

Report

SenseLabOnline listening test on sound insulation of walls - A feasibility study

Performed for COST Action TU0901 WG 2

SenseLab 012/12

Project no.: FR18303-04

Page 1 of 33

6 appendices

09 March 2012

DELTA

Venlighedsvej 4

2970 Hørsholm

Denmark

Tel. +45 72 19 40 00

Fax +45 72 19 40 01

www.delta.dk

VAT No. 12275110

Title

SenseLabOnline listening test on sound insulation of walls - A feasibility study

Journal no.

SenseLab 012/12

Project no.

FR18303-04

Our ref.

THP/SA

Client

COST Action TU0901 WG 2

Summary

This report describes a SenseLab listening test made on rating the annoyance potential of neighbour activities heard through walls. The main purpose was to test the methodology, but interesting results were found as well.

The SenseLabOnline Annoyance potential test methodology was employed with 22 assessors from 11 countries and the sounds were rated using the ISO 15 666 annoyance scale. The test was performed online by the participants with their own headphones in their own premises.

Six simulated walls (systems) were tested with four samples (sounds). The assessors completed 2 repetitions of the test.

Assumptions for the analysis of variance (ANOVA) were fulfilled. The ANOVA model showed high R^2 -values (91 %) and the largest effects came from the assessors, systems (simulated walls) and sound samples under test.

As an approximate level calibration procedure the participants were asked to adjust a speech sample to a natural level. A similar test with another group of persons gave a mean value of 64 dB(A) with a standard deviation of 4.5 dB for this methodology.

After this “calibration”, dose-response curves on the receiving side of the wall could be found with good accuracy ($R^2 = 0.95$).

The sound insulation of the simulated walls was rated significantly differently with respect to annoyance, potential except a single concrete wall and a double gypsum wall. A good correlation ($R^2 = 0.98$) was found between the annoyance potential and the R'_w -values.

DELTA, 09 March 2012

T. Holm Pedersen

Torben Holm Pedersen

Sonia Antunes

Sonia Antunes



Contents

1. Purpose.....	4
2. SenseLabOnline workflow overview	4
3. Equipment and software.....	4
3.1 Preparation of the sound files	5
3.2 Environment and play-back equipment	6
3.3 Level calibration and play-back levels	6
4. Test panel & instruction	7
5. Stimuli and test scheme	7
5.1 Methodology.....	7
5.2 Test scheme and stimuli.....	7
6. Data analysis.....	10
6.1 ANOVA.....	11
6.2 Overall results	15
6.3 Comparison with physical metrics.....	16
6.4 Noise sensitivity as moderator for annoyance potential	20
7. Conclusion.....	22
References	23
Appendix A - SenseLabOnline annoyance test	24
Appendix B - Assessors	26
Appendix C - Detailed instruction	27
Level adjustment with reference speech	27
Instructions for the listening test	27
Appendix D - Assessor performance and scale usage	28
Appendix E - ANOVA assumptions.....	30
Appendix F - Calculation of dose-response curves.....	32

1. Purpose

As a part of the COST Action TU0901 WG2¹ activities, listening test for building acoustics should be performed. The main purpose of the present study is to test an online methodology in order to perform listening tests for rating airborne sound insulation, relating to neighbour noise between residences at the same floor.

2. SenseLabOnline workflow overview

For a start a pilot test trial of the methodology was made with 3 different sounds (bath, flush, a tap sound) and 3 different systems (no attenuation, 5 dB attenuation and 10 dB attenuation). After analysing the tests results a more representative selection of 4 test sounds (samples) and 6 wall types (systems) were made.

The sound insulation of the walls was simulated by equalizing the 4 sound samples in order implement the frequency dependent attenuation curves for the sound insulation of the 6 types of walls selected.

The final 24 sound files representing combinations of sounds and walls were uploaded to SenseLabOnline which arranged the files in a random order for each assessor.

People from the COST TU0901 project were invited by e-mail to participate in an online listening test. 22 persons from 11 countries completed the test which was performed between 13 February and 15 February 2012.

A SenseLabOnline standard statistical analysis was performed including analysis of the assessor performance.

3. Equipment and software

The equipment used for the listening tests at DELTA SenseLab is listed in Table 1.

¹ COST Action TU0901 “Integrating and Harmonizing Sound Insulation Aspects in Sustainable Urban Housing Constructions” was established in 2009 and has Three Working groups:

WG1: Harmonized sound insulation descriptors and classification schemes in Europe

WG2: Subjective evaluation of sound insulation - Laboratory tests and harmonized field surveys

WG3: Design and acoustic performance of building constructions for multi-storey housing

Information about COST Action TU0901 and the WGs is found at websites, see [1].

Apparatus	Make	Model	AV no.	Remarks
PC	Lenovo	ThinkCentre 7103-A4G	–	Placed in control room
Sound Cards	Firestone	Audio Fubar III	–	Placed in control room
Headphones	Sennheiser	HD 650	1384L-1405L	Checked January 2012
Preamplifier	Brüel & Kjær	ZG 0350	1376L	Calibrated 13-07-2011
Acoustic calibrator	Brüel & Kjær	4230	567L	Calibrated 30-08-2011
Head and Torso Simulator	Brüel & Kjær	4100	–	Calibrated 16-09-2010
Measuring amplifier	Brüel & Kjær	2610	454T	Calibrated 17-11-2019
Sound Level meter	Brüel & Kjær	2236		Calibrated 22-03-2011

Table 1

Equipment used for the listening tests at DELTA. The play-back equipment for the non-DELTA participants is unknown.

The DELTA SenseLabOnline software version 1.5 was used for the listening test.

3.1 Preparation of the sound files

Both calibrated sound files and sound files from a sound effects library were selected for the test. The sound files represented sounds with different frequency content and different levels.

The Adobe Audition software was used for the sound editing. The sound files were first converted into 32 bit mono sound files. These files were level aligned with respect to the A-weighted levels as shown in Table 2.

Then the sound insulation of the walls was simulated by equalizing the sound samples in order to implement the frequency dependent attenuation curves for the sound insulation of the walls.

Finally the samples were converted to stereo and reverberation corresponding to a living room was added before conversion to a 16 bit stereo format.

The final 24 sound files representing combinations of sound samples and walls were uploaded to SenseLabOnline which arranged the files in a random order for each assessor.



3.2 Environment and play-back equipment

An online SenseLab test was performed. Each assessor was instructed to make the test in a silent room in their office or at home. The equipment needed was a computer with sound card, an internet connection and a good pair of headphones. The type of headphones used is not specified. It should be noted that especially for non-open headphones the bass reproduction may vary considerably.

3.3 Level calibration and play-back levels

Each assessor started the listening test by adjusting the play-back volume of an audio reference file with male speech, so the voice had a natural volume of a man talking 1 m from you.

In order to find out which average volume to expect, 24 persons were asked to make the same adjustment procedure and at the same time the correspondent level was registered. This experiment involved 6 women and 18 men, aged between 26 and 62 years, (DELTA employees - approximately half of them acousticians) and for this experiment a pair of Sennheiser HD 555 headphones was used. The level setting was made with a calibrated attenuator and after each trial the setting was noted.

In Table 2 the mean levels (measured with the headphones placed on a calibrated artificial head, B&K 4100 Head And Torso Simulator), standard deviation and 95 % confidence intervals are shown.

	L_{Aeq}, dB
Mean	63.9
Stand. Dev.	4.5
CI 95 %	1.8
Maximal difference	15.0

Table 2

Mean values, standard deviation and 95 % confidence intervals for the level adjustment of male speech at 1 m made by 22 persons.

With a sound level meter the voice of a male talking in a 1 meter distance was measured to $L_{Aeq} = 60$ dB. As seen from Table 2 the reference audio file was in average adjusted to 64 dB (i.e. 4 dB higher than the natural sound level).

4. Test panel & instruction

34 persons of the three working groups of the COST TU0901 action were invited to participate. Of these 22 persons from 11 different countries completed the online test (see Appendix B).

In order to simulate the context of the occurrence of neighbour noise, the participants were instructed as follows:

“Close your eyes and concentrate on imagining that you are sitting and relaxing at home and hear the sounds from your neighbours. Imagine that the sounds will appear approximately every 10 minutes with the same duration as in this test”.

The participants were informed that it was a blind test in the sense that they would not get further knowledge about the test signals until after the test. The duration of the listening test was estimated to approximately 30 minutes.

5. Stimuli and test scheme

5.1 Methodology

The overall attribute evaluated in the SenseLabOnline test was “Annoyance”. The user interface for the assessors is shown in Appendix A.

When Annoyance is measured in an experimental context, it is called annoyance potential, but in this report the full name is not used consequently.

Stimuli were rated on a 0-10 continuous annoyance scale. The scale is divided into five equal intervals labeled: ‘Not at all’, ‘Slightly’, ‘Moderately’, ‘Very’, and ‘Extremely’ according to ISO 15 666.

For each assessor a randomized presentation order was used within each test block and each repetition block, yielding a double blind paradigm.

5.2 Test scheme and stimuli

After editing of the sound stimuli (see clause 3.1) the test was created online at: www.senselabonline.com. This included uploading of samples and preview of the stimuli for the test.

