19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

AUDITORY QUALITY OF PERFORMANCE SPACES FOR MUSIC —THE PROBLEM OF THE REFERENCES—

PACS: 43.55-n, 43.55-Hy, 43.66-Ba

Blauert, Jens[1] & Jekosch, Ute[2]
[1]Inst. of Communication Acoustics, Ruhr-Universität Bochum, Germany, jens.blauert@rub.de
[2]Chair of Communication Acoustics, IAS, TU Dresden, Germany, ute.jekosch@tu-dresden.de

ABSTRACT
In the last two decades, sound quality has become a hot item of discussion in acoustics. For architectural acousticians this does not come as a surprise, as they have always dealt with this issue, particularly regarding aspects of *auditory-scene quality* and *transmission quality* – since an enclosed space is basically a three-dimensional transmission system. In sound-quality research, current models of the quality-judgment process are based on a comparison of the character of sounds with the character of appropriate references. For a more general modelling approach, it has further been proposed to put all the individualities and particularities, which are due to specific items, tasks, functions, and individual observers, into the character of the references. In this contribution, some problems regarding the identification and specification of proper references are discussed in some detail. Thereby spaces for musical performances are taken as examples. Issues like typicalness, functional adequacy, and tradition, are briefly addressed.

1. THE CONCEPT OF QUALITY IN ACOUSTICS
It has become clear, particularly during the last decade, that quality is not a feature inherent to an item. Rather, it is something that evolves from a process in which judgments take place [4, 6, 12]. The judgements concern comparison of the item under consideration – or, more precisely, all of its recognized features – with some kind of a reference pertinent to it.

The following definition, a translation after [13], may be taken as an example of this modern concept of quality in acoustics. It was originally formulated for speech quality, but replacing speech by sound makes it more general without jeopardising the basic idea.

> "*Sound Quality* is the result of an assessment of the suitability of a sound item, considering all of its recognized and nameable features and feature values. The assessment specifies as to which amount the sound item under consideration complies with a reference arising from aspects such as individual expectations and/or social demands and/or pragmatic necessaries − again considering all recognized and nameable features and feature values"

In trying to transform this concept into a manageable algorithm, a taxonomy in form of a block diagram has been proposed [5] – as illustrated in Fig. 1. This block diagram is likewise intended to provide a basic architecture for an algorithmic model of sound-quality assessment.

The different blocks in the diagram represent

(a) The character of the sound item under consideration
(b) The character of the reference that the sound item is compared to
(c) The judgment process

The sound character is formed from all recognized, nameable and – if possible – quantitatively assessed (i.e. measured) features of the sound item under consideration*). The relevant data are usually provided in the form of profiles of feature values, such as an acoustical profile and a psycho-acoustical profile.

*) *Footnote*: In older literature on sound and speech quality the totality of features of an item is sometimes called "quality" instead of character. This usage of the term quality is now discarded [5, 9, 14, 15].

Further profiles might be added, e.g., an emotional profile, a semiotic profile (…incl. semantics). Acoustical data can be collected with physical measuring instruments. Perception- or/and cognition-based data are typically rendered by panels of experts who have been trained to recognizing the features under consideration selectively, and to assess them quantitatively**.**

