inter.noise 2000
The 29th International Congress and Exhibition on Noise Control Engineering
27-30 August 2000, Nice, FRANCE

I-INCE Classification: 0.0

PERCEPTUAL SOUND QUALITY EVALUATION

M. Bodden

Ingenieurbuero Dr. Bodden, Ursulastr. 21, D-45131, Essen, Germany

Tel.: +49 201 430 84 17 / Fax: +49 201 430 84 18 / Email: db@product-sound.de

Keywords:
SOUND-QUALITY, PERCEPTION, GEAR RATTLE

ABSTRACT
The perception of Sound Quality is not only based on the pure physical signal, it also depends on other sensorial modalities and even non-sensorial factors. Sound Quality evaluation thus becomes a complex, difficult and interdisciplinary task, and methods to evaluate Sound Quality cannot only restrict themselves to the acoustical signal. Other modalities and the specific situation and background of the subject have to be considered. Anyhow, for the application in the industrial environment the corresponding needs and restrictions have to be considered: methods have to be time-efficient, render results with a sufficient accuracy, and give direct clues on how to improve products. The process of Sound Quality Evaluation is put into the above mentioned context in this paper.

1 - INTRODUCTION
For a long period of time Sound Engineering basically dealt with the reduction of the overall sound level that is emitted by a product. But, within the last decade the focus started to switch more and more towards the aspect of the quality of the resulting sound. This development of Sound Engineering results in the fact that sound engineers have to cope with completely different tasks and methods − the requirements for this profession have been significantly extended. In contrast to traditional Sound Engineering which is restricted to the investigation of pure physical and mechanical dimensions, Sound Quality Engineering also has to consider human perception. Thus besides the traditional mechanical and physical knowledge Sound Quality Engineers also have to acquire knowledge in psychoacoustics and even in psychology.
A basic problem resulting form this change is that completely different measurement procedures are necessary. While physical signals like the overall sound pressure level can directly be measured with an instrument and following a method well defined in international standards, now human perception has to be measured. From the view of the traditional engineering education it might even be stated that such a "measurement" is impossible, because no instrument can directly measure this perception. But, instead of an instrument here different measurements methods have to be applied, methods which are based on perceptual test with subjects. The development of these tests have a long tradition in the field of psychoacoustics, which offers the basic solution for the problem of Sound Quality evaluation: physical signal parameters are related to aspects of human perception. These methods can thus be used to build the bridge between parameters which can be measured with traditional instruments and human perception. But, the methods have to be extended in order to cope for non-acoustical and even non-sensory moderators, so that they can not be standardized as traditional sound engineering methods − knowledge in human perception is required.

2 - MODERATING FACTORS FOR SOUND QUALITY
In contrast to other quality measures which can be defined by pure physical quantities, Sound Quality is based on human perception. Human perception itself is not only based on the acoustical signal which is received by the two ears of listener, is also depends on other sensorial modalities like visual, tactile or haptic information. Furthermore and even more complicated, also non-sensorial aspects have an influence on the judgment of Sound Quality − cognition controls our perception.
The cognitive influences can be divided into three groups:

- Source (product) -related: a source/product usually represents an image;
- situation-related: a product is used in a specific activity situation, the user can interact with the source;
- person-related: people have their personal expectation, motivation, taste, preference or aversion.

Sound Quality thus is a multidimensional consisting of three different factor groups:

- physical factors (the acoustical signal);
- psychoacoustical factors (describing acoustical sensorial aspects, e.g., loudness, sharpness, fluctuation strength, see e.g., [1]);
- psychological factors.

An important point is that humans only use three to four of these factors to create their judgment (see, e.g., [2]). The selection of the respective factors is driven by cognition. As a consequence, the same physical sound can result in completely different Sound Qualities. Sound Quality is product specific, which means that each product (or class of products) has its own specific requirements for Sound Quality. It is the first step of Sound-Quality Evaluation to identify these product-specific requirements. Sound Quality evaluation thus is a complex task, and that it requires multidisciplinary knowledge. The appropriate methods has to be selected based on the specific product and task.

3 - PROCEDURES OF SOUND-QUALITY EVALUATION

In each type of measurement all factors which have an influence on the quantity to be measured have to be controlled. This is also true for measurements of Sound Quality. Thus the first task in setting up an experiment is to identify the moderating factors for the specific product or sounds to be evaluated. This can be a tedious task, because in most cases it is not known in advance which factors do have an influence and which do not.

