutter echoes	Diffusion	Appealing/unappealing
9	Liveness	Balance	Blunt/sharp
10	Pitch changes	Blend	Diffuse/concentrated
11	Scattering	Ensemble	Overbearing/reticent
12	Singing tone	Immediacy of response	Light/dark
13	Slap back	Texture	Muddy/clear
14	Standing-wave system	Freedom from echo	Dry/reverberant
15		Freedom from noise	Weak/strong
16		Dynamic range	Treble emphasized/not emphasized
17		Tonal quality	Bass emphasized/not emphasized
18		Uniformity throughout the hall	Beautiful/ugly
19			Soft/loud

and Douglas 1971, p. 249), the authors applied both MDS and factor analyses, finding 4–6 orthogonal factors. The solutions they obtained, however, were different both for the different stimulus settings and for the different types of analyses, i.e., both the number of factors and the relation of factors and items were different for each of the sub-studies. This problem will be further addressed in Sect. 3.

While Hawkes and Douglas collected data in the field, i.e., by interviewing concertgoers, Lehmann and Wilkens (1980) used an experimental approach by presenting dummy-head recordings of the Berlin Philharmonic Orchestra in six different halls and capturing the assessment of subjects on a semantic differential with 19 different attributes shown in Table 1. These were selected from a list of originally 27 items by eliminating those with an excessive inter-rater variance, indicating an inconsistent interpretation between subjects. The factor analysis of the ratings delivered three orthogonal factors, explaining 89% of the total variance. Considering the weigths of the original attributes on these variables, these were interpreted as *strength* *and extension of the sound source*, *definition* and *timbre of the overall sound* (Wilkens 1977). In an attempt to overcome the limitations of the experimental approach lacking the ecological validity of live concert situations, Sotiropoulou et al. (1995) used a questionnaire to be rated at three concerts in two different concert halls in London. Similar to Lehmann and Wilkens (1980), they started with a larger vocabulary of about 100 labels, which was reduced to 54 bipolar attributes based on a relevance rating collected in pretests. Analysing the ratings of about 80 participants by factor analysis, the authors obtained four factors explaining roughly 66% of the total variance, which they interpreted as *body*, *clarity*, *tonal quality*, and *proximity*. In both experiments it became obvious that linguistic descriptors are not necessarily suitable as measuring instruments if their meaning is inconsistently interpreted by different raters or if their immediate relationship to the perceptual object under consideration is not assured.

The numerous investigations dedicated to finding suitable *technical* parameters to predict specific perceptual categories are not the subject of this contribution. Only some of them are also interesting here because they highlighted the importance of specific perceptual aspects which did not appear in earlier studies (compare Table 1). Most importantly, a group of studies emphasized the importance of spatial aspects of room acoustics, in particular of an increased perceived "source width" and the perceived acoustic "envelopment" of the auditorium (Barron 1971; Barron and Marshall 1981; Bradley and Soulodre 1995). All of these studies, however, employed synthetic sound fields created by loudspeakers in the anechoic chamber, and asked participants to evaluate these qualities as isolated items. Since they were not evaluated as part of a multidimensional measuring instrument, it is not apparent to what extent they form *independent* aspects of the room acoustical impression, or whether they are physically or perceptually correlated to other aspects.

In this context, it is essential to bear in mind that the ratings of two objects can be correlated because two labels refer to similar perceptual impressions (such as "loudness" and "strength" of sound), or because different perceptual qualities covary in the physical objects of the stimulus pool. For example, rooms providing more "reverberance" could—for physical reasons—always provide more "envelopment", although the perceptual concepts are clearly different.

After 2000, the study of the perceptual space of room acoustic conditions as a whole attracted a renewed interest directed to the *individual* vocabularies used to describe room acoustic conditions. Several studies, first aiming at the evaluation of spatial-audio reproduction systems (Berg and Rumsey 2006) and then also on the perception of natural acoustical environments, included a qualitative part for the verbal elicitation of the terminology and a quantitative part for the statistical analysis of the generated terms, allowing to identify clusters of attributes with a similar meaning. An initial of two studies conducted at Aalto University, Helsinki, produced room-acoustical stimuli by impulse-response measurements of a loudspeaker orchestra in three different concert halls, encoded in Ambisonics B-Format, processed with directional audio coding (DirAC, Pulkki 2007) and reproduced by a 16-channel loudspeaker system. A second study used impulse responses of eight different concert halls, encoded in Ambisonics B-Format, processed with the spatial impulse-

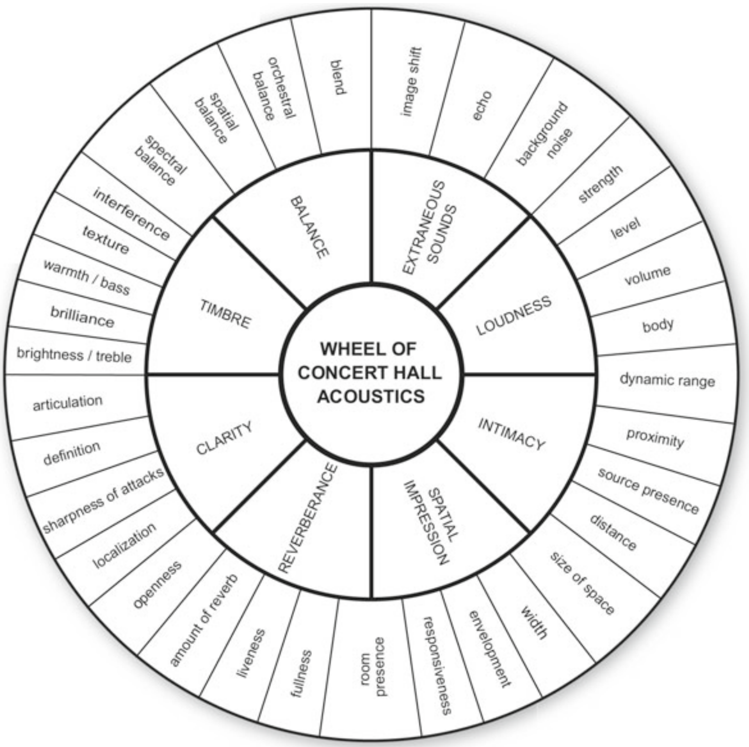

Fig. 2 Wheel of concert hall acoustics (Kuusinen and Lokki 2017)