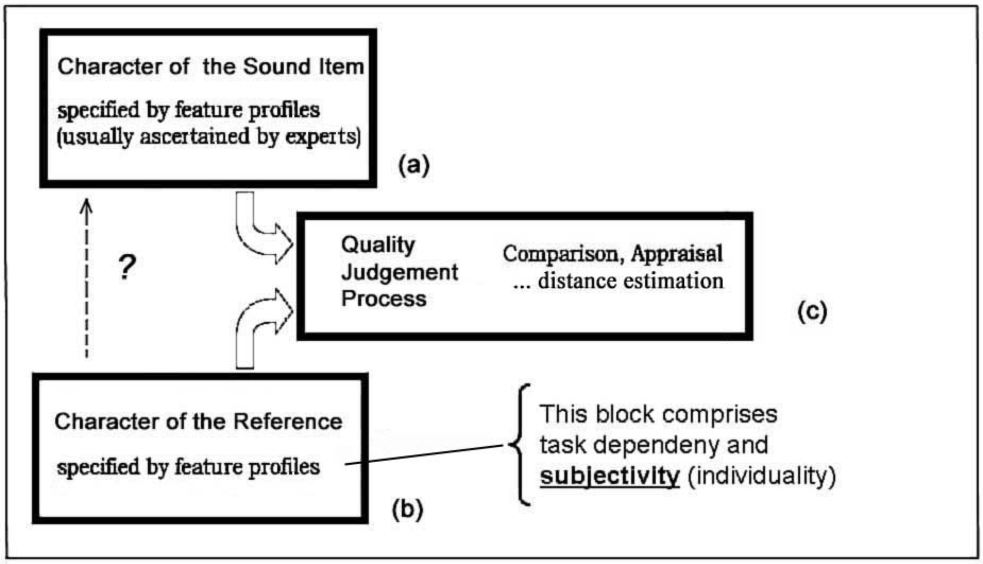


Figure 1 –Architecture for a model of sound-quality assessment

The character of the references is much harder to evaluate, since it usually contains a set of concepts rather than a set of measurable sensory perceptions or, even less so, physically defined data. Since references often have a profound basis in cognition, it is actually required to measure peoples' thoughts to arrive at quantitative reference profiles. Apart from that, emotions have to be taken into account as well. To this end, when investigating the character of references, psychological and sociological expertise is usually needed. This fact is, certainly, not always well accepted by engineers, as the latter are used to specifications compiled from purely physical, i.e. acoustic data. Unfortunately, this is only rarely possible when it comes to quality judgments regarding the "acoustics" of spaces for musical performances – as we shall discuss in detail below.

Once the characters of both the sound item and the reference are known, the quality judgement can take place. The process may be thought of as the determination of a multidimensional distance between sound item and reference. Yet note, a distance of zero does not necessarily represent the highest quality – this is due the fact that expectations may well be exceeded.

In the course of proposing Fig. 1 as a basic architecture for an algorithmic recognizer and predictor of quality in acoustics, the following further simplifications have been discussed [5]. Blocks (a) and (b) could be built in such a way as to function task-independently and objectively. In other words, these blocks could be designed as components which function in a very general way, i.e. disregarding particularities of items, individuals, tasks, functions and situations for which quality-appraisal is to be performed. Under this regiment, all individuality as well as all particularities concerning specific tasks, functions and situations would then be put into the reference block. In this way, we get a model where the blocks (a) and (b) are fairly universal and all particularities are put into block (c).

What has not yet been discussed, but should be noted in passing by, is that some feedback from block (c) to (a) may have to be included into the model, as the data required to be included in (a) may depend on what is contained in (c) in an actual case.

In any case, the paramount role of the reference for any quality judgment is obvious. For this reason, the question of references is the main object of discussion in this contribution. The authors are sure that a lot of future research effort will have to be put into the problem of proper references.

2. SOUND-QUALITY REFERENCES – ORDERED BY THEIR LEVEL OF ABSTRACTION

References in the context of different categories of quality in acoustics can have different levels of abstraction and, consequently, the measuring methods to be applied are different. This fact will be elucidated in the following – see Fig. 2.

The author would like to note as this point that the following reasoning should be taken as an essay. Further analysis will be necessary to develop the reported ideas into a valid theory. This also holds for the terminology used.

