# Evaluation of Quality Features in Non-Standardized Rooms: Overcoming the problem of the missing reference using MSIPM

Ulrike Sloma[1]

[1] *Technische Universität Ilmenau, 98693 Ilmenau, Germany, Email: ulrike.sloma@tu-ilmenau.de*

## Abstract

An implementation of the multiple stimulus ideal profile method (MSIPM) for the evaluation of spatial audio signals is presented in this contribution. This new method in the field of audio evaluation, was only employed by a few studies since now. The research in this study focuses on the evaluation of quality features of spatial audio signals in non-standardized rooms. The playback is done through a loudspeaker system and a soundbar. The aim is to evaluate differences in perception according to the differences in room acoustic parameters. In previous studies assessing this question, no differences in quality perception were found. An explanation may be that there is no constant reference throughout the tested rooms. Each audio signal is modified by the special room acoustics. Furthermore, it is assumed that listeners adapt to the room acoustics. The idea behind the usage of MSIPM is that the listeners have an expectation of the sound of one spatial audio signal. The anticipated specification of a quality feature can be defined as its perceptual ideal point. The value may be consistent over different rooms and the evaluation of the audio signals is done relative to this ideal point value.

## Introduction

The presented research deals with audio quality evaluation considering the context of the use of technical audio reproduction systems. Mostly the evaluation of audio quality is done in specified listening rooms where the acoustic parameters, like size, external noise or reverberation, are stipulated through standards. These rooms are not like the listening room where the expected user listens, e.g a typical living room. For that reason quality evaluation limited to standardized rooms does not seem to be sufficient for quality evaluation. From own experience the assumption that the room acoustics of a listening room influence the perceived quality of spatial audio is nearby. The initial aim of the research was to find suitable attributes to describe spatial audio quality and how the room acoustic parameters influences these descriptions. This question was assessed in two mixed methods studies by the author [1]. The studies focused on the question if a difference between several listening rooms exist. A quantitative non-individual approach asking for basic audio quality and five generalized attributes was used. The second study was employed with an individual descriptive approach. Both studies used the same evaluation principle – one listener panel evaluates the same signals through a 5.1 loudspeaker setup in three acoustically different rooms.

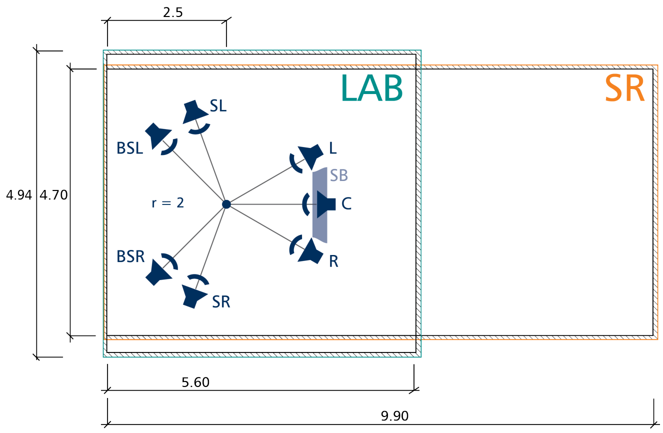



**Figure 1:** Scheme of listening test setup in the audio-video laboratory (LAB) and the seminar room (SR). Placement of loudspeaker system (LS)[L, R ,C, SL, SR, BSL, BSR], soundbar (SB) and listener (midpoint). The LS height was at 1.20 m. The SB was placed directly under the C LS. The listener was sitting on a chair. The distance back wall to midpoint was 2.50 m. The room dimensions are given in meters. Not shown are the subwoofers, placed in front of the listener.