The test scheme was to evaluate 6 different systems for 4 types of sounds: music, people talking (voices), party sounds (people talking, laughing and music) and a toilet flush. The music had bass and heavy drums. The bass drum had the main components at 65 and 130 Hz, and at 50 Hz the level had dropped 6 dB with a steep slope down to lower frequencies. This means that the energy in the frequency bands below 50 Hz is inferior.



The natural levels of the sound samples on sending side were for a start adjusted the levels indicated in the table below (Table 3).

	Natural level		Estimated test level	
	L_{Aeq} , dB	$L_{A\ max,F}$, dB	L_{Aeq} , dB	$L_{A\ max,F}$, dB
Flush	69	79	83	93
Music	85	91	99	105
Party	80	86	94	100
Voices	65	74	79	88
Reference speech	60	69	64	73

Table 3

The natural A-weighted sound pressure levels and the average levels presumably used in the test (estimated test levels) of the sound samples on the sending side of the walls.

Taking into account the results obtained in the pilot test, namely the low volume of the sounds files, after the attenuation, two alternative strategies were considered:

1. Play-back at natural levels:
Realistic assessments of annoyance potential may be obtained, but some sound samples would be inaudible on the receiving side of the walls.
2. Play-back at increased levels:
Unrealistic high annoyance potentials but probably a good relative discrimination among the six walls.

The last option was chosen with an intended level increase of 10 dB.

As the mean level of the reference speech probably is adjusted 4 dB higher than assumed this means that the levels of the sound samples was in average 14 dB higher than the natural level.

These sounds have been processed in order to simulate their transmission through 6 different walls, see Table 4.

System no.	Name/code	Details	R'_w , dB
1	Single concrete Code: S_concrete	200 mm concrete 2400 kg/m ³	56
2	Single lightweight concrete Code: S_light_concrete	260 mm lightweight concrete 1400 kg/m ³	50
3	Single brick Code: S_brick	115 mm brick 1200 kg/m ³ , render 2 x 10 mm	41
4	Single gypsum Code: D_light_gypsum	2 x 1 layer of gypsum board, single frame, 45 mm mineral wool	40
5	Double gypsum Code: D_3x_gypsum	2 x 3 layers of gypsum board, double frame, 190 mm mineral wool	57
6	Double concrete Code: D_heavy	2 x 80 mm concrete, 2400 kg/m ³ 60 mm space, 50 mm mineral wool	63

Table 4

The walls simulated in the test and their weighted apparent sound reduction index (R'_w).

In Figure 1 the apparent sound reduction index for the 6 walls (systems) are presented as function of the frequency.

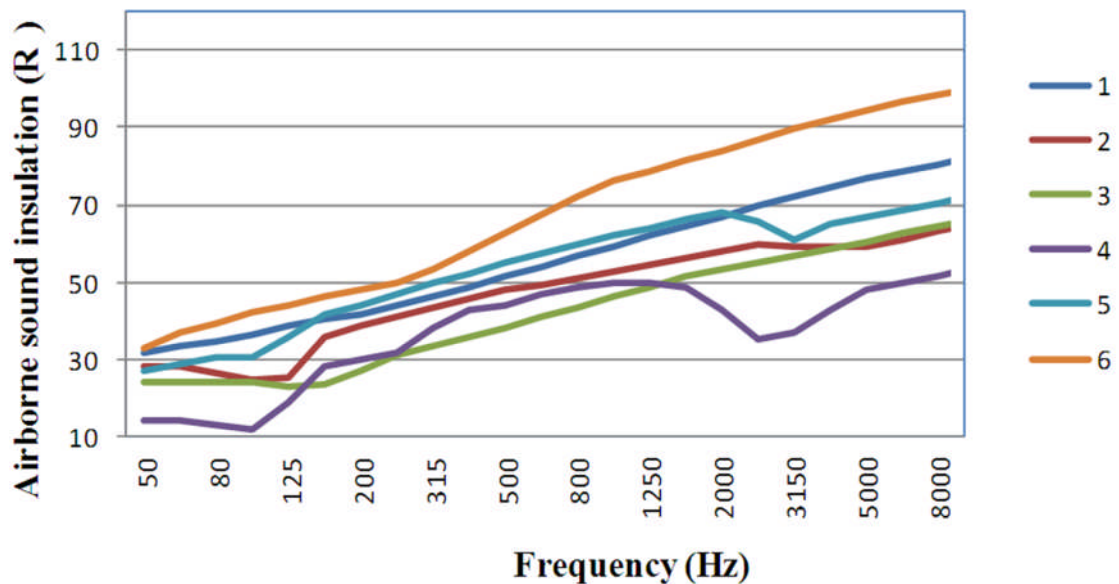


Figure 1
Apparent sound reduction index for the 6 walls (systems).

For system 1 and 5 the sound insulation field data was taken from reference [11]. The data for other solutions came from Bastian database (R_w -values). In order to take account for flanking sound transmission, 4 dB were subtracted in the attenuation curves of the systems 2, 3, 4, and for the double heavy wall (system 6) 8 dB were subtracted. In the frequency range between 5000 Hz to 20 000Hz we simulated the sound insulation curves using the mass law, for all systems. This means an increasing attenuation by 6 dB/octave.

For system 1 (single concrete) the sound insulation data started at 100 Hz, so below this frequency (until 50 Hz) we also extended the frequency range by using the mass law. Below 50 Hz we didn't make any further processing of the sound files.

6. Data analysis

The analysis was performed in accordance with the ITU-R BS.1534-1 recommendation. The scores are assumed to be quantitative, and the standard statistical analyses are conducted. Results are reported as mean values and a confidence interval. A Tukey Honestly Significant Difference is also available to sustain the results.

Before the data was processed for the results of the test the assessor performance (scale usage), agreement and consistence (of repetitions) were inspected (see Appendix D - Assessor performance and scale usage).

6.1 ANOVA

The results in this section show the effect of the variables (assessors, systems and samples) and the interactions between these variables. In Figure 2 the P-values and the F-values are shown. The P-values give the significance of the effects and the F-value give the size of the effects. Overall the dataset fulfilled the assumptions for the ANOVA (see Appendix E - ANOVA assumptions).

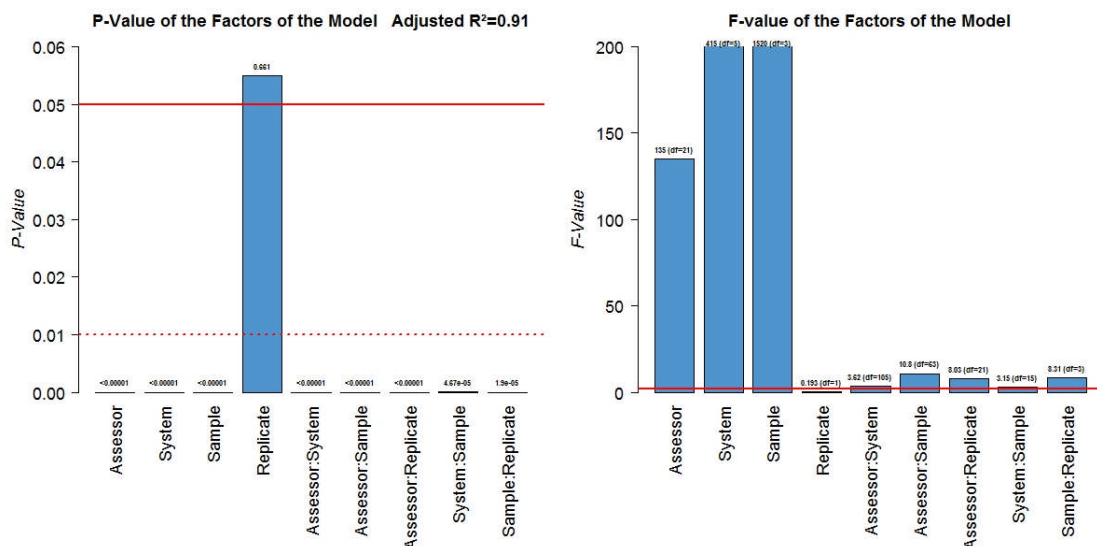


Figure 2

The results of the Analysis Of Variance (ANOVA).

In the left graph the P-values on a 95 % level are shown; the 95 % confidence level is represented with a red full line and the 99 % confidence level with a dotted line. In the right graph the F-values are shown. In both graphs columns that are larger than the length of the y-axis are cut at the top (with the actual numeric value given in the column).

The P-values indicate whether a variable has a significant effect. The P-value gives the probability of being wrong in saying that a variable has an effect. A P-value of 0.05 means that there is 5 % chance of being wrong in anticipating that a certain variable affects the results. The smaller P-values the more significant the effect is.

The F-values indicate the level of power for each of the variables. In the F-value graph a red line is drawn at the level 2. The larger F-values the larger the influence of the variable is. Variables with F-values less than 2 do not have much effect on the results.

Overall we see significant effects from the variables Assessors, Systems and Samples. The only insignificant variable is the replication, meaning that the assessor generally can repli-

cate repeat their assessments. The most powerful variables are the System (walls) and Samples (sounds) followed by the Assessor effect.