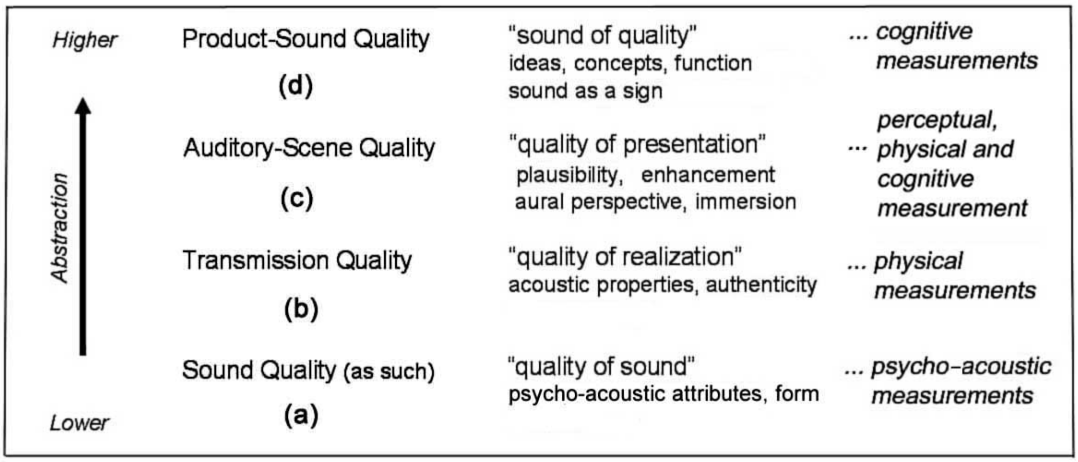


Figure 2 –Sound-quality categories – ordered by the level of abstraction

Along our lines of thinking, the lowest level of abstraction of the reference is given when basic psycho-acoustic attributes are judged upon, such as loudness, pitch, sharpness, roughness, and timbre. This is the case when the *sound quality* concerned is the quality of the auditory event *as such*, i.e. in its pure, non-interpreted form – see Fig.2 (a). The assessment methods in this case are perceptual to be sure, namely, psycho-acoustical ones. In other words, they require listening tests which employ basic psychometric methods.

A higher level of abstraction occurs in connection with instrumental (…physical) measurements. Instrumental measurement methods have been developed for providing ultimate objectivity, i.e. independence of the results from a particular laboratory or experimenter. It is this kind of data, engineers would generally prefer – if this were feasible. These kinds of measurements are, for example, used when the *quality of transmission* is the issue of interest – Fig. 2 (b). This is, for instance, often the case in information technology. Physical (acoustical) reference data may then be available.

We have termed the next higher level of abstraction *auditory-scene quality* here – Fig.1 (c). What is actually meant is the quality as to which the *presentation* of the auditory scene meets the reference. Hereby, it is not necessarily aimed at authentic reproduction, rather terms like plausibility, perceptual coherence, presence and immersion become issues of interest. The measurement methods to capture these issues are perceptual ones, though more complex than those applied under level (a) – but also cognitive ones.

An even higher level of abstraction is reached when the role of sounds as signs comes to the fore [4, 6, 10, 11, 12, 13]. This is, for example, the case for *product-sound quality* – see Fig. 2 (d). Products sounds are considered to have a high quality when they signal to the potential or actual product user that the product itself has a high quality. In these cases, it is not the sound quality as such that is the issue of interest. Rather, the focus lies on the fact that the product sound supports the quality of the product by acting as a sign for the actual product quality. The methods to measure these effects and to transform the results into profiles to be included into references for quality assessment draw heavily upon cognition. It certainly takes expertise in cognitive psychology (incl. semiotics) to manage these tasks.

3. REFERENCES FOR THE ACOUSTICS OF PERFORMANCE SPACES FOR MUSIC

We shall now attempt to apply the ideas expressed above to the problem of references for room-acoustics quality, i.e. the so-called quality of the "acoustics" of performance spaces for music. Actually, the character of the references in this context show elements of all levels of abstraction as laid out in Section 2. This holds for real spaces as well as for those which have been electronically modified or electronically created (e.g., virtual spaces).

(a) Sound-quality level

As has been elucidated above, the features at this abstraction level are those which are purely psycho-acoustic i.e. without any meaning or interpretation attached to it. In [16], such features have been named *unbiased.* Psycho-acoustics items with relevance to room acoustics are, e.g, loudness, roughness, tonal balance, spaciousness (… auditory source width), reverberance, audible echoes. However, it has to be mentioned that doubts have been expressed as to whether unbiased perceptual items really exist at all. At least with meaningful signals, such as music, this is rather unlikely. Nevertheless such items, sometimes called *sensations*, are often uncritically discussed as valid quality features in the room-acoustics literature.