Once the factors are known, it can either be decided if they can be controlled in the experiment − or, if this is not possible, if they at least can be kept constant during the experiment and for all subjects. The methods to evaluate Sound Quality can thus not only restrict themselves to the pure acoustical signal, they also have to consider other modalities and the specific situation and background of the subjects. Although they are based on traditional psychoacoustics, these basic methods have to be extended to cope with the requirements.

Usually Sound Quality evaluation test are performed in a laboratory. It is obvious that the moderating factors in a such a laboratory situation can significantly differ from those which are present in the normal life situation where the product is handled by a user. This context information is better considered by field tests, but this type of test shows some drawbacks compared to laboratory tests. Advantages of laboratory tests are:

- the test is reproducible;
- all subjects have identical test conditions;
- if products are compared, they can be evaluated in identical states of operation;
- different sounds can directly be compared;
- stimuli can adaptively be modified depending on the subjects answer, e.g., to efficiently identify target sounds;
- the test is time-efficient.

In contrast a field test shows the following advantages:

- it is a representative situation for the usage of a product in daily life;
- a typical handling of the product is possible;
- interaction with the product is possible;
- subjects can individually select typical or critical states of operation.

If Sound Quality should be evaluated with regard to customer relevance, in general a field test is indispensable. But, especially due to the effort and time consumption such an investigation often is not possible or practicable.
If the experiments have to be conducted in the laboratory, they have to be carefully planned and in general it has to be checked if the results can be transferred to the field. Differences in judgments in the field and laboratory are usually due to the fact that subjects can derive different information in both cases, so that their cognition might select different factors to build their judgment.
Resulting from the discussion of moderating factors above in general the following aspects have to be considered for a laboratory experiment.
With regard to the physics sophisticated methods for aurally-adequate sound recording and playback are available. Using for example a dummy head for recording and equalized headphones for playback the acoustical signal at the eardrums of a listener can nearly perfectly be reproduced. But, since humans also perceive low frequencies by the whole body, a pure headphone reproduction does not lead to authentic perception. To avoid this sometimes subwoofers are used if sounds have strong low frequency components. The acoustical channel can thus normally be reproduced in a satisfactory manner. This is different for other modalities since corresponding reproduction methods are either still missing or very expensive. Optical information can be presented by images or videos, but true 3-dimensional reproduction is not applicable. Other modalities can only be presented as with strong simplifications or restrictions [3].
The most problematic factor group are the cognitive factors. In the laboratory a reduced amount of information is available for the subjects, and this specially concerns non-acoustical and the non-sensory information.
The *source-related factors* are not present in a pure acoustical experiment, so that they have to be made available by presenting additional information about the product, e.g., in form of a verbal description, pictures, videos, or models.
Situation-related factors are hard to reproduce in the laboratory. Here subjects usually are passive in listening to a sound, so that they are not included into the activity. Furthermore, interaction with the source usually is not possible. It is thus necessary to explain the situation carefully to subjects.
Person-related factors have a stronger influence the more the subject knows about the product and the situation, so that the remarks above have to be applied. It is important for the interpretation of the results to identify and record these factors, e.g., in form of a questionnaire.
As a consequence a general applicable and standardized method to evaluate Sound Quality does not exist. The specific aspects of the product, its application, and the target group have to be well considered in planning and running evaluation experiments.
An appropriate evaluation method consists of two blocks: a kernel procedure, usually implemented as one of the standard or modified psychoacoustic test methods, and a framework which contains the presentation an documentation of all non-acoustical information. A variety of different psychoacoustic test methods are available from literature (see, e.g., [4]), and the selection of the appropriate method depends on the character and number of stimuli and the required type of output. Most common methods are absolute and relative methods. An example of an absolute method are direct-magnitude estimation tests, in which subjects listen to a stimuli and directly quantify the feature to be evaluated. The most popular relative method is pair-comparison, in which two stimuli are presented as a pair, and the subject has to select the one which better fulfills a given criterion. Anyhow, both methods have their advantages and disadvantages. Especially for the application in the industrial environment the corresponding needs and restrictions have to be considered: methods have to be time-efficient, render results with a sufficient accuracy, and give direct clues on how to improve products.
An appropriate methods was presented in [5]. The so-called *individual test* combines the advantages of pair comparisons (direct comparison of the feature to be evaluated) and direct estimation (absolute judgment of the feature) but avoids their disadvantages (time consumption and difficulty for similar stimuli).
In this test the subject has access to all stimuli, and he can decide by himself how often and in which order he wants to listen to sounds. His task is to arrange the stimuli on a graphic board in such a manner that the feature to be evaluated is rated on a scale, e.g., from bad (bottom) to good (top). The result thus represents both, a ranking and an absolute judgment. The experiment is time-efficient since subjects can perform pair-comparisons only for those stimuli which are similar. A further advantage of the individual test is that the subject controls the experiment himself. He thus is actively involved in the experiment, which usually results in a higher motivation. Furthermore the subject has no longer the impression to be controlled by the test, so that his self-reliance increases and his stress is reduced.