response-rendering algorithm (SIRR; Merimaa and Pulkki 2005) and reproduced by a 14-channel loudspeaker system. Analysing the large set of about 100 individual attributes generated and rated by 20 resp. 23 participants in the two studies, two resp. three main components could be extracted explaining 66% resp. 67% of the total variance. These were interpreted as *loudness/distance* and *reverberance* in the first study (Lokki et al. 2011), and *loudness, envelopment and reverberance*, *bassiness and proximity* and *definition and clarity* in the second study (Lokki et al. 2012).

As a summary of their own and other work, authors of the same group suggested a "wheel of concert hall acoustics", including eight main categories and 33 items to visualize the main perceptual aspects of concert halls with unamplified musical instrument sounds (Kuusinen and Lokki 2017). The wheel format, which has a longer tradition in the domain of food quality and sensory evaluation (compare Noble et al. 1987) is a structured and hierarchical form to present a lexicon of different sensory characteristics. Pedersen and Zacharov (2015) used the wheel to present such a lexicon for reproduced sound, with the selection of the items and the structure of the wheel based on hierarchical cluster analysis and measures for discrimination, reliability, and inter-rater agreement of the individual items—an empirical basis which was not provided in Kuusinen and Lokki's original wheel for concert halls—see Fig. 2.

3 Psychometrics and Scale Development in Room Acoustics

The studies summarized above have, in different ways, confirmed the multidimensional character of room acoustical conditions as a mediator for the sound qualities of music and speech, and they provided different lists of attributes to describe these qualities. As underlying dimensions of the room acoustical impression, the *loudness* or *strength* and the *reverberance* of rooms were consistently extracted from the ratings of these attributes, as well as a factor for the *timbre* of the room. None of these studies, however, attempted to construct a standardized measuring instrument that can be used with different groups of listeners to describe the whole width of the room-acoustic perception space. From today's point of view, this was impossible due to shortcomings in terms of the experimental stimuli, the participating subjects and the statistical analysis techniques employed in these works. According to modern standards in social research, any psychological measurement instrument has to meet established quality criteria in terms of *reliability*, *validity*, and *invariance* concerning a typical sample of users and stimuli. To establish perceptual dimensions, the investigation has to take care of *representativeness* and *breadth* of stimuli, possible hidden *confounders* and a proper *sample selection* in order to prevent bias concerning the generalizability of results.

A first problem that pertains to most of the early studies in acoustical room impression measurement is the lack of experimental control concerning the stimuli presented. Studies that collected ratings in physically existing rooms always risked the influence of hidden confounders such as the audio content, the visual impression, or the musical performance, all of which co-vary with the auditory impression of the room. Presenting the whole breadth of possible room acoustical conditions while keeping these confounding variables constant seems only possible with state-of-the-art technologies for auralization. It is, of course, true that such an experimental approach can not account for all visual, architectural, and social aspects that constitute the multi-modal impression of a concert venue. To determine the *acoustical* properties of a room, however, these influences act as confounders increasing the measurement error of the test.

One may ask to what extent full control of all non-acoustic factors makes sense, since, for example, an interaction between the room-acoustic conditions and the playing style of musical performers is also present in the real situation and thus not an experimental artifact in the narrower sense. However, incorporating such interaction into the experimental design might, on the one hand, conceal characteristics of space if, for example, musicians were tempted to compensate for the effect of space by adjusting their timbre and volume in the opposite direction. On the other hand, the effect of space on the performer's playing has proven to be quite individual. Different musicians react in very different ways to room-acoustic conditions

(Schärer Kalkandjiev and Weinzierl 2013, 2015; Luizard et al. 2020). The consideration of the interaction of space and performance would, therefore, face the problem of which reaction pattern should be used as a basis here. Nevertheless, the final question that must be taken into account regarding the overall aesthetic judgment of space—"*Is the room suitable for a musical content?*"—cannot be conclusively evaluated in the laboratory.

A second challenge lies in the sample of rooms presented in terms of representativeness. The identification of latent dimensions of room acoustical perception, i.e., a stable factor-analytic solution of the measured data, which is valid beyond the specific sample of rooms used in the test, cannot be expected for a too-small set of stimuli. In order to identify five largely independent perceptual dimensions, a set of at least $2^5 = 32$ stimuli would be required so that all perceptual qualities can be varied systematically and independently from each other and hence can be adequately identified by factor analysis. Furthermore, only with a sufficiently large sample of rooms, the results can be considered representative of the targeted population of room acoustical conditions. Comparing these requirements with the sample sizes used in the studies mentioned above, with typically less than ten rooms, it becomes evident that neither the dimension of the perceptual space, i.e., the number of latent variables, nor the structure and interpretation of the adopted factor solution could be reliably determined.

A third challenge is the size of the sample of listeners. In order to reliably assess the dimensionality of perceptual constructs represented by questionnaire item batteries by factor analyses, it is recommended to use sample sizes of at least 100–200 subjects or at least a sample of three times the number of employed items, even in the favorable case of a good fit of attributes and factors (high commonality). If this requirement is not met—and it never was in the aforementioned studies, stable solutions can be expected only for the most critical factors, i.e., those carrying most of the variance, while all other factors, representing the more subtle aspects of room acoustic impression, are affected by a large sampling error (MacCallum et al. 1999).