(b) Transmission-quality level

This is the level at which physical measurements are taken. As we have discussed above, physical measurement methods have been designed for objectivity, i.e. to exclude any individuality of judgements. Measurement methods that would produce results which depend on the individual observer are not tolerated at this physical level. Consequently, physical measurements do not measure percepts directly. They concern with items in the physical world – so called quality elements [13]. In other words, they represent specific constructs. This is why they are of higher abstraction level than sensations. For a closer discussion of this argument see [3]. We have adopted the term transmission-quality level here, as these items typically refer to the (system) response of performance spaces to specific text signals – that is, the spaces are considered as transmission systems. Typical examples for items to be measured are sound-pressure level, impulse responses and transfer functions, reverberation time, clarity, and lateral energy fraction. Such physical, i.e. instrumentally measured data are in use, for example in the consultancy busyness, to predict perceptual features, but unfortunately, the correlation is usually unsatisfactory – compare for example [8]. Nevertheless, they are useful in the construction process of spaces for musical performance, since they may, for instance, help to track the effects of modifications in the design and its realisation.

(c) Auditory-scene-quality level

This level is, roughly, the one at which audio expert such as sound engineers and/or tonmeisters use to act, for example, when producing musical recordings or arranging play-backs for musical performances. They design the auditory scene such as to achieve a desired artistic effect, thereby certainly taking psycho-acoustic and physical criteria into account – but always drawing upon their rich professional experience. In other words, they certainly interpret what they hear, and then they design and realize the auditory scene as a coherent entity. Quality features at this level of abstraction may be such as, auditory transparency, balance of timbre, clear auditory perspective (…incl. proper distance cues), plausibility, immersion; further enhancement of specific features, such as word intelligibility, presence, immersion, spatial impression. Although perceptual judgements are used in this context, they are again no longer purely psycho-acoustical ones − as they take the specific program material into account. Thus, there is indeed very much interpretation going on in the context of the realization of an auditory product with optimum artistic effect – among other things by considering emotions, actions and prior knowledge of the prospective listeners. The musical program itself – i.e. the composition – is not part of the sound engineer's job. In other words, the sound engineers do their best to transmit and present artistic contents effectively, but they do not create the contents themselves.

(d) Product-sound-quality level

Although the term product-sound is mainly used for industrial products, the idea can be applied to spaces for musical performance as well. The point is that, at this level, the sign carrier is the issue, i.e. the sound is considered a medium for communication. In the case of music, these are is the artistic form, content, and function, as the composer had in mind when writing the composition. Product-sound quality, in this sense, includes the instruments selected, the particular players of these instruments, plus the space where the performance takes place. This is certainly the highest level of abstraction in the context of quality in acoustics. Cognitive assessment methods become necessary here. Product-sound quality becomes an item for the architectural acoustician and/or sound engineer when the performance spaces, and the auditory effects assigned to them, become an integrated part of the musical program material, i.e. the composition. This is regularly the case with modern electronic music, but there is also music for acoustical instruments where spatial features are part of the composition – see [7] for some interesting examples.

4. CONSIDERATIONS REGARDING THE COGNITIVE COMPONENTS OF REFERENCES

As to the character of the references for "good acoustics", items at all four levels of abstraction are in use. Yet, while psycho-acoustic and physical criteria are usually well known to architectural acousticians, this is to a lesser degree the case when it comes to criteria of higher abstraction. Audio experts such as sound engineers and/or tonmeisters are usually better experts at the auditory-scene-quality level, while cognitive psychologists, semio-acousticians, sound designers and, of course, musicians may be the experts of choice at the product-sound quality level. In the following, we pick a few representative examples for issues that come up for discussion when dealing with these levels of higher abstraction, particularly the cognitive level.

Typicalness and functional adequacy

Quality requires *typicalness* (…typicality) as a prerequisite. For example, a good steak is expected to taste like a steak. If it tasted like an excellent Peking duck, it would certainly not be considered a good steak. To produce dishes which display fine shades of tastes, yet typical of their kind, makes a good chef. The typical sound of the *Vienna Musikvereinssaal* is regarded excellent for orchestral concerts of the classic/romantic genre, but may be less adequate for pop concerts. Question: Is a good room acoustician somebody who is able to produce rooms which sound typical as compared to prototypes which are accepted to be excellent with regard to specific performances?

There is something to it, but it is not that simple. The quality of the acoustics of a hall must certainly be rated in the context of what is typically expected for specific types of performances. German speaking acousticians take account of this fact by using the term *acoustic qualification (Hörsamkeit)* instead of acoustic quality. Yet, it finally depends on the individual judgement what actually makes the acoustical qualification of a hall. In cases of well established acoustic traditions, different listeners share similar expectations. They have developed an internal prototypical sound image and, in comparing the sound to be judged with this prototype, they are able to assess a hall quasi *objectively,* i.e. there is high agreement between different listeners. In other cases, where tradition is less pronounced, listeners judge more *subjectively,* i.e. the individual judgement differ more among each other.