Also the interactions Assessor*System, Assessor *Sample, Assessor *Replicate, System*Sample, Sample *Replicate are significant (P-values <0.05).

Figure 3 shows the assessor responses for each of the systems (walls) averaged over the 4 sound samples. It is seen that the assessors 5014, 5028 and 5032 give low ratings that do not discriminate between the systems.

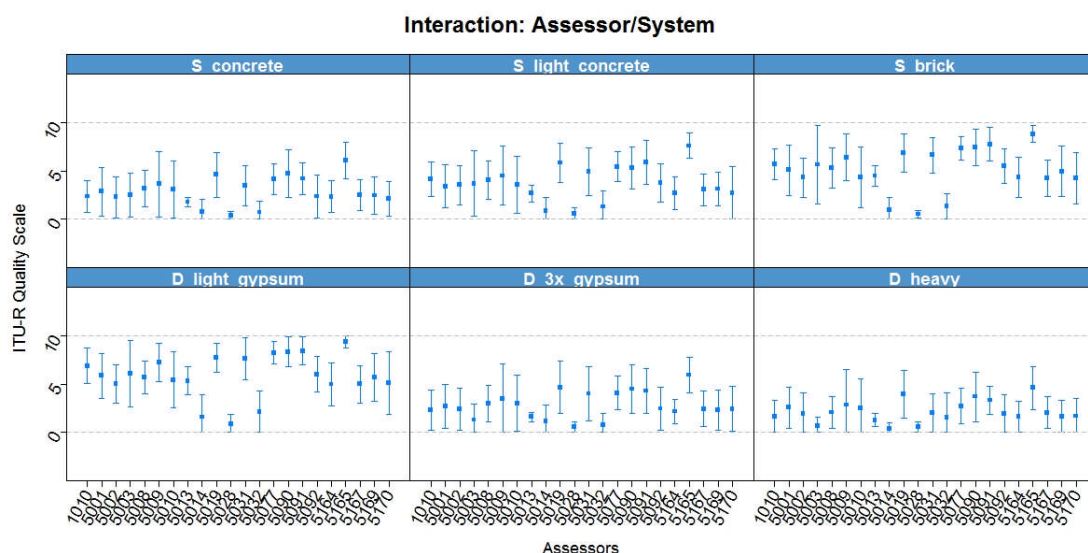


Figure 3

*Representation of the interaction Assessor*System. The system names and codes are described in Table 4.*

Figure 4 shows the assessor responses for each of the samples (sounds) averaged over the 6 walls (systems). It is seen that the assessors 5014, 5028 and 5032 give low ratings that do not discriminate between most of the sound samples.

Figure 3 shows the assessor responses for each of the replicates averaged over the 4 samples (sounds) and the 6 walls (systems). The left and right graph shows more or less the same picture meaning that the replicated do not differ much.

Figure 6 shows the responses for each of the walls (systems) for each of the samples (sounds) averaged over all assessors. The ranking of the annoyance from the four samples seem to be independent of the type of wall.

Figure 7 shows the responses for each of the replicates for the 4 samples averaged over the 6 systems (walls) and all assessors. The left and right graph shows more or less the same picture meaning that the replicated do not differ much.

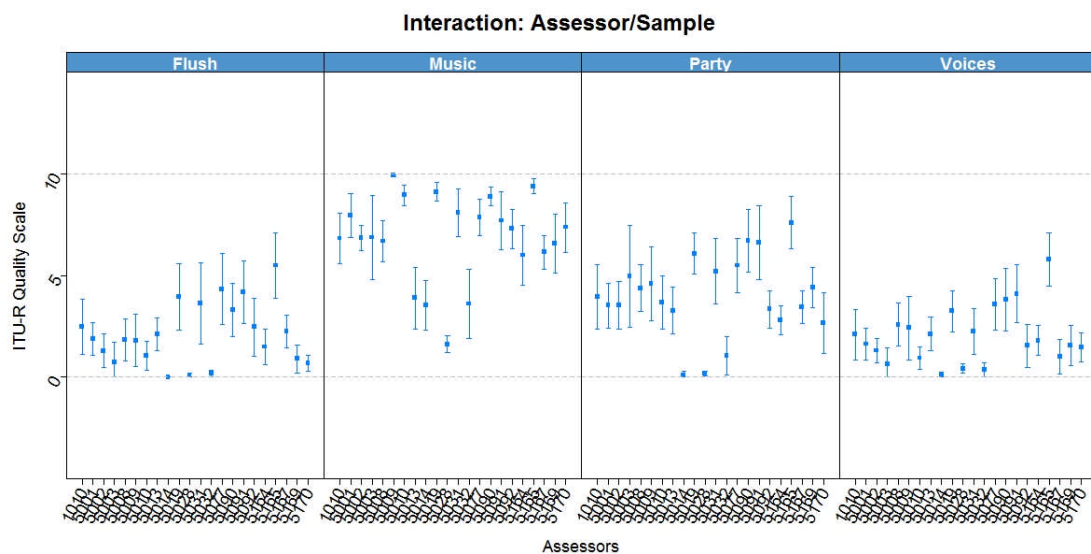


Figure 4
*Representation of the interaction Assessor*Sample.*

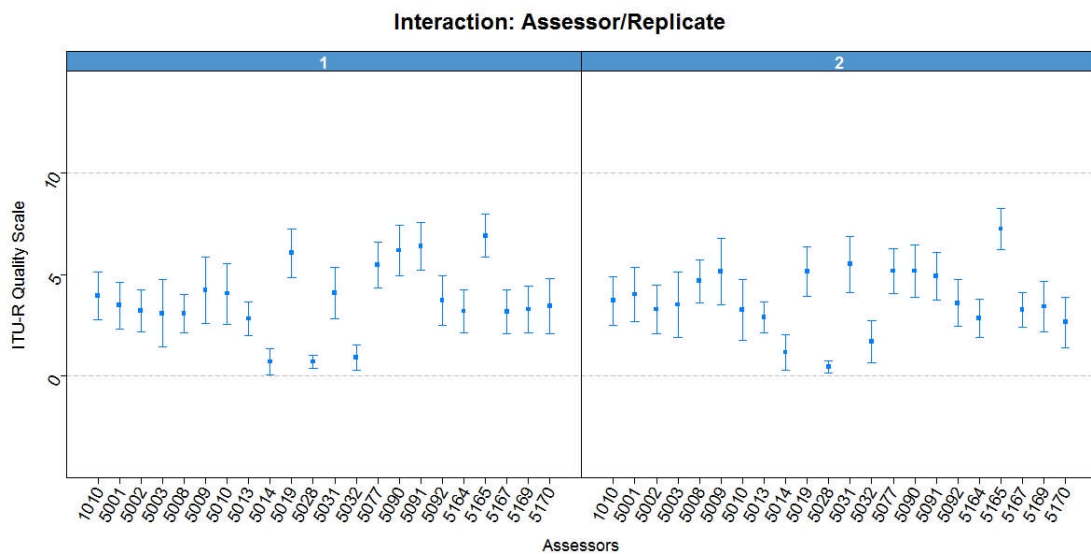


Figure 5
*Representation of the interaction Assessor*Replicate.*

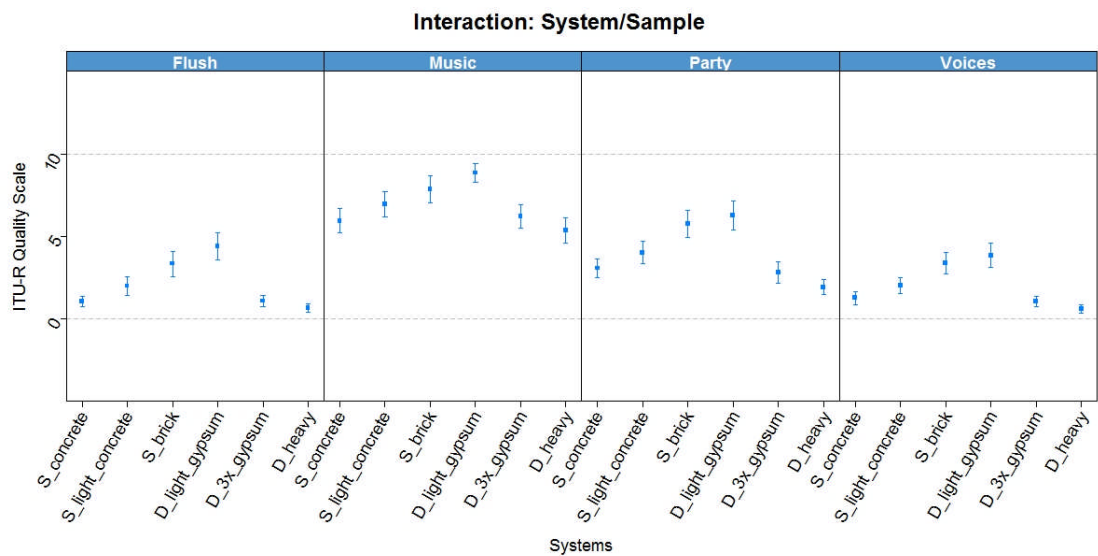


Figure 6
Representation of the interaction System*Sample. For at least one system the ratings given for one sample is different from another (average across all assessors). The system names and codes are described in Table 4.