Finally, any psychological measuring instrument requires an analysis of the psychometric qualities of the perceptual constructs and questionnaires based on them in terms of validity, reliability, and measurement invariance. Techniques and criteria for this purpose have been developed extensively in the social sciences (Vooris and Clavio 2017). Typical requirements for psychological questionnaire instruments comprise the use of *latent measurement models*, the demonstration of *convergent and discriminant validity* ("Do the scale's subdimensions actually measure what they are supposed to measure and are sub-dimensions sufficiently different from each other?"), as well as demonstrations of sufficient *construct reliability* ("How precise does the scale measure?") and *measurement invariance* (Millsap 2011) across time, stimuli, and populations of interest ("Are the scale's measurements independent of the experimental factors employed?").

In order to achieve and demonstrate an acceptable degree of validity, reliability, and measurement invariance, and to deal with different sources of measurement error (Schmidt and Hunter 1999), psychological scale development today typically relies on latent-variable models (Loehlin 2004). In this approach, it is assumed that every manifest measurement of a questionnaire item is, in fact, the expression of underlying latent psychological constructs. When at least three items are measured for any construct exclusively ("simple structure"), it is possible to not only estimate the degree of item-measurement error, construct loading and resulting construct reliability, but also to calculate error-free construct scores and work with these in later analyses. The latent-variable approach also allows checking for reliability across time (retest reliability), which is considered the most important indicator of reliability for scales since Cronbach (1947), and invariance across experimental conditions. Since past studies on room acoustics predominantly drew on principal-component analyses (PCA) and clustering techniques, where none of these tests is possible (Fabrigar et al. 1999), only a minority of reliability-, validity-, and invariance-related questions could be addressed systematically.

4 The Room-Acoustical-Quality Inventory (RAQI)

A multi-stage investigation was conducted by the Audio-Communication Group at TU Berlin to develop a language-based measuring instrument for the different qualities of room-acoustical environments for music and speech and to address the methodological gaps described above (Weinzierl et al. 2018). In a first step, expert knowledge from different professional domains in room acoustics was acquired by help of a focus group in order to provide a comprehensive terminology covering all aspects of the room-acoustical impact on music and speech performances. In a second step, listening experiments with acoustical experts and non-specialists were conducted using 35 rooms of different architectural types, different size, and different average absorption values in order to address the most important types of acoustic performance venues and their specifics. Different audio content was used, including solo music, orchestral music, and dramatic speech. The goals of the subsequent statistical analyses were to,

- Find an exhaustive list of verbal attributes that describes all relevant room acoustical properties
- Identify the best-suited items of this list to form a standardized measurement instrument
- Analyze the underlying dimensions of room-acoustical impressions
- Construct a measurement instrument based on these dimensions and corresponding items
- Demonstrate the reliability of the new instrument across and within raters
- Demonstrate measurement invariance of the new instrument across experimental conditions such as audio content type and subject samples
- Demonstrate sufficient discriminant validity of its subdimensions.

In order to realize this in an experimental setting that permitted controlling for any possible confounders, the study drew on room-acoustical simulation and auralization by dynamic binaural synthesis. The consensus vocabulary generated by the expert focus group consisted of 50 perceptual qualities related to the timbre, geometry, reverberation, temporal behavior, and dynamic behavior of room-acoustical environments, as well as overall, holistic qualities. While some attributes reflect lower-order qualities closely related to temporal or spectral properties of the audio signal ("loudness"), ("treble/mid/bass range tone color"), perceived "size", and "width" of sound sources), other attributes reflect higher-order psychological constructs, supra-modal, affective, cognitive, aesthetic, or attitudinal aspects such as "clarity", "intimacy", "liveliness", "speech intelligibility", "spatial transparency", or "ease of listening" (Weinzierl et al. 2018, SuppPub 1). For the listening experiment, binaural room-impulse response (BRIR) datasets were simulated for 35 rooms at 2 listening positions for solo music and speech. For the orchestral piece, 25 rooms at 2 listening positions were selected, leaving off 10 rooms where the stage area would not be large enough for an orchestra. Thus, in total, 190 room-acoustical conditions (rooms $\times$ listening positions $\times$ source characteristics) were simulated for the listening experiment. Fourteen of these 190 possible stimulus combinations were rated by each of the 190 participants in a balanced incomplete block design, using 46 items selected from the focus group terminology.

An exploratory factor analysis (EFA) based on the common factor approach was conducted to estimate the number of independent latent dimensions contained in the full-item data matrix. The scree- and Kaiser-criterion was used as a starting point for constructing a multidimensional measurement model. For each of the possible solutions, a series of confirmatory factor analyses (CFA) was conducted to consecutively remove single items from the measurement models up to a point where an implied removal would have led to less than three items per factor, or otherwise, the overall fit of the measurement model was already good. The latter was read from Root-Mean-Square Errors of Approximation (RMSEA), Comparative Fit Indices (CFIs), and Standardized Root-Mean-Square Residual (SRMR) coefficients, as well as congeneric Construct Reliability (CR), indicating the internal consistency of a factor construct, and Average Variance Extracted (AVE) indicating how well a factor explains the scores of its underlying items (Fornell and Larcker 1981).

The factor analysis suggests possible solutions with 4, 6 or 9 factors. These can be interpreted as a general room-acoustical *quality* factor, *strength*, *reverberance*, *brilliance* (4-factor solution), *irregular decay* and *coloration* (6-factor solution), *clarity*, *liveliness* and *intimacy* (9-factor solution). The corresponding item batteries consist of 14, 20, and 29 attributes as shown in Fig. 3. From a statistical point of view, the 6-factor RAQI scale with 20 items is the best compromise between a comprehensive assessment of the full complexity of room-acoustical impressions while at the same time ensuring sufficient statistical independence of the different factors.