It may be helpful at this point to look at what architects do when designing buildings. First of all, they aim at constructing buildings for specific functions, thus arriving at specific built forms (…functional adequacy, form follows function). Often the decision for a specific form evolves gradually over a series of projects and then stabilizes at a generally accepted type of room, such forming a tradition. Good *traditions* evolve as the result of an optimization process. The traditions persist as long as the prerequisites, e.g., the architectural hardware, persist. The range of accepted forms is restricted in these cases. An extreme is given when historic buildings are rebuilt or restored authentically.

In a less traditional context, architects have more freedom in their design. Nevertheless, for example in the case of concert halls, even a very unconventional hall must still show typical features of a concert hall – otherwise it would not be one. Obviously, in such a case, *typicalness* does not necessarily mean that the sound of a hall is identical with or even close to that of a prototype – yet, that it shares some essential features with it that are invariant across a class of halls. These *invariant features* have something to do with the function of the halls, i.e. with their ability to support specific kinds of acoustical performances. It is worth noting that these invariant features are certainly cognitive ones, as, although the sound signals at different seats are very different, the listeners can nevertheless usually identify the hall from most seats. These invariant feature have still to be identified in future research. They are certainly of paramount importance for the characters of quality references.

Listening tradition and aesthetic form
In the course of an optimization process across a series of architectural projects and their realization and use, there are also the processes of evolving of *listening traditions* in the audiences. Since the tradition plays a decisive role in the context of listener expectations, and thus becomes part of the reference characters for quality judgments, the following line of thinking may be noteworthy at this point. To this end, we distinguish in the following between four periods to present the essence of our arguments − although finer categorization would of course as well be possible.

Period 1: During the period in time when the so-called classical music originated, i.e. at the end of the 18th century, music was played with acoustical instruments that had developed to their classical form in a process which lasted for quite some decades and was determined by technological feasibility of that period. For these instruments, music was written to be performed in halls that, as well, were determined by the *hardware* possibilities of these times, and then optimized. This led to the classical *shoe-box*-shaped concert hall with a rectangular structure, namely, brick wall and a wooden ceiling. Altogether, the listening tradition for classical music had almost two centuries to establish itself firmly – and was then traded on up to our times. Figure 3 (a) gives an overview of the architectural features and of the functional capabilities of halls of this kind.

Hardware, persistant (bricks, stones and wood)	
Life span > 100 years	
Rectangular "shoe-box" form (Basilica)	
Sound sources:	acoustic musical instruments human voice
Program:	baroque, classical and romantic music

(a)

Hardware, persistant (concrete)	
Life span > 100 years	
Fan-shaped and arena-shaped plans	
Sound sources:	acoustic musical instruments human voice
Program:	as with classical halls, plus modern music

(b)

Figure 3 –Features of spaces for musical performance without electro-acoustics
(a) classical shoe-box halls, (b) fan-shaped and arena-shaped halls

Period 2: Since about the beginning of the 20th century, construction technology allowed for larger halls, due to the availability of steal and concrete. As visibility (…sight lines) often gained precedence over the acoustics, *fan-shaped* and *arena-shaped* plans were usually favoured. Nevertheless, this period of more a century by now also produced a number of commonly accepted sound traditions. Halls from these times that are rated excellent acoustically are rare, but the exception makes the rule, for example, *Carnegie Hall* in New York. Today there is a strong tendency to modify these halls during refurbishment – what, by the way, offers good job opportunities for qualified acoustical consultants. See Fig. 3 (b) for the architectural features and of the functional capabilities of such halls.

Period 3: At about 1920, i.e. later in the 20th century, electro-acoustics arrived at the musical scene – the vacuum tube as prerequisite for amplification was invented about 1910. This made it possible to develop *analogue* electrical instruments and audio systems, though with a limited variability with respect to the musical sounds that they produced and/or provided to the halls. For example, the *Hammond* organ, certain electric guitars and electric pianos, the steel-plate reverberation unit, are still considered typical for this period, which, however, lasted not for a century but only for a little more than four decades so far. An overview of pertinent acoustic features and functional capabilities is given in Fig. 4 (a). Please note at this point that electro-acoustics became a determinant of the "acoustics" of hall besides their built architecture!