In the previous studies no or just minor effects of the room acoustics could be observed in the evaluation data. The highest effects on quality depend on the played audio signal, and the reproduction system. One conclusion could be that room acoustics have no influence on the evaluation of quality and it does not matter whether a spatial audio system is tested in a laboratory or in a real listening environment. But the participants reported on heard differences between the rooms. It could be thought of several reasons why the test results are not conclusive. First reason might be that the listener adapts to the room acoustics of the listening room. The second reason is that in both tests no "real" reference was presented which is stable across the different rooms and therefore a direct comparison was not possible. [1]

To overcome the problem of the missing reference the idea to use a thought example or expected spatial quality as reference came up. These ideal profile method is a well known sensory profiling method in food and product industry [2]. The first documented transfer of this method to audio evaluation was done by Zacharov et al. [3] in 2016 as MSIPM - Multiple Stimulus Ideal Profile Method to assess next generation audio systems. Liebl et. al. [4] used this method to assess the quality of a large number of soundbars and compare them among each other.

For the presented study the basic idea to use this MSIPM method was to take the ideal point as representation of the internal reference, which may stay constant between the rooms or will also be adapted.

The standardization of the Multiple Stimulus Ideal Profile Method is in progress by the ITU within WP6C, currently named as Recommendation ITU-R BS.[NO-REF] ("Method for the subjective quality assessment of [highly differentiable] audio systems using multiple stimuli without a given reference"). However the study presented here refers to the procedures described in [3] and [4].

## Multiple Stimulus Ideal Profile Method

The method evaluates audio quality in two separated steps. After an evaluation of the basic audio quality (BAQ) selected attributes are evaluated sequentially. All evaluations are done in a multiple stimulus way, were mostly the systems or signals to be tested are rated parallel with possible direct comparison. Combining the results of the two steps allows a deeper analysis and interpretation of the listeners perception. [3]

The test procedure is structured as follows:

1. Assessor instruction
2. Familiarisation and rating of the basic audio quality (BAQ)
3. Familiarisation and rating of quality attributes and assessment of the ideal characteristics of the attributes (ideal point)

## Listening test design

The test was conducted in two listening rooms, with same reproduction systems and audio signals. Figure 1 shows the rooms and the positioning of the listener and the reproductions systems.

### Rooms:

Two acoustically different rooms, located at TU Ilmenau, were chosen as listening rooms. A relative dry audio-video laboratory (LAB) and a more reverberant seminar room (SR). The parameters in table 1 are measured with and averaged over the used loudspeakers with the measuring microphone MM210 and MK223 capsule from Microtech Gefell at the listener position.

**Table 1:** Room parameters of the both listening test rooms.

| Parameter | LAB | SR |
|---|---|---|
| $V$ | $77.38\,m^3$ | $144.24\,m^3$ |
| $T_{30}$ | $0.25\,s$ | $1.17\,s$ |
| $EDT$ | 0.09 | $0.87\,s$ |

### Playback systems:

The first reproduction system was a standardized loudspeaker (LS) setup placed according to ITU [5]. Used are Genelec 1030A loudspeakers with a Genelec 7050B subwoofer. The second system was a Yamaha YSP-2500 soundbar with subwoofer (transition frequency 500 Hz) (SB) which was calibrated for each room. The used soundbar produces its sound field using the reflections from the walls of the reproduction room.

### Attributes:

Two attributes were selected by a panel discussion. The number is limited to reduce the length of the listening test. The scale ranged from 0 to 100 with two sided run-offs of 10. This scale is also applied to the BAQ rating.

- **Perceived scene extension:** Perceived extension of the scene in all directions.
- **Clarity:** Impression of how clearly different elements in a scene can be distinguished from each other, how well can various properties of individual scene elements be detected.

### Listeners:

The test was done in two separate sessions at different days. The quality in both rooms was evaluated by the same listener panel in the same order. First LAB and second the SR. 13 listeners participated, four of them were female and nine male. The average age is 33 years. All listeners were experienced with listening tests but not expert listeners. Each listening test took about 45 minutes in each room.

## Stimuli

### Signals:

As audio signals a complex effect scene (fantasy) and voices distributed throughout the room at up to three discrete positions (voices) were chosen. Three versions of the signals were mixed for different numbers of reproduction channels. A Stereo mix, a 5.1 and a 7.1 mix. These items are played through both playback systems.