With *strength* and *reverberance*, two of the sub-dimensions are omnipresent in the room-acoustical literature. Also, *clarity* and *intimacy* as additional factors have been frequently highlighted by previous studies (Hawkes and Douglas 1971; Lokki et al. 2012). With *brilliance*, *coloration*, and *intimacy* appearing as largely independent factors, it seems that timbre-related qualities play a greater role with more dimensions than previously assumed. The importance of perceived *irregularities in decay* and of *liveliness* as an independent construct has, however, hardly been considered so far.

9-factor RAQI	6-factor RAQI	4-factor RAQI	Factors	Items	Poles	W	I
9-factor RAQI	6-factor RAQI	4-factor RAQI	Quality	Liking	I like it – I don't like it	1.0	2
				Room acoustic suitability	suitable – not suitable	1.0	56
				Ease of listening	difficult – effortless	0.9	10
				Global balance	balanced – unbalanced	0.8	4
			Strength	Size	small – large	1.0	57
				Loudness	soft – loud	0.7	64
				Width	small – large	0.8	57
			Reverberance	Duration of reverberation	short – long	1.0	47
				Reverberance	dry – reverberant	1.0	54
				Strength of reverberation	weak – strong	1.0	51
				Envelopment by reverberation	weak – strong	0.7	48
			Brilliance	Brilliance	not brilliant – very brilliant	1.0	48
				Tone Color bright/dark	bright – dark	-0.8	-7
				Treble range characteristic	attenuated – emphasized	0.7	3
			Irregular decay	Flutter Echo	none – very strong	1.0	26
				Echo	none – very strong	0.7	40
				Irregularity in sound decay	none – very strong	0.9	32
			Coloration	Boominess	not boomy – very boomy	1.0	37
				Roughness	not rough – very rough	0.7	31
				Comb filter coloration	none – very strong	0.8	34
			Clarity	Temporal clarity	clear – blurred	1.0	10
				Spatial transparency	blurred – transparent	1.0	1
				Precision of localization	precise – diffuse	-0.8	-6
			Liveliness	Liveliness	dead – lively	1.0	11
				Spatial presence	low – high	1.0	63
				Dynamic range	small – large	0.9	50
			Intimacy	Intimacy	remote – intimate	1.0	-3
				Distance	close – distant	-0.8	51
				Warmth	cool – warm	0.5	3
Single items				Metallic tone color	not metallic – very metallic		
				Openness	open – constricted		
				Attack	soft – crisp		
				Richness of sound	low – high		

Fig. 3 The Room-Acoustical-Quality Inventory (RAQI). Four, six, and nine factors as possible sub-dimensions of room acoustical impression can be measured with questionnaires containing 14, 20, and 29 items, which are given with corresponding poles. Weights (W) and Intercepts (I) should be used to measure factors and for structural-equation analysis. Four additional single items with high retest reliability, which could not be assigned to any of the factors, are given below

In terms of psychometric quality, the factors of the 6-factor RAQI exhibit good across-rater consistency and within-rater stability. With regards to measurement invariance, scalar measurement invariance across measurement occasions could be demonstrated for a rather long distance of approximately 42 days. Scores from all RAQI sub-dimensions can thus be directly compared across studies as long as experimental conditions and test subject sample are identical. Similar results pertain to changes in experimental listening position: Scores taken from different listening positions in the same room did not differ systematically. Although the acoustical transfer functions might be quite different, as was demonstrated even for minor changes of the listening position (de Vries et al. 2001), listeners are able to identify the room and its acoustical properties as a consistent cognitive object.

5 Low Retest Reliabilities of Experts versus Laymen: A Problem of Language or a Problem of Perception?

As part of the RAQI development study, the listening test with 88 subjects of 190 in total was repeated six weeks later with identical stimuli. Based on these data, test-retest reliabilities could be determined, calculated as the correlation of measurements within individuals across time, as a measure of the precision, with which certain room-acoustical features could be evaluated. For a majority of the 46 items rated by all participants, the reliabilities turned out to be rather low. Only three items related to reverberation: "*reverberance*", "*strength*", and "*duration of reverberation*", exceeded values of $r = 0.7$, which is usually considered as a criterion of *good* reliability. Many other items, including popular ones in room acoustics such as "*sharpness*" or "*transparency*", turned out to be based on somewhat unreliable judgements ($r = 0.37/0.43$), using the variation over time within subjects as an indicator.

The low stability of most single item scores indicates that room acoustical impressions appear to be strongly influenced by time-varying situational factors, such as variations in attention, mental efficiency and distraction (*random response errors*) and variations in mood, feeling and mindset (*transient errors*, Schmidt and Hunter 1999). The extent of these psychological measurement errors, however, also depends on the expertise of the listeners. Since the aforementioned subject sample consisted of 60 music-interested non-specialists and 28 individuals with professional education in room acoustics, the relevance of this personal trait could be determined, showing a mean retest reliability of 0.50 across all items for non-specialists versus 0.59 for acoustical experts. Since there is no evidence for differences in the sensory performance between the two groups, the reasons for this difference have to be sought in the perceptual back-end, to pick up on a term from Sect. 1.