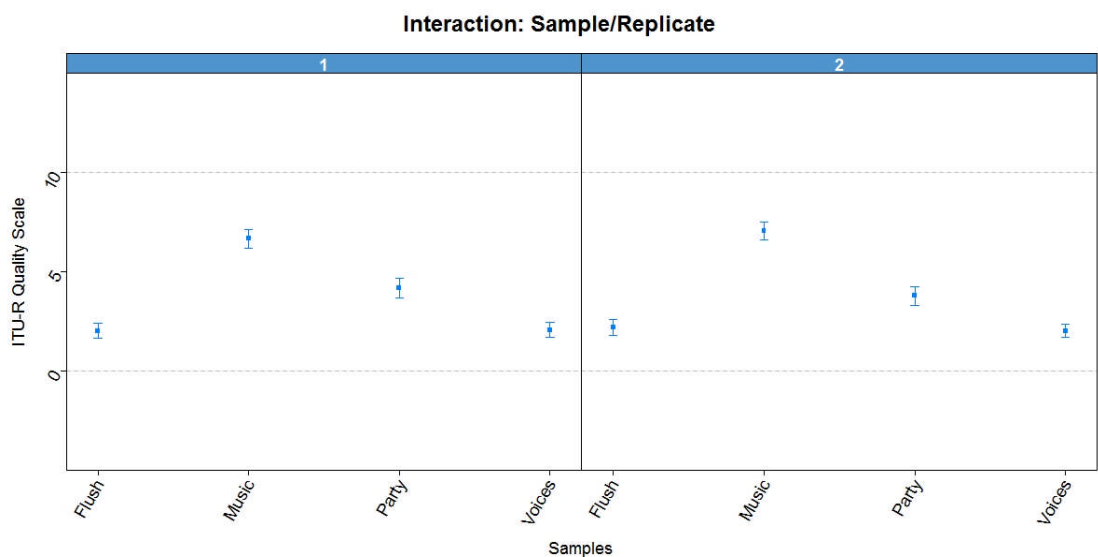


Figure 7
Representation of the interaction System*Replicate.

6.2 Overall results

This chapter focuses mainly on the overall scores and significance tests between systems under test. In Figure 8 the mean ratings of the systems (simulated walls) under testing are presented. It can be seen that the single light gypsum wall (D_light_gypsum) was rated having the worst performance, whereas the double heavy wall (D_heavy) has the best. Due to the large level differences of the 4 samples (sounds) the confidence intervals are rather large. Overlapping confidence intervals indicate that the differences may not be significant.

Table 5 shows the results of a Tukey Honestly Significant Difference test. The results show that sound insulation of all the simulated walls was rated significantly differently with respect to annoyance, except for system 1 (single concrete) and system 5 (double gypsum).

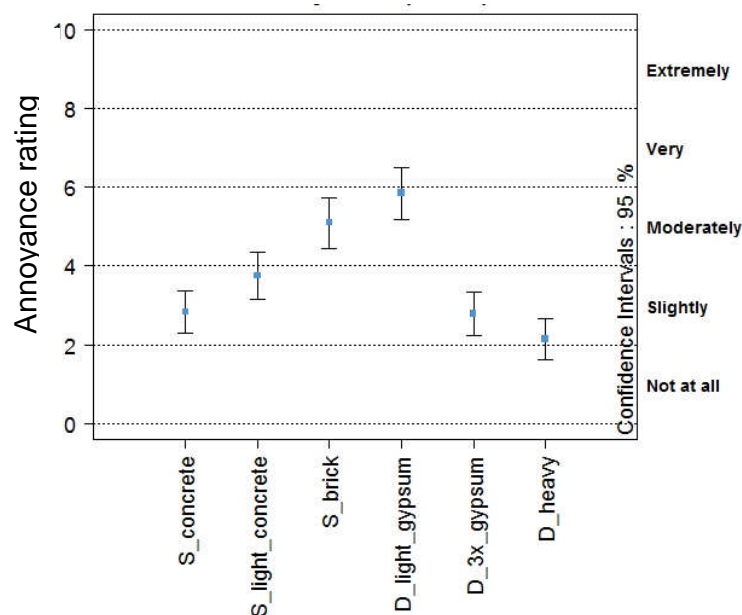


Figure 8

The mean rating of the systems (walls) under test averaged over all 4 samples and all 22 assessors. The bars show the 95 % confidence intervals. The system names and codes are described in Table 4.

testname

Tukey HSD (System)

	S_concrete	S_light_concrete	S_brick	D_light_gypsum	D_3x_gypsum	D_heavy
S_concrete	1	2.6e-13	2.1e-13	2.1e-13	1	1.7e-10
S_light_concrete	2.6e-13	1	2.1e-13	2.1e-13	2.6e-13	2.1e-13
S_brick	2.1e-13	2.1e-13	1	2.9e-12	2.1e-13	2.1e-13
D_light_gypsum	2.1e-13	2.1e-13	2.9e-12	1	2.1e-13	2.1e-13
D_3x_gypsum	1	2.6e-13	2.1e-13	2.1e-13	1	2.9e-09
D_heavy	1.7e-10	2.1e-13	2.1e-13	2.1e-13	2.9e-09	1

p-value of the Tukey HSD test with the unsignificant test in red

Table 5

Matrix showing the Tukey Honestly Significant Difference (HSD) test. Systems which have not been rated significantly different rated are marked with red ($P > 0.05$). The system names and codes are described in Table 4.

6.3 Comparison with physical metrics

Based on the findings from Figure 3, Figure 4, Figure 5 and Appendix D, the graphs in this section are without the results from assessors 5014, 5028 and 5032.

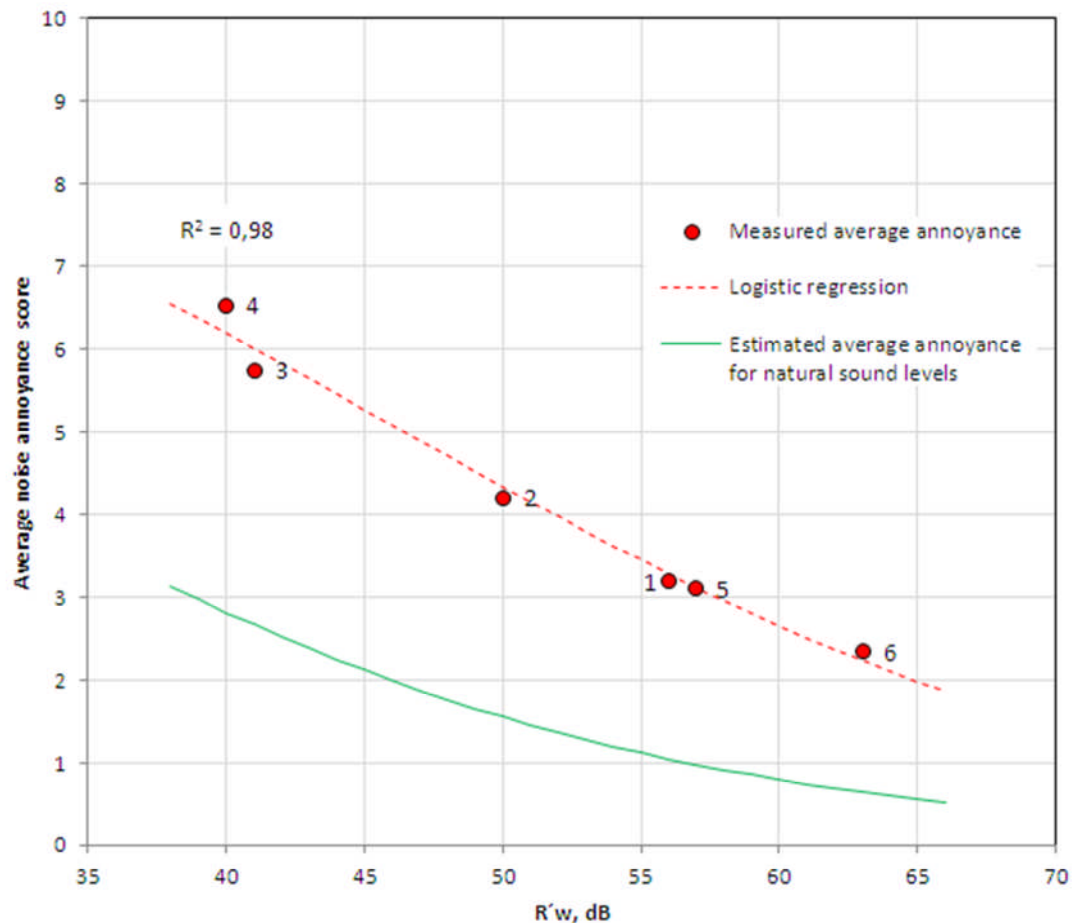


Figure 9

The red dots indicate the measured annoyance averaged over the assessors' responses and the four different stimuli played back approximately 14 dB to loud. The green line indicates the estimated annoyance averaged over the four stimuli at natural level. The parameters for the green curve are: $s = -0.0753$, $f = 27.5$ dB - see Appendix F. The numbers in the symbols indicate the wall construction, see Table 4.