### Trials:

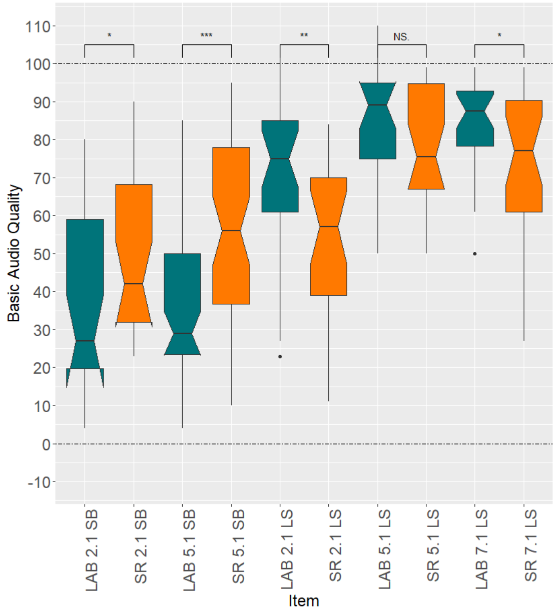
To avoid that the number of channels will have a overlaying effect the items were presented in groups shown in table 2. These triples are evaluated in randomized order and with one repetition. The 5.1 mix played over SB is included in both trials, to see if it the rating is influenced by the more or less spatial items.

**Table 2:** Overview grouped triples – items in one trial are presented within one multiple stimulus interface (BAQ and attribute evaluation).

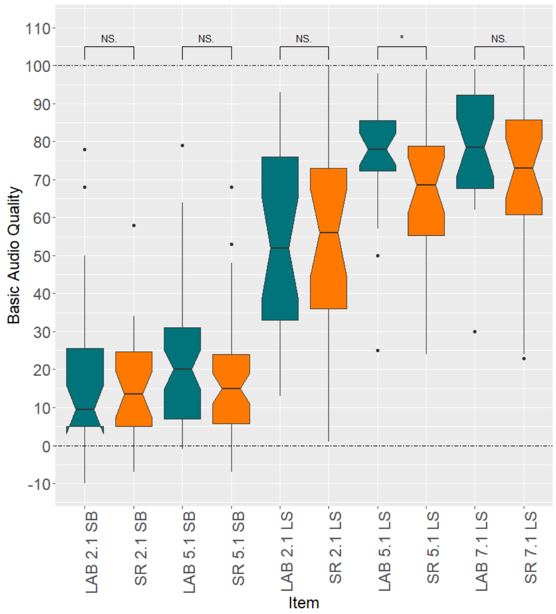
| | | | |
|---|---|---|---|
| **StereoTrial** | LS 2.1 | SB 2.1 | SB 5.1 |
| **SpatialTrial** | LS 7.1 | LS 5.1 | SB 5.1 |

## Results

The main results are displayed in the figures 2 and 3. The test for normal distribution of the data was done with the Shapiro-Wilk test. Since most of the data was not normal distributed the ongoing analysis was done with non-parametric statistics. The Wilcoxon rank-sum test was performed to display significant differences in the ratings between the two rooms.




(a) Signal - fantasy scene




(b) Signal - voices

**Figure 2:** Results for the **basic audio quality** (BAQ) ratings for all systems and rooms, 5.1 SB rating is summed up over both trials – $^{***}$ $p \leq 0.001$,$^{**}$ $p \leq 0.01$,$^{*}$ $p \leq 0.05$ (Wilcoxon rank-sum-test)

## Basic Audio Quality

Figure 2 shows the resulting box plots for the basic audio quality for all systems and versions evaluated in the test for both signals.

### Trial:

The item 5.1 mix played over SB for both signals was included in both trials. No difference in the evaluation between the StereoTrial and SpatialTrial could be found inside one room. Therefore the SB 5.1 ratings for both trials were evaluated together.