Laymen could assess properties, which are clearly room-related and accessible to everyday experience such as the *size* of the room ($r = 0.59$ vs. 0.58 for experts vs. non-experts), the occurrence of *echoes* ($r = 0.67$ vs. 0.68) and even the degree of *liking* ($r = 0.59$ vs. 0.62) with the same, sometimes even better reliability than experts. Whenever, however, the influence of the acoustical source and the influence of the room on the same auditory qualities had to be separated, such as when rating the "*brightness*" ($r = 0.65$ vs. 0.49) and the "*bass-range characteristic*" ($r = 0.64$ vs. 0.17), or the impact of the room on *temporal clarity* ($r = 0.72$ vs. 0.49) and *speech intelligibility* ($r = 0.70$ vs. 0.58), experts were clearly in the advantage. In these cases, the ability to judge this reliably depends on the extent to which *the performance* and *the room* are separated cognitive objects attracting differentiated attention. In addition, experts then also cultivate a specialized vocabulary, including attributes such as *comb-filter coloration* ($r = 0.56$ vs. 0.44) or the *spatial transparency* ($r = 0.56$ vs. 0.38), which are hardly required to describe everyday experiences with music and speech.

Hence, the question of whether the higher precision of experts in evaluating the acoustic properties of rooms is due to a better-trained perception or a more sophisticated vocabulary points to the same interwoven phenomenon: The cognitive performance in the separation of source and space requires a more sophisticated vocabulary, but the sophisticated vocabulary, in turn, can lead to a more differentiated perception.

6 The Perceived Quality of Performance Spaces

Most studies dedicated to room acoustic qualities have, in one way or another, also examined what makes up the *overall quality* of performance spaces. To avoid terminological confusion between the two concepts, Blauert has suggested distinguishing between the "aural quality" of a room and a set of "quality features" making up the "aural character" of the room (Blauert 2013). In the following, however, we will stick to a distinction between "quality" and "qualities", because these are fairly well established terms in the psychological literature.

Quality, in the sense of liking or preference, was already an issue in early investigations aiming at preferred values for the reverberation time of concert halls (Sabine 1906; Watson 1923; Bagenal 1925; Sabine 1928; Knudsen 1931). Multidimensional approaches have often tried to find correlations between the rating of individual attributes or factors and the overall *pleasantness* (Wilkens 1977), the degree of *enjoyment* (Hawkes and Douglas 1971) or the *preference* (Soulodre and Bradley 1995; Lokki et al. 2012) of the room acoustical impression. The relation between the rating of individual qualities and overall quality judgements, however, turned out to be dependent on many factors beyond the room acoustical properties, such as the musical repertoire, the musical performance, and the taste of individual listeners. Some of the studies could even identify different preference groups, one of which preferred more reverberance while the other preferred more clarity and definition (Lehmann and Wilkens 1980; Lokki et al. 2012).

The proportion of the variance in overall quality judgments which can be explained only by acoustic qualities of the rooms themselves was estimated by the authors of this chapter, based on the ratings of 190 participants, collected in the development of the Room-Acoustical-Quality Inventory (Weinzierl et al. 2018). Since an unspecific *quality* factor turned out to be one dimension in each factor solution, the relationship of this general factor to the other dimensions could be estimated. For this purpose, the measurement model of the RAQI was turned into an equivalent structural-equation model regressing the scores of the quality factor on the other factors to estimate their influence and the overall explained variance for the quality factor. Not only linear but also quadratic influences were tested, since psychoacoustic influences on cognitive percepts often show a u-formed (or inverse u-formed) relationship, for example, between preference and reverberation, where an optimal range and a decrease in quality on both sides is the most plausible relation. To account for this, the LMS approach for non-linear effects in structural-equation models was used (Harring et al. 2012).

In the 6-factor version of the RAQI shown in Fig. 3, the model was able to explain about half of the variance in *quality*—see Fig. 4. While *strength* and *brilliance* scores

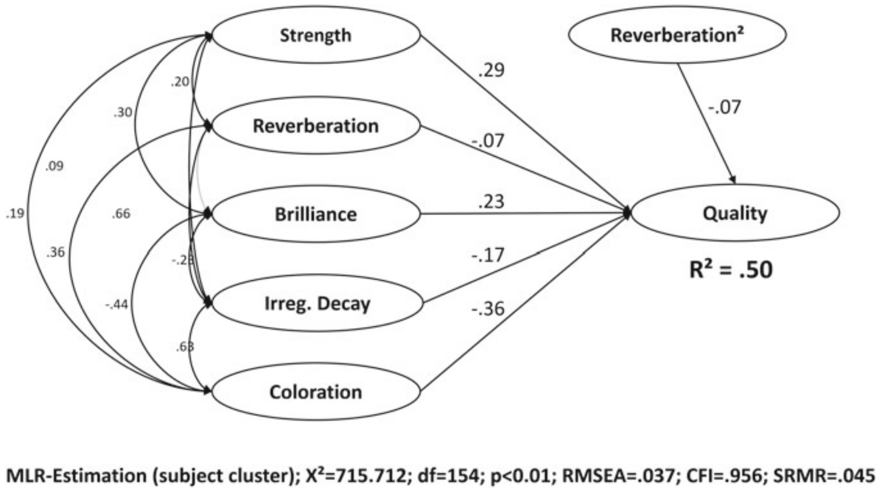

Fig. 4 Structural-Equation Model estimating the influence of five dimensions in the 6-factor-RAQI on the overall quality factor. Path coefficients are beta-weights/correlations. The parameters indicating the fit of the model are explained in Sect. 4

exhibit the largest positive influence on *quality* judgements, a decrease in quality arises with higher *coloration* and higher *irregular decay* values. *Reverberance* had both a linear and an inverted-u relationship to *quality*.

One should be aware that the influence of the individual factors—indicated by the beta weights—as well as the explanatory power of this *quality* model—indicated by the measure of determination, R^2—depends significantly on the properties of the stimulus pool from which it is derived. The less the presented rooms and the music-and-speech content correspond acoustically to the expectations of the listeners, the higher will be the proportion of the overall quality that can be explained solely by acoustic properties. Also, the sign and the value of the linear beta weights initially only indicate in which direction and to what extent parameters, for example, the rooms' reverberance, deviated on average from the perceived optimum as seen from the listeners' point of view. In order to obtain values for these relationships that correspond to a certain cultural practice, it is therefore vital to work with a stimulus pool that is an adequate representation of this practice.