Figure 9 shows the relation between the annoyance scores and the R'_w -values for the 6 systems (simulated walls). It is seen that there is a very good relation ($R^2 = 0.98$ by a logistic regression - see Appendix F) between the annoyance scores (red dots) and R'_w . Due to the large level differences of the 4 samples (sounds) the confidence intervals are rather large, see Figure 8.

The estimated annoyance potential as function of R'_w averaged over the 4 sounds (toilet flush, music, party and voice sounds) at natural levels is shown with the green line in the

figure. This result is found from the L_{Aeq} -Annoyance graph, Figure 11, by decreasing the levels by 14 dB and reading the resulting annoyance potential.

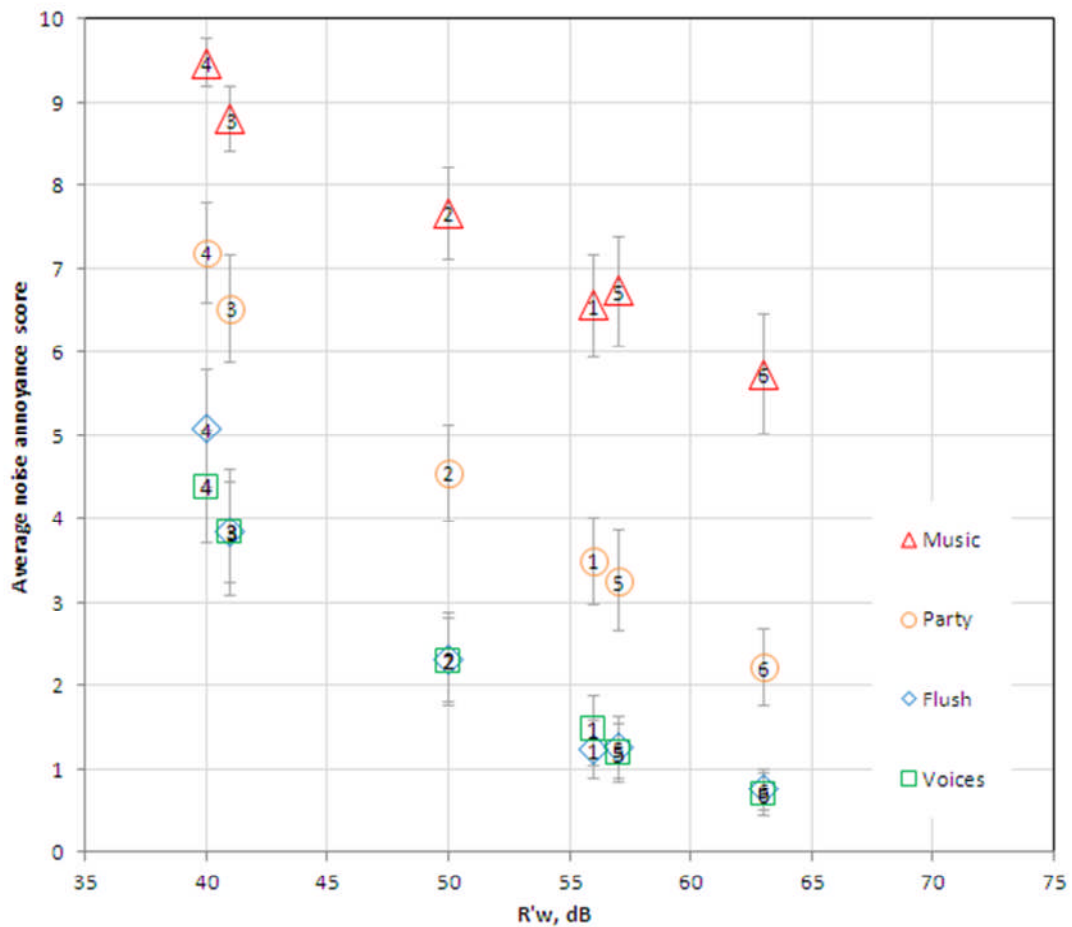


Figure 10

The measured annoyance response averaged over the assessors' responses for each of the four different stimuli played back approximately 14 dB to loud. The vertical bars is the 95 % confidence intervals. The numbers in the symbols indicate the wall construction, see Table 4.

Figure 10 show the annoyance scores for each of the sounds heard through each of the walls as function of the R'_w -values. It is seen that the loudest sounds (see Table 3) have the highest annoyance potential and that within the confidence intervals the same ranking of the walls is obtained independent of the sound samples.

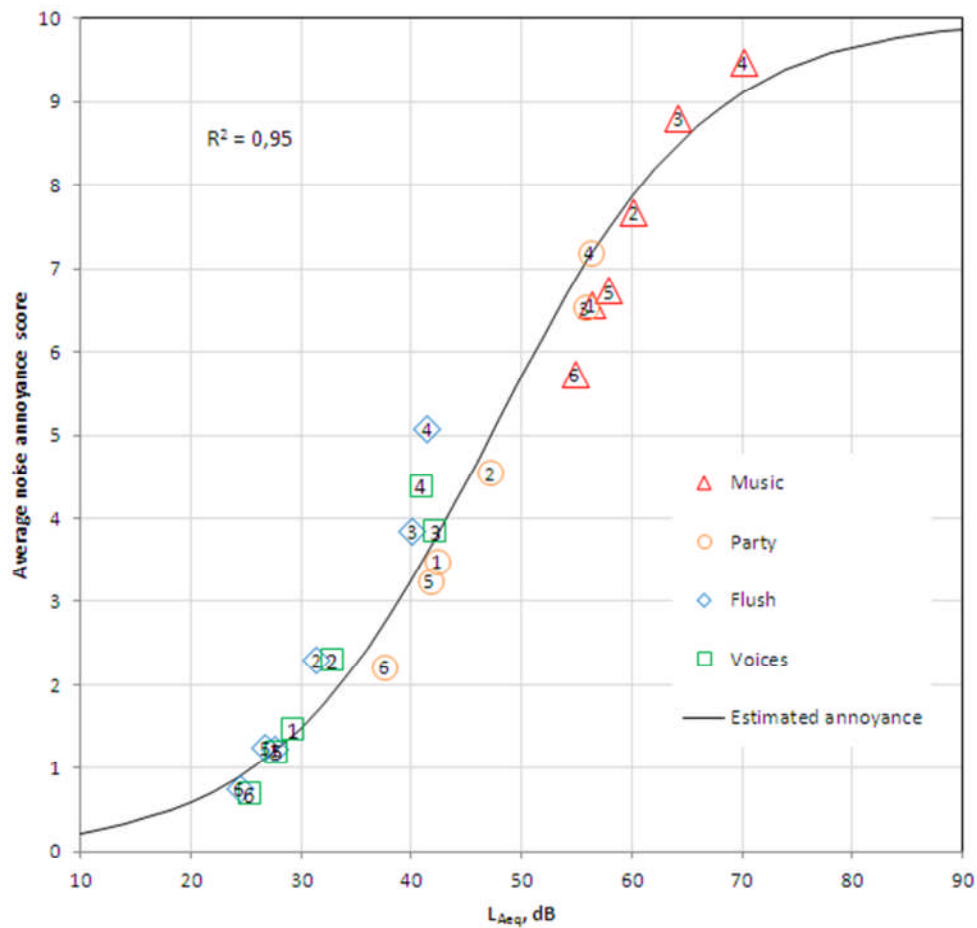


Figure 11

L_{Aeq} - Annoyance. The L_{Aeq} -levels refer to the receiver side of the wall.

Parameters for the estimated annoyance potential are: $s = 0.1016$, $f = 47.2$ dB - see Appendix F. The numbers in the symbols indicate the simulated wall construction, see Table 4.

Figure 11 shows the annoyance ratings as function of the A-weighted sound pressure levels on the receiving side of the wall. It is seen that there is a very high correlation between the A-weighted levels and the annoyance even if the spectra of the sound samples differ.