### Signal:

The ratings of the BAQ depend on the reproduced signal content. Looking at the signal voices, only one significant difference, 5.1 mix over LS, can be found between the rooms. For the complex fantasy scene differences in the evaluation can be found for all systems and versions, except the 5.1 mix over LS. The ratings for the fantasy scene have a high variance.

### System:

The found differences in the signals further depend on the reproduction system. The LS setup was rated higher in BAQ in the LAB than in the SR. In contrast to that the SB was rated higher in the SR than in the LAB. Further the ratings for the signal voices and the system SB have a lower variance than the other combinations.

## Attribute rating

The overall results regarding the dependency on trial, signal and system are similar to those of the BAQ evaluation. The resulting box plots are shown in figure 3. For the signal voices no significant differences are observed. The signal fantasy scene has some differing evaluation. The 5.1 mix played over SB was summed up for the attribute evaluation, too.

### Ideal point

The ideal point was summed up over both trials, since there was no significant difference between them. The results for the Ideal Point do not show differences between the rooms.

### Perceived scene extension:

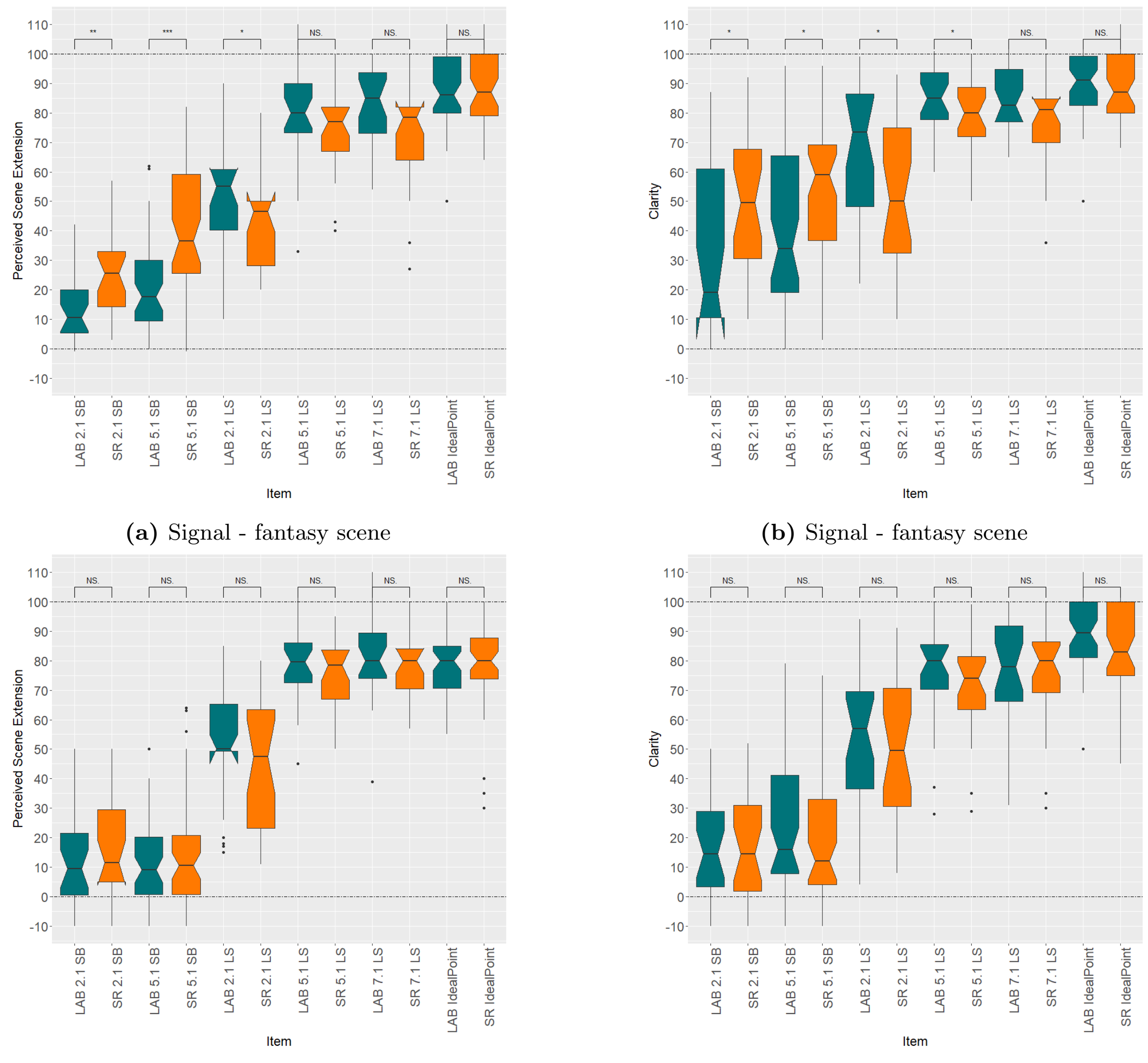
A moderate variance over all ratings can be seen. The ideal point for this attribute is higher for the signal fantasy scene than for the signal voices.