With the stimulus pool used, fifty percent of the variance in *quality* could be explained by perceptual attributes of the room in the presented model. This part of the variance can be considered as the context-independent part of the overall preference judgements, not accounting for the musical repertoire, the musical performance, and the individual taste of the listeners. To explain the preference of music listeners for specific concert halls in a specific situation, a significantly extended model would thus be required. Although various potential influencing factors related to the cultural context of such an overall judgement have been proposed, for example, by Blauert (2013), who pointed out the importance of the *typicality* of concert halls as a result of two hundred years of Western concert culture, a comprehensive model for the overall aesthetic impression of performance venues, validated by empirical data, has not yet been proposed.

A promising candidate for such a model could emerge when taking into account that judgments about concert halls are always embedded in music-cultural practices in which the music piece, its performance, the performance space and the predisposition of listeners are intricately interwoven, and in their entirety shape the aesthetic judgment of a musical event. Thus, when music psychology examines the factors that influence the aesthetic judgment of a concert performance, the spatial and social context under which that judgment is made will always form a part of that judgment. Listeners can try to consider individual aspects and, for example, try to analytically separate the contributions of the sound source and the performance space to the perceived sound event (Traer and McDermott 2016). However, this separation will always be incomplete. Hence a reasonable approach will possibly lie in applying models that have already been proven empirically to explain the aesthetic judgment of music and music performances also to the evaluation of the *venues* in which music is performed.

Such an approach is exemplified by the model of Juslin et al. (2016), shown schematically in Fig. 5. It assumes that listeners make aesthetic judgments in particular situations in which they adopt what the authors call an “aesthetic attitude”. It is, not least, the concert ritual and the concert hall itself that encourages listeners to adopt this attitude. Once this condition is met, aesthetic processing may be influenced by several factors in the artwork, the perceiver, and the situation. These influences are mediated through the perception, cognition, and emotion of the listener. For one thing, a clear separation of these processes is difficult to draw, and the same musical cues can be processed perceptually (i.e., as sensory impressions), cognitively (i.e., depending on conceptual knowledge) and emotionally (i.e., aroused by other psychophysiological mechanisms). However, even more important is the observation that different listeners use different criteria that determine which of this information and which of those channels have an impact on the resulting aesthetic judgment.

These criteria can be related to varying degrees both to the musical work, to a performance, and to a performance space. One of these criteria is *beauty*. Concerning the performance space, beauty could be understood as the sum of the room acoustical qualities, for which a multidimensional measuring instrument has been developed with the RAQI (Fig. 3). Beauty, however, should not be identified with aesthetic value in general. Other criteria such as the degree of *originality* of a musical event and the related performance venue, the *skill* in its realization, the *typicality* with respect to performance traditions, the degree of *expression* and *emotional contagion*, and the *message* related to the socio-cultural connotations of the musical event, can play a major role for many listeners. In the study of Juslin et al. (2016), most listeners appeared to use a small number of three to five criteria in their judgments, and there were significant individual differences among the listeners, both in how many and which criteria were used.

It is tempting to assume that it is the individual choice of criteria which may account not only for the different aesthetic judgments about music as an integrated

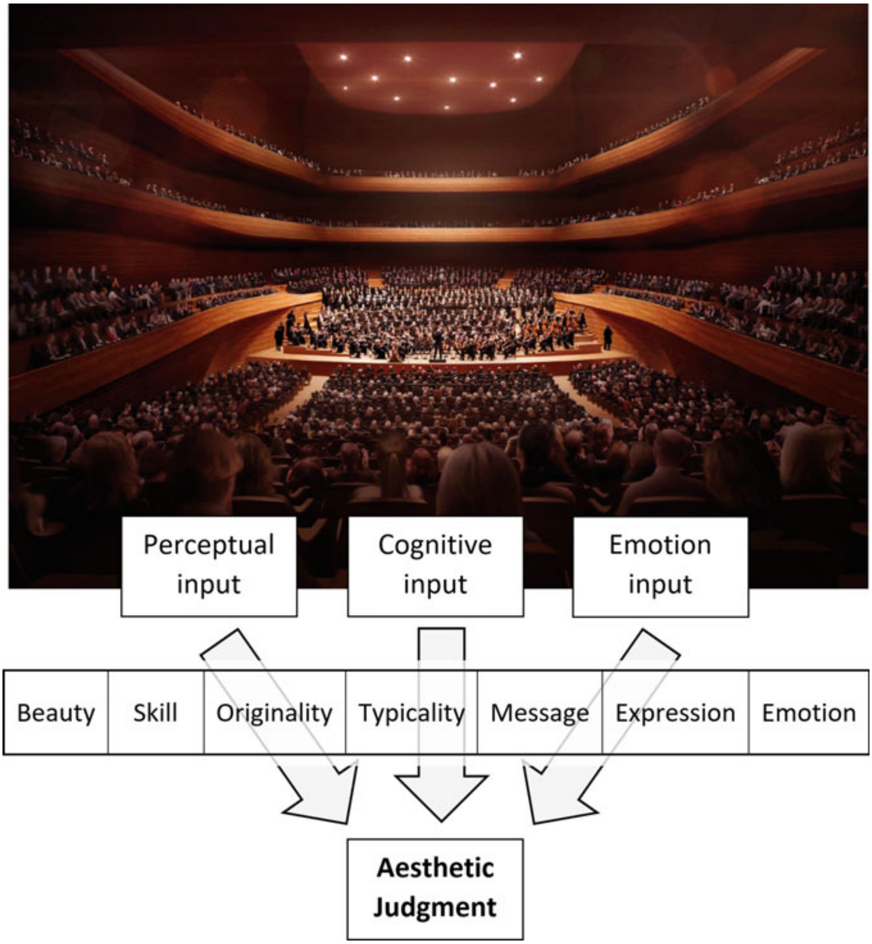

Fig. 5 A model for the formation of aesthetic judgments about music, according to Juslin et al. (2016). The analysis of a musical event is channelled through the perception, cognition, and emotion of the listener. Whether these inputs will affect the resulting aesthetic judgment depends on the listener's criteria, which act as filters for the processed information

experience but also for the different judgments about musical performance venues. Empirical verification of this assumption could be an essential contribution to a problem that might have been considered for too long only from a psychoacoustic perspective.