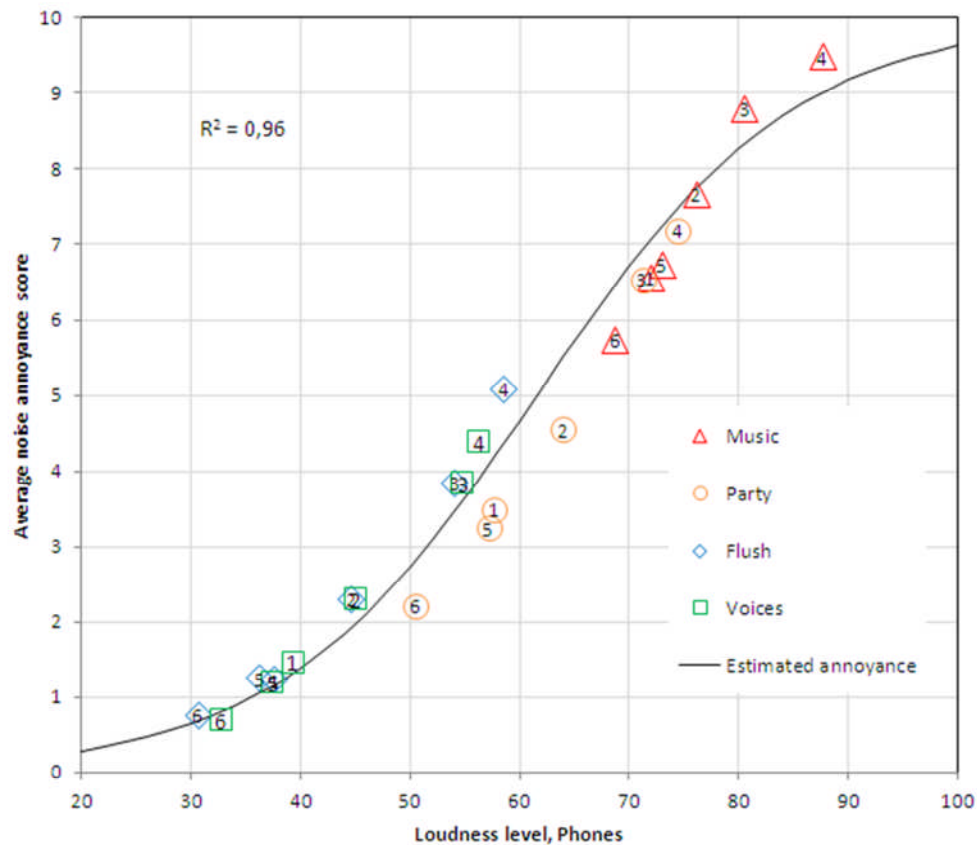


Figure 12

Loudness level - Annoyance. The Loudness levels refer to the receiver side of the wall. $s = 0.0846$, $f = 61.6$ phon, see Appendix F. The numbers in the symbols indicate the wall construction, see Table 4.

Figure 12 shows the annoyance ratings as function of the loudness levels on the receiving side of the wall. It is seen that there is a slightly higher correlation than with the A-weighted levels and the annoyance. As the spectra of the 4 sound samples differ one could expect a better fit with loudness level, but there is not much room for improvement.

6.4 Noise sensitivity as moderator for annoyance potential

In connection with the listening test, the test persons filled out two different questionnaires about their noise sensitivity, see reference [13]. In contradiction to many other investigations the noise sensitivity did not appear as a moderating factor for the annoyance. The main reason for this is supposed to be that the level calibrations are not absolute but depend to some degree on the individual, see clause 3.3.

The two questionnaires were a 9-item questionnaire for noise sensitivity formulated by Zimmer and Ellermeier [14] and a one question scale the noise sensitivity used in the COST questionnaire for building acoustics [12]. The two approaches give very different results, for details see the reference [13].

7. Conclusion

The main purpose was to test online methodologies for listening tests within building acoustics. In the interpretation of the results it should be emphasized that this was not a “real life” setting for neighbour noise (the context is only imagined) and that the attenuations of the wall constructions were simulated.

The SenseLabOnline test

A SenseLabOnline test was made on rating the annoyance potential of neighbours’ activities heard through different simulated walls. This type of application presents new challenges for listening test, mainly related with the large level range of sounds that could be present in a session. Some of the soft sounds may not be audible, if high insulation values are present. The advantage of making a pilot trial at first in order to have a realistic idea about problems that can occur is emphasized. A careful selection of the sound samples should be made. The SenseLabOnline test made it possible to carry out a test with 22 assessors from 11 countries in two days.

A simple adjustment procedure for approximate level calibration

No accurate level calibration was made, but the approximate level adjustment procedure seemed to be sufficient for this test. Even with only approximate calibrated levels at the listener and no premises about the user’s headphones, the overall results of this listening test seems very realistic.

Relations between the average annoyance potential and R'_w

The sound samples were played back at average supposed levels that were 14 dB higher than the natural levels and that made it possible to achieve results for walls with both high and low R'_w -values. By the help of the dose-response curves (see below) it was possible to find the estimated annoyance potential for different R'_w -values for the sounds at natural levels.

A dose-response curve for neighbour noise

The annoyance potential of 24 samples (4 sound samples heard through 6 wall-types) was assessed. This made it possible to find a reliable dose-response curve for the annoyance potential of neighbour noise heard on the receiving side of the walls.

References

- [1] COST Action TU0901 “Integrating and Harmonizing Sound Insulation Aspects in Sustainable Urban Housing Constructions”, 2009-2013, http://w3.cost.eu/index.php?id=240&action_number=TU0901 (public information at the EU COST website) or www.costtu0901.eu (Action website).
- [2] EBU Tech 3276 and EBU tech. 3276 supplement-1
Listening conditions for the assessment of sound programme material: multi-channel sound.
- [3] ISO 8586-2 “Sensory analysis - General guidance for the selection. Training and monitoring of assessors - Part 2: Experts”. International Standards Organization. 1994
- [4] ISO/TS 15666. Technical Specification: Acoustics - Assessment of noise annoyance by means of social and socio-acoustic surveys. 2003-02-01.
- [5] ITU-R Recommendation BS.1116-1
Methods for the Subjective Assessment of small Impairments in Audio Systems including Multi channel Sound Systems. 1997.
- [6] ITU-R Recommendation BS.1534-1
Method for the subjective assessment of intermediate quality level of coding systems (2001-2003). The ITU Radio communication Assembly.
- [7] Le Ray, G. & Pedersen, T. H. (2011). Statistical analysis used in SenseLabOnline. AV 1083/11. DELTA.
- [8] Legarth, S. V. & Zacharov, N. V. (2009). Assessor selection process for multisensory applications. Proceedings of the 126Th Convention of the Audio Engineering Society (Munich. Germany).
- [9] Lorho, G.; Le Ray, G.; Zacharov, N. V. (2010). eGauge - A Measure of Assessor Expertise in Audio Quality Evaluations. AES 38th (Piteå. Sweden).
- [10] Pedersen, T. H. (2007)
The “Genlyd” Noise Annoyance Model. Dose-Response Relationships Modelled by Logistic Functions. AV 1102/07. 2007.
- [11] Rasmussen, B.; Petersen, C. M.; Hoffmeyer, D. (2011). Lydisolering mellem boliger-nybyggeri. SBI-anvisning 237. Aalborg Universitet.
- [12] Simmons and Hagberg (2011) “A Questionnaire for Correlation of Subjective Evaluation of Dwellings with their Objective Building Acoustic Parameters”. Forum Acusticum 2011 paper 414.
- [13] Suhr, A. and Pedersen, T. H.
Methods for investigation of reactions to neighbour noise and noise sensitivity - a pilot study. SenseLab 013/12, 2012. www.madebydelta.com
- [14] Zimmer, K. and Ellermeier, W. (1998). “Ein Kurzfragebogen zur Erfassung der Lärmempfindlichkeit” in: Umweltpsychologie Heft 2, pp. 54-63.

Appendix A - SenseLabOnline annoyance test

This method is intended for the assessment of annoyance potential for sound signals.

Description

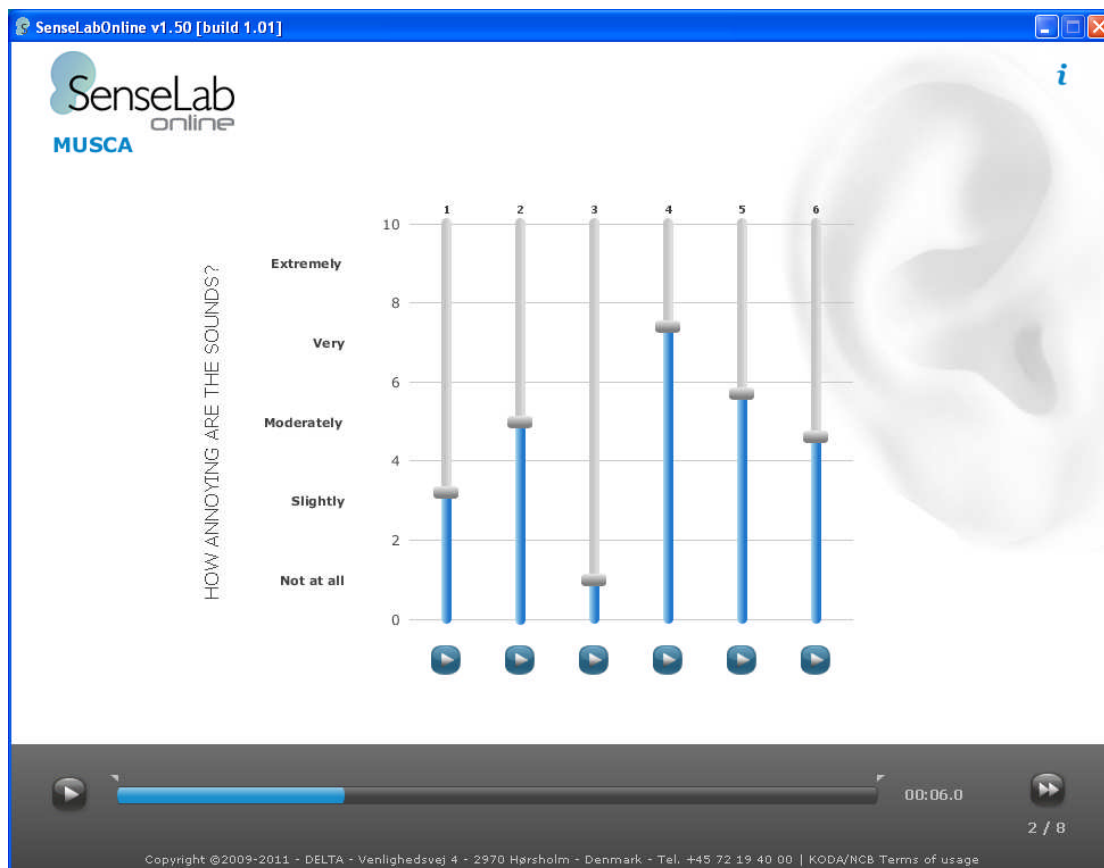


Figure 13

The user interface for the assessors in the SenseLabOnline test on annoyance potential.