### Clarity:

In the StereoTrial a very high variance in the ratings can be observed, especially for the signal fantasy scene. The SpatialTrial shows a lower, moderate, variance. The ideal point is similar for both signals.

## Discussion and Conclusion

The results of the listening test show a dependency of room acoustical parameters on the evaluation of the perceived audio quality. This can be seen in the basic audio quality rating and in the attribute rating. The evaluation between the systems seem to be more distinct in a slightly reverberant environment (LAB) than in a high reverberant environment (SR). As in the previous studies [1] a high signal and system dependency compared to the room acoustical influence can be seen in the derived data. Regarding the different variances in the ratings it can be said that the listeners are rather in agreement with the characteristics of the term perceived scene extension than with the term clarity.




**(a)** Signal - fantasy scene

**(b)** Signal - fantasy scene

**(c)** Signal - voices

**(d)** Signal - voices

**Figure 3:** Results for the ratings of the attributes **perceived scene extension** (3a, 3c) and **clarity** (3b, 3d) for all systems and rooms, 5.1 SB and the ideal point ratings are summed up over both trials – $^{***}\ p \leq 0.001,^{**}\ p \leq 0.01,^{*}\ p \leq 0.05$ (Wilcoxon rank-sum-test)

The assumption of a constant ideal point throughout the rooms was confirmed with the listening test results. It is not clear if the rating of the ideal point supported the resulting observable quality differences between the rooms. Caused by the calibration process the reproduction through the playback system SB is not stable between the rooms. Therefore the two changing variables, room and SB, make a comparison difficult.
Further it should be discussed if the MSIPM method is suitable for the proposed listening test design. To apply the method in a correct way the rooms should have been the items to be compared in the multiple stimulus interface.

## Acknowledgement

The author likes to thank all the voluntary participants who were taking part in the listening test.

## References


[1] Sloma, U.: Evaluation of Quality Features of Spatial Audio Signals in Non-Standardized Rooms: Two Mixed Method Studies, 140th AES Convention, Paris - France, 2016.

[2] Worch, Thierry, et al.: Ideal profile method (IPM): The ins and outs., Food Quality and Preference 28.1 (2013): 45-59.

[3] Zacharov, Nick, et al.: Next generation audio system assessment using the multiple stimulus ideal profile method., 8th Int. Conf. on Quality of Multimedia Experience (QoMEX). IEEE, 2016.

[4] Liebl, T, et al.: Audio quality evaluation of soundbars using the multiple stimulus ideal profile method, Proc. of 44nd annual convention for acoustics - DAGA, Munich - Germany, 2018.

[5] Rec. ITU-R BS.775-3: Multichannel stereophonic sound system with accompanying picture, 2012.