Acknowledgements The work reported here was produced within the research unit on "Simulation and Evaluation of Acoustical Environments (SEACEN)", supported by the Deutsche Forschungsgemeinschaft (FOR 1557). The authors are indebted to all colleagues in this project who contributed to this work, such as D. Ackermann, F. Brinkmann, D. Grigoriev, H. Helmholz, M. Ilse, O. Kokabi, L. Aspöck, and M. Vorländer, as well as to Jens Blauert for many inspiring discussions on the topic. Further, we want to thank two anonymous reviewers for their comments on this book chapter.

References

Bagenal, H. 1925. Designing for musical tone. *Journal of the Royal Institute of British Architects* 32 (20): 625–629.

Barron, M. 1971. The subjective effects of first reflections in concert halls–the need for lateral reflections. *Journal of Sound and Vibration* 15 (4): 475–494.

Barron, M., and A.H. Marshall. 1981. Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure. *Journal of Sound and Vibration* 77 (2): 211–232.

Beranek, L.L. 1962. *Music, Acoustics & Architecture*. New York: Wiley.

Berg, J., and F. Rumsey. 2006. Identification of quality attributes of spatial audio by repertory grid technique. *Journal of the Audio Engineering Society* 54 (5): 365–379.

Blauert, J. 2013. Conceptual aspects regarding the qualification of spaces for aural performances. *Acta Acustica united with Acustica* 99 (1): 1–13.

Bradley, J.S., and G.A. Soulodre. 1995. Objective measures of listener envelopment. *Journal of the Audio Engineering Society* 98 (5): 2590–2597.

Bureau of Standards. 1926. *Circular of the Bureau of Standards, No. 300. Architectural acoustics*. Washington: G.P.O. https://archive.org/details/circularofbureau300unse, https://archive.org/details/circularofbureau300unse.

Cronbach, L.J. 1947. Test 'reliability': Its meaning and determination. *Psychometrika* 12 (1): 1–16. https://doi.org/10.1007/BF02289289.

Dabrowska, E., and D. Divjak. *Handbook of Cognitive Linguistics*. Berlin, Boston: De Gruyter Mouton.

Dancygier, B. 2017. *The Cambridge Handbook of Cognitive Linguistics*. Cambridge, UK: Cambridge University Press

de Vries, D., E.M. Hulsebos, and J. Baan. 2001. Spatial fluctuations in measures for spaciousness. *Journal of the Acoustical Society of America* 110 (2): 947–954.

Evans, V., and M. Green. 2007. *Cognitive Linguistics. An Introduction*. Edinburgh: Edinburgh University Press.

Everett, C. 2013. *Linguistic Relativity: Evidence Across Languages and Cognitive Domains*, vol. 25. Berlin/New York: De Gruyter Mouton.

Fabrigar, L.R., D.T. Wegener, R.C. MacCallum, and E.J. Strahan. 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* 4 (3): 272–299.

Fornell, C., and D.F. Larcker. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18 (1): 39–50. https://doi.org/10.1177/002224378101800104.

Harring, J.R., B.A. Weiss, and J.C. Hsu. 2012. A comparison of methods for estimating quadratic effects in nonlinear structural equation models. *Psychological Methods* 17 (2): 193–214. https://doi.org/10.1037/a0027539.

Hawkes, R.J., and H. Douglas. 1971. Subjective acoustic experience in concert auditoria. *Acustica* 24 (5): 235–250.

Johnson-Laird, P.N. 1983. *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Harvard University Press.

Juslin, P.N., L.S. Sakka, G.T. Barradas, and S. Liljeström. 2016. No accounting for taste?: Idiographic models of aesthetic judgment in music. *Psychology of Aesthetics, Creativity, and the Arts* 10 (2): 157–170.

Kahle, E. 1995. *Validation d'un modèle objectif de la perception de la qualité acoustique dans un ensemble de salles de concerts et d'opéras (Validation of a perceptual model of the acoustic quality in an ensemble of concert halls and opera houses)*. Le Mans: Le Mans Universite. Ph.D. thesis.

Knudsen, V.O. 1931. Acoustics of music rooms. *Journal of the Acoustical Society of America* 2: 434–467.

Kuusinen, A., and T. Lokki. 2017. Wheel of concert hall acoustics. *Acta Acustica united with Acustica* 103 (2): 185–188.

Lehmann, P., and H. Wilkens. 1980. Zusammenhang subjektiver Beurteilungen von Konzertsälen mit raumakustischen Kriterien (Relation between subjective elaluations of concert halls and room-acoustical criteria). *Acustica* 45: 256–268.

Levinson, S.C. *Space in Language and Cognition: Explorations in Cognitive Diversity*. Cambridge: Cambridge University Press.

Loehlin, J.C. 2004. *Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis*. Mahwah: Routledge.

Lokki, T., J. Pätynen, A. Kuusinen, and S. Tervo. 2012. Disentangling preference ratings of concert hall acoustics using subjective sensory profiles. *Journal of the Acoustical Society of America* 132 (5): 3148–3161.

Lokki, T., J. Pätynen, A. Kuusinen, H. Vertanenen, and S. Tervo. 2011. Concert hall acoustics assessment with individually elicited attributes. *Journal of the Acoustical Society of America* 130 (2): 835–849.