Hardware, slowly variable	
Life span ≈ 25 years	
Acoustic modification of classical and modern halls	
Sound sources:	amplified musical instruments amplified human voice
Program:	amplified, electroacoustically modified music, and music generated by analogous electronic devices

(a)

Software, instantly variable	
Life span ≈ a few years	
Acoustic modification of classical and modern halls, auditory virtual environments	
Sound sources:	amplified musical instruments amplified human voice sampled and electronically generated sounds
Program:	amplified, electro-acoustically modified music, all kinds of electronic music

(b)

Figure 4 –Features of spaces for musical performances including electro-acoustics
(a) analogue electro-acoustic equipment, (b) digital electro-acoustic equipment

Period 4: Today's technology now takes advantage of digital hardware and signal-processing algorithms. The possibilities to create, vary and manipulate instrumental sounds (…digital synthesis) and to modify the auditory conditions in the spaces for performance (…electronic architecture) are almost unlimited. Further, for creating and changing sounds it is no longer necessary to change hardware. *Software* does the job. Due to this fact, variations can be performed instantly and with no physical effort – the mass of electrons is negligible. Sound can therefore be changed all times. This may have the consequence that tradition can hardly develop, as everything is always on the move and there is not enough persistence to gather hearing experience.

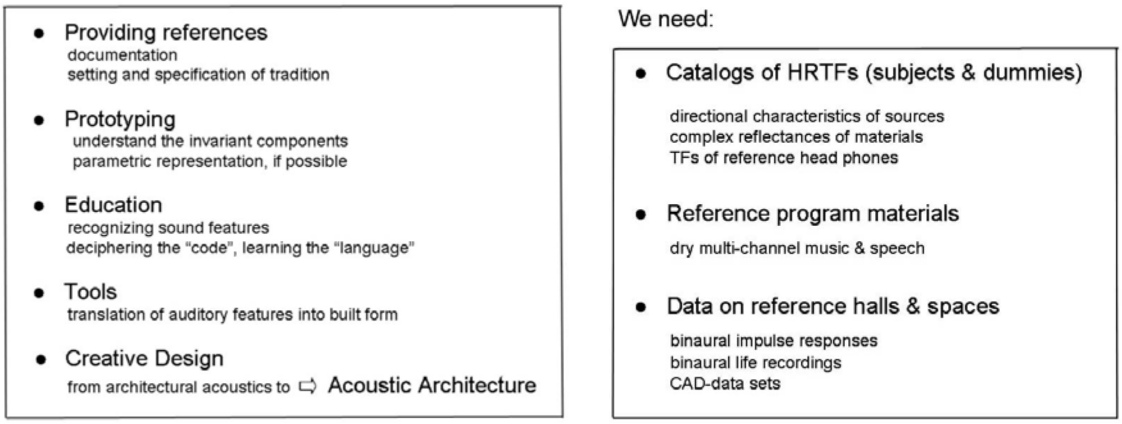
5. SOME CONSEQUENCES

The listeners react to non-traditional listening scenarios in such a way that they try to classify and categorize what they hear. However, the categories that they arrive at in this way do, more and more represent the presence of *fashion* rather than longer-lasting lines of cultural development. While in the sixties, a listener would have described his preferred kind of music with either classical music or jazz (or both), these days one tends to hear statements like: music of the fifties, the sixties, or the seventies. or, possibly, Heavy Metal vs. Hard Rock. The life cycles of musical stiles have become considerably shorter.

Yet, it is essential for the artistic effect of music that the listeners are acquainted with the *musical language* of the music under consideration. *Aesthetic form* finds its expression in the variations of the musical attributes within a (proto)typical range, such that typicality is preserved. One must be capable of decoding the *language of the music* to be able to grasp the aesthetic contents of music. This requires that the language of music must be taught, studied and learned, i.e. the audience must be educated. Please note at this point that musical language is to a large extent stile-specific. Listeners who understand classical music, do not necessarily understand jazz as well.