4 - APPLICATION EXAMPLE: GEAR RATTLE

A typical evaluation task is the investigation of the effect of a sound component on Sound Quality. Such a component is gear rattle, which is a noise produced by the gear and which only becomes audible in specific driving conditions. Although it is of relative low level compared to the overall sound level of a car, it focuses the attention of subjects to it once it is detected. The phenomenon was investigated in [6] and [7], the focus here is more to put the evaluation method in the theoretical context discussed in the previous chapters.

The specific requirements for Sound-Quality in this case are not that the sound is not audible at all, but that it is not noticed by subjects or that it is below their acceptance-level, respectively. A typical driving condition in which gear rattle can be noticed is a stop-and-go situation, e.g., in a traffic jam. Here the driver subsequently engages and disengages the clutch, so that he gets a direct comparison between the situations with and without gear rattle at nearly identical background noises. This representative situation was selected to conduct the experiments.

In general, the threshold or acceptance level of such a sound component can easily be measured in the laboratory. But, since the gear rattle is only one component of a complex acoustical signal, it is not obvious that the results in the laboratory can be transferred to a situation where a customer is driving the car. We thus developed a test design which gives evidence about the difference between field and laboratory and between the situation where subjects do not know which sound component should be evaluated (non-sensitized) and where they are sensitized to gear rattle.

In the laboratory sounds of 7 different vehicles with a diesel engine where investigated, and two of them extended by a vehicle with gasoline engine were tested in the field. The vehicles for the field were of same type and color. Two groups of subjects participated in the test which was constructed of three different phases. One group started non-sensitized in the field, was then sensitized there and finally went into the laboratory, while the other group started non-sensitized in the laboratory, was then sensitized there and finally went into the field.

Fig. 1 shows the results of the ratings of the first phase, where the task of the subjects was to rate the annoyance of the engine noise.

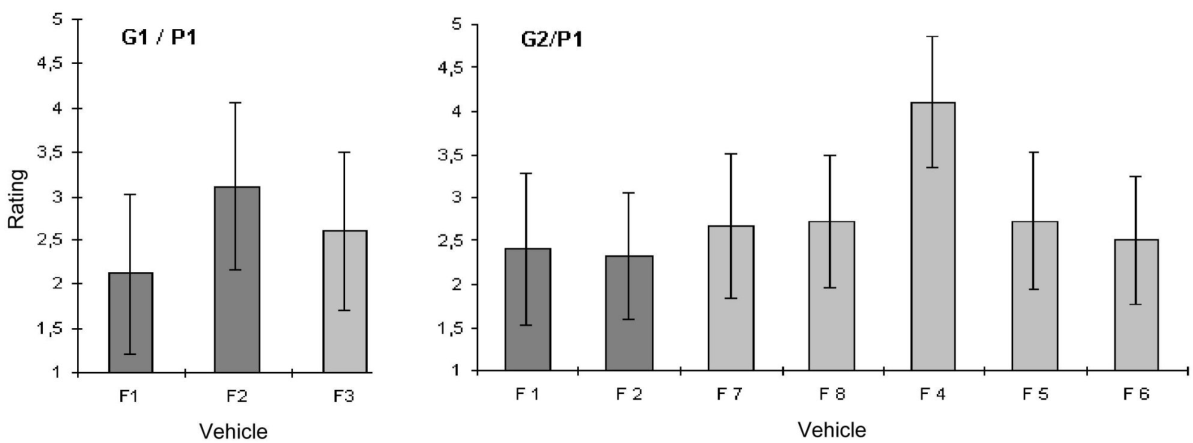

Figure 1: Rating in the field (left) and laboratory (right), non-sensitized subjects.