Luizard, P., J. Steffens, and S. Weinzierl. 2020. Singing in different rooms: Common or individual adaptation patterns to the acoustic conditions? *Journal of the Acoustical Society of America*. 147 (2): EL132–EL137.

MacCallum, R.C., K.F. Widaman, S. Zhang, and S. Hong. 1999. Sample size in factor analysis. *Psychological Methods* 4 (1): 84–99.

Merimaa, J., and V. Pulkki. 2005. Spatial impulse response rendering I: Analysis and synthesis. *Journal of the Audio Engineering Society* 53: 1115–1127.

Millsap, R.E. 2011. *Statistical Approaches to Measurement Invariance*. New York: Routledge.

Minsky, M. 1977. Frame-system theory. In *Thinking. Readings in Cognitive Science*, ed. P.N. Johnson-Laird and P.C. Wason. Cambridge: Cambridge University Press.

Murphy, G. 2004. *The Big Book of Concepts*. MIT Press.

Noble, A.C., R.A. Arnold, J. Buechsenstein, E.J. Leach, J.O. Schmidt, and P.M. Stern. 1987. Modification of a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture* 38 (2): 143–146.

Osgood, C.E., G.J. Suci, and P.H. Tannenbaum. 1957. *The Measurement of Meaning*. Urbana, Ill.: University of Illinois Press.

Pedersen, T.H., and N. Zacharov. 2015. The development of a sound wheel for reproduced sound. *Audio Engineering Society Convention*. 138 Preprint No. 9310.

Pulkki, V. 2007. Spatial sound reproduction with directional audio coding. *Journal of the Audio Engineering Society* 55: 503–516.

Sabine, P.E. 1928. The acoustics of sound recording rooms. *Transactions of the Society of Motion Picture Engineers* 12 (35): 809–822.

Sabine, W.C. 1900. Reverberation. *The American Architect and Building News*. 68: 3–5, 19–22, 35–37, 43–45, 59–61, 75–76, 83–84.

Sabine, W.C. 1906. The accuracy of musical taste in regard to architectural acoustics. *Proceedings of the American Academy of Arts and Sciences* 42 (2): 53–58.

Schärer Kalkandjiev, Z., and S. Weinzierl. 2013. The influence of room acoustics on solo music performance: An empirical case study. *Acta Acustica united with Acustica* 99: 433–441.

Schärer Kalkandjiev, Z., and S. Weinzierl. 2015. The influence of room acoustics on solo music performance: An experimental study, Psychomusicology: Music. *Mind and Brain* 25 (3): 195–207.

Schmidt, F.L., and J.E. Hunter. 1999. Theory testing and measurement error. *Intelligence* 27 (3): 183–198. https://doi.org/10.1016/S0160-2896(99)00024-0.

Somerville, T. 1953. Subjective comparisons of concert halls. *BBC Quarterly* 8: 125–128.

Somerville, T., and C.L.S. Gilford. 1957. Acoustics of large orchestral studios and concert halls. *Proceedings of the IEE* 104: 85–97.

Sotiropoulou, A.G., R.J. Hawkes, and D.B. Fleming. 1995. Concert hall acoustic evaluations by ordinary concert-goers: I, Multi-dimensional description of evaluations. *Acta Acustica united with Acustica* 81 (1): 1–9.

Soulodre, G.A., and J.S. Bradley. 1995. Subjective evaluation of new room acoustic measures. *Journal of the Audio Engineering Society* 98 (1): 294–301.

Spearman, C., and L.W. Jones. 1950. *Human Ability*. London: Macmillan.

Susini, P., G. Lemaitre, and S. McAdams. 2011. *Psychological Measurement for Sound Description and Evaluation in Measurement with Persons: Theory, Methods, and Implementation Areas*, eds. B. Berglund and G. B. Rossi. New York: Psychology Press.

Thiering, M. 2018. *Kognitive Semantik und Kognitive Anthropologie: Eine Einführung [Cognitive semantics and cognitive anthropology: An introduction]*. Berlin: De Gruyter.

Tkaczyk, V., and S. Weinzierl. 2019. Architectural acoustics and the trained ear in the arts: A journey from 1780 to 1830. In *The Oxford Handbook of Music Listening in the 19th and 20th Centuries*, eds. C. Thorau and H. Ziemer. New York, NY: Oxford University Press.

Torgerson, W.S. 1952. Multidimensional scaling: I. Theory and method. *Psychometrika* 17 (4): 401–419.

Traer, J., and J.H. McDermott. 2016. Statistics of natural reverberation enable perceptual separation of sound and space. *Proceedings of the National Academy of Sciences* 113 (48): E7856–E7865.

Vooris, R., and G. Clavio. 2017. Scale development. In *The International Encyclopedia of Communication Research Methods*. eds. C.S.D.J. Matthes and R.F. Potter. American Cancer Society.

Watson, F.R. 1923. *Acoustics of Buildings*. New York: Jon Wiley and Sons.

Weinzierl, S. 2002. *Beethovens Konzerträume. Raumakustik und symphonische Aufführungspraxis an der Schwelle zum modernen Konzertwesen [Beethoven's concert halls. Room acoustics and symphonic performance practice on the threshold to modern concert life]* (Bochinsky, Frankfurt am Main).

Weinzierl, S., S. Lepa, and D. Ackermann. 2018. A measuring instrument for the auditory perception of rooms: The Room Acoustical Quality Inventory (RAQI). *Journal of the Acoustical Society of America* 144 (3): 1245–1257.

Wilkens, H. 1977. Mehrdimensionale Beschreibung subjektiver Beurteilungen der Akustik von Konzertsälen [Multidimensional description of subjective evaluations of the acoustics of concert halls]. *Acustica* 38: 10–23.