The user interface for the SenseLabOnline annoyance test. The sliders are used for the assessment of the stimuli. Below the sliders are play buttons for one of the samples (sounds) heard through the 6 simulated walls in random order. When you click a new play button there is a soft crossover to the new sound within 40 ms. The labels on the scale are in accordance with ISO 15 666. The assessor is allowed to switch forth and back at will between any of the systems under test. The 4 sounds are presented on one screen each and the order is random for each assessor.

SenseLabOnline provide an advanced feature that allows assessors to focus by zooming in on and looping a part of the sound sample, which is found to most relevant.

Assessor requirements

A minimum of 20 assessors is recommended.

Technical requirements

Depending on the requirements the SenseLabOnline test is normally performed either in DELTA's listening room (which fulfills the ITU-R BS.1116-1) or in the assessors' homes with soundcards and headphones supplied by DELTA for less demanding tests. In this special case the test was made in the assessors' homes or offices with their own headphones.

Test planning, administration and reporting

The standards include specific instructions as to the planning (including randomization etc.), administration (instruction to the assessors) and content of the test report. All these matters are handled by SenseLabOnline.

Analysis approach

The scores are assumed to be quantitative and the standard statistical analyses are conducted. Results are reported as mean values and a confidence interval. A Tukey Honestly Significant Difference is also available to sustain the results.



Appendix B - Assessors

People from the COST TU0901 project and a few relative to these were invited by e-mail to participate in an online listening test. The majority of the test persons have acoustic insight but are not trained assessors. 22 persons from 11 countries completed the test, for details see Table 6 below.

Country	Invited	Started	Completed
Czech rep.	1	1	1
Denmark	8	7	6
Finland	1	1	1
France	2	2	2
Italy	3	1	1
Poland	2	2	2
Portugal	3	3	3
Romania	1	1	1
Slovenia	1	1	0
Spain	3	3	3
Slovakia	2	1	1
Sweden	3	2	1
UK	4	2	0
Total	34	27	22

Table 6

A list of participating nationalities and test status for the assessors.

Appendix C - Detailed instruction

Level adjustment with reference speech

.... For PC users:

- 1) Open the Windows Volume control (Control Panel\Sound and audio devices: Audio tab, press Volume for Sound playback)
- 2) Set the sliders for Wave and Volume control at 10-15%
- 3) Open Windows media player, adjust the Media player volume control to 100% and play the file enclosed in this mail.
- 4) Adjust the Windows Volume control to a setting where the voice has the natural volume of a man talking 1 m from you. (You may compare with a real male talking)
- 5) Do not change these settings before and during the listening tests!

For Mac users:

- 1) Adjust the overall audio out level up to 3 “bars”. You can do this by using the speaker icons on your keyboard (shortcut), or via System preference > Audio > Audio Out
- 2) Open iTunes and adjust the iTunes volume control to 100%. Play the file enclosed in this mail in iTunes
- 3) Adjust the overall audio output level to where the voice has the natural volume of a man talking 1 m from you. (You may compare with a real male talking). Do not adjust the volume in iTunes, but use the procedure described in 1).
- 4) Do not change these settings before and during the listening tests!

Then you go to: www.senselabonline.com

David male reference +10dB mono 32 bit adj(-10dB.1)16bit_COST2.wav

Instructions for the listening test

When you are ready to start the test. Close your eyes and concentrate on imagining that you are sitting and relaxing at home and hear the sounds from your neighbour's.

Imagine that the sounds will appear approximately every 10 minutes with the same duration as in this test.

Set the sliders according to how annoyed would you feel.

Don't forget to imagine that you are sitting at home.



Appendix D - Assessor performance and scale usage

This appendix focuses on assessor performance and scale usage, and may be applied for assessor feedback and training purposes.

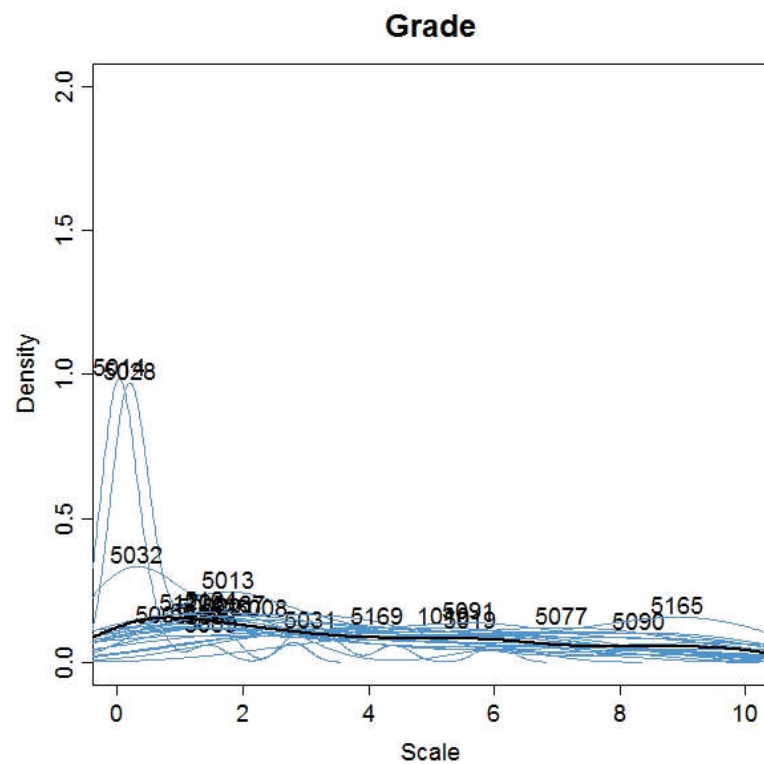
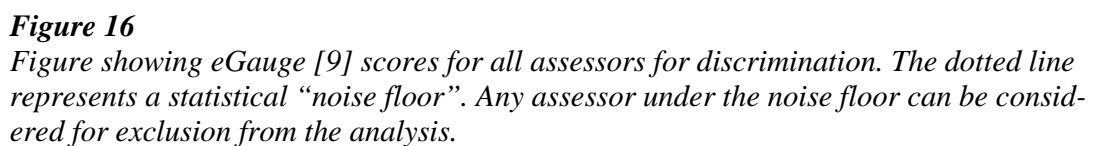
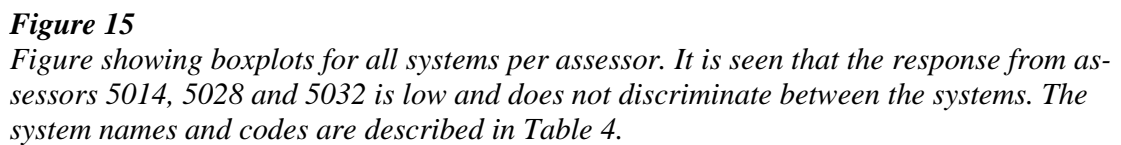


Figure 14

Figure showing the distribution density of ratings for all assessors. Scale usage is well distributed across the scale for most assessors. Only assessors 5014, 5028 and 5032 use a minor part of the scale.



Appendix E - ANOVA assumptions

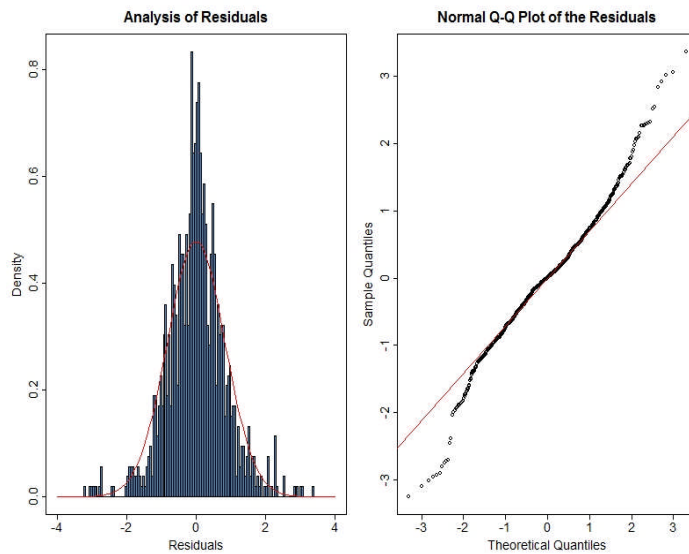


Figure 17
Graphs used for checking normality of the residuals

The first assumption in the ANOVA is normality of the residuals. In Figure 17 the left panel shows a histogram of the residuals which can be seen to fit well with the normal distribution (red curve).