It can be seen that in the field vehicles 1 and 2 are rated differently, while they get the same rating in the laboratory. The annoyance of the noises is thus based on the gear rattle in the field condition, while it is based on other features in the laboratory condition. The explanation for this effect is the interaction in the field: the subject engages and disengages the clutch and gets the respective feedback in form of the gear rattle, so that he is automatically sensitized to this sound component.

Fig. 2 shows the results of the other two experiment phases when subjects were sensitized. Here the task of the subjects was to rate the strength of the gear rattle.

The differences between the sounds is now also rated in the laboratory. Since in that case also stimuli with strong synthetic gear rattle were presented, the difference between the rating of vehicles one and two is slightly lower than in the field.

5 - SUMMARY

Methods to evaluate Sound-Quality have to consider the product-specific requirements and the background of human perception. They thus have to go beyond the pure acoustical signal and have to

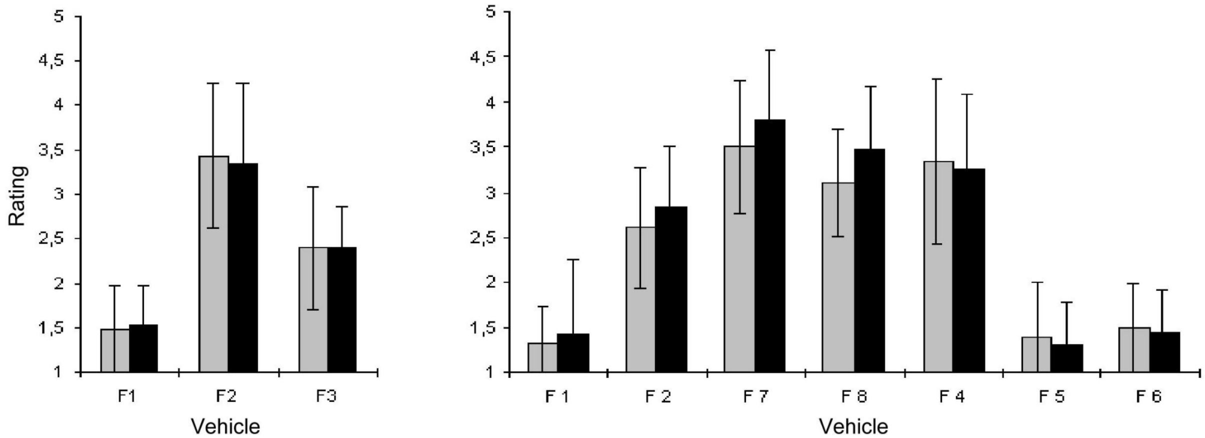

Figure 2: Rating in the field (left) and laboratory (right), sensitized subjects; light bars: subjects which first where in field; dark bars: subjects which first where in laboratory.

consider other sensory quantities and non-sensorial moderating factors. These moderating factors have to be identified first in each evaluation. The application example has shown that the rating of subjects can significantly depend on the non-acoustical and even non-sensory factors, so that they have to be considered and controlled or documented in each evaluation.

REFERENCES

1. **Fastl, H.**, The Psychoacoustics of Sound-Quality Evaluation, *Acustica - Acta Acustica*, Vol. 83 (5), pp. 754-764, 1997

2. **Blauert, J; Jekosch, U.**, Sound-Quality Evaluation - A Multi-Layered Problem.*Acustica - Acta Acustica*, Vol. 83 (5), pp. 747-753, 1997

3. **Blauert, J,; Lehnert, H,; Sahrhage, J.; Strauss, H.**, An Interactive Virtual-Environment Generator for Psychoacoustic Research.*Acustica united with Acta Acustica*, Vol. 86, pp. 94-102, 2000

4. **Green, D.M.; Swets, J.A.**, *Signal Detection Theory any Psychophysics*, Krieger, 1974

5. **Bodden, M.; Heinrichs, R.**, Evaluation of interior vehicle noise using an efficient psychoacoustic method, *Proc. of the Euronoise 98*, 1998

6. **Heinrichs, R.; Bodden, M.**, Perceptual and instrumental description of the gear-rattle phenomenon, *Proc. Int. Congress on Sound and Vibration, Copenhagen, Denmark*, 1999

7. **Bodden, M.; Heinrichs, R.**, Analysis of the time structure of gear rattle.*Proc. Of the Internoise 99, Fort Lauderdale, USA*, 1999