The less the listeners' understanding of the listener, the lower the abstraction level of the character their reference. For example, being exposed to modern electronic music without prior experience leads to a listening situation where basic psycho-acoustic features come to the fore that experts of this kind of music would usually not pay any attention to – as they are more focussed on the function than on the form.

Consequences for the consultants are that they have to guide and educate the listening experience of the clients in the course of their consultancy activities. A list of tasks, following from this idea, is given in Fig. 5 (a). Please note that the traditional one, namely translating acoustical features into built architecture, is only one out of five tasks here. To be able to provide references to the client and to the public, data as listed in Fig. 5 (b) should be readily available to consultants. It would thus be a good idea to establish a data bank at an adequate site, in order to make these data widely available in the field. This is particularly necessary since computer modelling, incl. *auralization*, has become a common practice in acoustical design.



(a) **(b)**

Figure 5 –From architectural acoustics to acoustic architecture
(a) relevant tasks of the consultant/designer, (b) support material needed

Further, acoustical consultants have to consider features of higher level of abstraction than they used to do traditionally. In other words, in addition to psycho-acoustical and physical criteria, they more and more may have to include higher-level perception and cognition into their considerations. Skills as those that are found in sound engineers and/or cognitive psychologist and musicians may have to be acquired by them – or at least be adequately understood. In other words, the may have to deal not only with form but with function as well. One of the current authors had thus stated some time ago that architectural acoustics will soon develop into acoustic architecture [1], i.e. the acoustical consultants will act as architects in their own right.

Recently, the term *aural architecture* has been proposed for this field at large [7]. The authors of this paper could not agree more.

References

[1] Blauert, J, (1994) Auditory perception and room acoustics: What next? (invited lecture). Wallace–Clement–Sabine–Centennial Symp., 27th meeting Acoust. Soc. Amer., Cambridge MA
[2] Blauert, J.(2000), Audio technology & virtual reality, Proc. AGORA 2000 – L'espace compose, IRCAM, Paris
[3] Blauert, J. (2006) Reality, virtual reality … and the roots of psychoacoustics, Proc. 2nd Int. Conf. on the Perceptive Quality of Systems, 149−151, ISCA, Grenoble
[4] Blauert, J. & Jekosch, U. (1996) Sound-quality evaluation – a multi-layered problem, ACUSTICA/acta acustica 83, 747–753
[5] Blauert, J. & Jekosch, U. (2003), Concepts behind sound quality: some basic consideration, Proc. InterNoise`03, 72-76, KSNVE/ASK, KR−Seoul
[6] Blauert, J. & Jekosch, U. (2005), Noise quality: What's the problem? Proc. 3rd Congr. Austria–Alps–Adria Acoust. Soc., 18-25, HR–Opatje
[7] Blesser, B. & Salter, L.-R. (2007) Spaces speak, are you listening? Experiencing aural architecture, MIT Press, Cambridge MA
[8] Heß, W. (2006) Time-variant binaural-activity characteristics as indicator of auditory spatial attributes. Doct. diss., Ruhr-Univ. Bochum, dissertation.de – Verlag im Internet, Berlin
[9] ISO 8402 (1994), Quality management and quality assurance – vocabulary
[10] Jekosch, U. (1999), Meaning in the context of sound-quality assessment. ACUSTICA/acta acustica 85, 681-684
[11] Jekosch, U. (2005) Assigning meaning to sounds – semiotics in the context of product-sound design. In: J. Blauert, ed., Communication Acoustics, Springer Berlin–Heidelberg–New York
[12] Jekosch, U. (2005) Basic concepts and terms of "quality", reconsidered in the context of product-sound quality, ACUSTICA united with acta acustica 90, 999–1007
[13] Jekosch, U. (2005), Speech quality – assessment and evaluation, Springer, Berlin–Heidelberg–New York
[14] Letowski, T. (1989) Sound-quality assessment: concepts and criteria, preprint 2825, Proc. 87th Conv. Audio Engrn. Soc.
[15] Pedersen, T. H. (1998) Optimisation of perceived product quality, Proc. EuroNoise'98, 633–638, Munich
[16] Zwicker, E. (1991) A proposal for defining and calculating the unbiased annoyance. In: A. Schick, ed. Contributions to psychological acoustics, 187−202, Bibliotheks- und Informationssystem der Univ. Oldenburg, Oldenburg