The right panel shows the same information in another way. In this Q-Q plot the normal distribution appears as the straight red line and the actual residuals are shown as black points. It may be easier to see the deviation from the normal distribution in this plot.

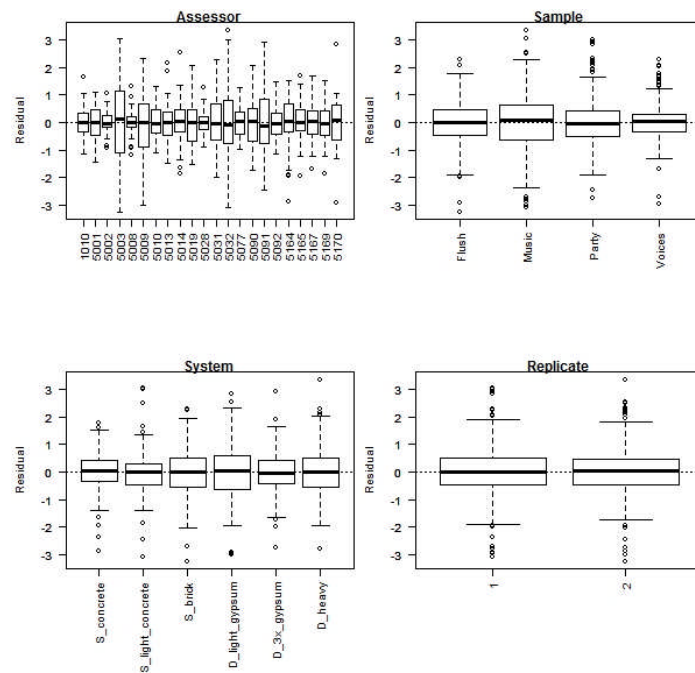


Figure 18

Graph showing the homogeneity of the variance for the test question.

The second ANOVA assumption is the homogeneity of the variance of the residuals, which is represented with the boxplots in Figure 18.

The boxes show the 0.25 quartile and the 0.75 quartile points, i.e. the points on the scale that are exceeded by 25 % respectively 75 % of the data points. In the box the median, i.e. the point on the scale that is exceeded by 50 % of the data points, is shown. The distance between the 0.25 quartile and the 0.75 quartile point are called the “inter quartile distance”. 1.5 times this distance from the upper and lower box limits is often used as a criterion for outliers.

In the linear model methodology it is assumed that the different levels of a factor are evaluated in the same way then the error from this evaluation should be the same and resulting in the homogeneity of the residuals from the different levels.

Therefore the residuals are divided between the different levels of the main factors that are evaluated in the ANOVA model. It is usual that the variance of the residuals from the assessors is not homogeneous. The assessors are not performing with the same accuracy. Whereas one can expect the assessor to evaluate on average the system and the sample with the same accuracy

Appendix F - Calculation of dose-response curves

In this report the dose-response relationships are shown by logistic curves. These curves describe the fundamental properties of dose-response relations. See reference [10]

The dose-response curve is described by:

$$A = \frac{u}{1 + e^{-s(E-f)}} \quad (1)$$

where

- A is a measure for the **A**nnoyance
- u is the **u**pper limit for the scale for A
- s is the **s**lope
- E is the noise **E**xposure), e.g. L_{Aeq} in dB
- f is the level for 50% (**f**ifty) annoyance

The sound pressure level for a certain amount of annoyance can be found from:

$$E = f - \frac{\ln(\frac{u}{A} - 1)}{s} \quad (2)$$

The graphs next page illustrates the meaning of s and f . Further information can be found in reference [10]

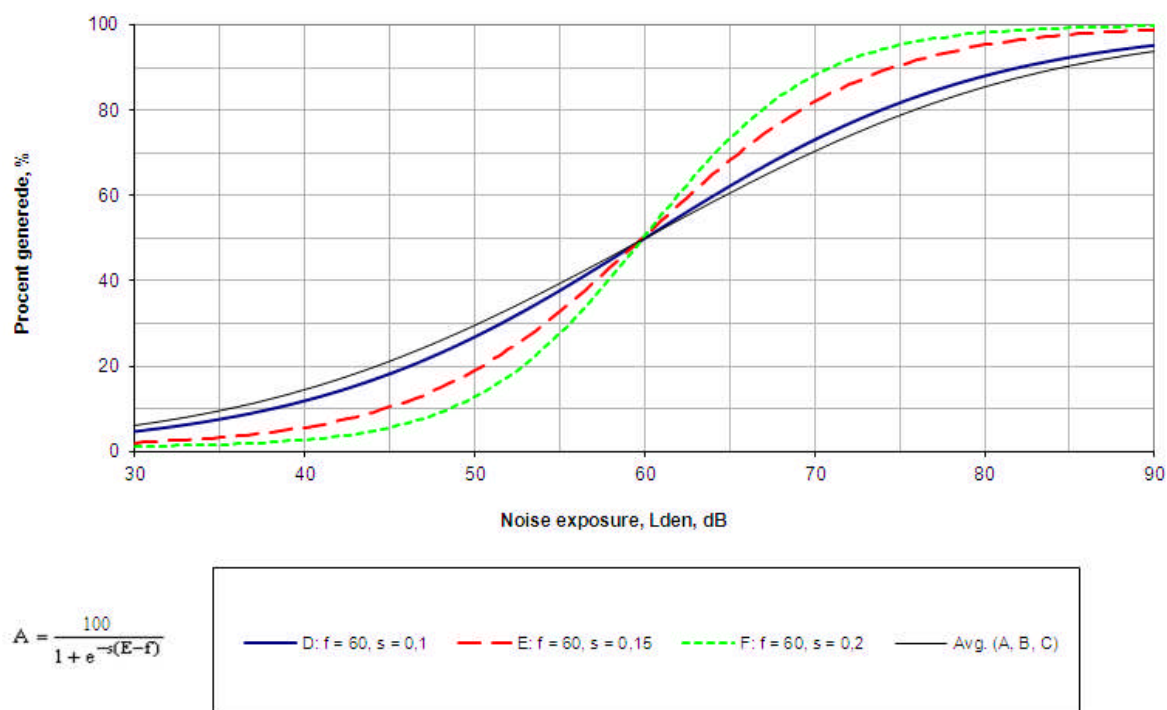
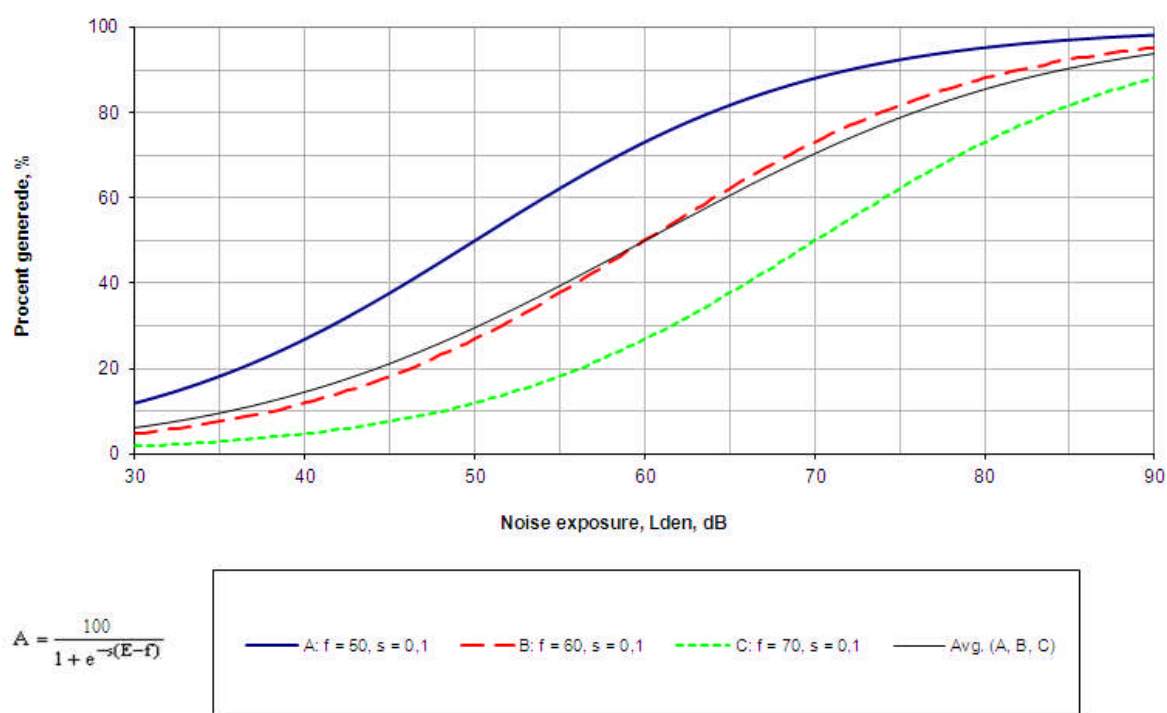


Figure 19

Examples on graphs with different values of the parameters s and f .