Working Report No. 1 2002
Arbejdsrapport fra Miljøstyrelsen

Laboratory Evaluation of Annoyance of Low Frequency Noise

Torben Poulsen and Frank Rysgaard Mortensen
Technical University of Denmark, Acoustic Technology

Danish Environmental Protection Agency
Danish Ministry of the Environment

The Danish Environmental Protection Agency will, when opportunity offers, publish reports and contributions relating to environmental research and development projects financed via the Danish EPA.

Please note that publication does not signify that the contents of the reports necessarily reflect the views of the Danish EPA.

The reports are, however, published because the Danish EPA finds that the studies represent a valuable contribution to the debate on environmental policy in Denmark.

Foreword and acknowledgements

The project described in the present report was initiated by the Danish Environmental Protection Agency and made possible by a collaboration contract between the Agency and The Department of Acoustic Technology at The Technical University of Denmark in Lyngby, Denmark.

We would like to thank all test subjects for collaboration and patience during the time consuming – and sometimes boring - listening tests. Especially the four persons who are suffering from low frequency noise in their daily life are thanked for their willingness to participate in the investigation.

DELTA, Lyngby, provided sound recordings for the listening tests.

The Danish Environmental Protection Agency has followed the investigation with great and encouraging interest during all parts of the project.

Tina Hjort is thanked for providing contact to the special group of test subjects and for a continued interest in the investigation and the results.

Lyngby, May 2001

Torben Poulsen

Table of contents

Introduction. Objective of the Project

Several different assessment methods for the annoyance due to low frequency noise have been proposed during the later years. Also the noise limits or criteria values used with the different assessment methods differ. As a consequence of the ambiguity the Danish Environmental Protection Agency has asked for an investigation, where the subjectively assessed annoyance due to a number of practical examples of low frequency noise was compared with the predicted annoyance using different assessment methods.

Such an investigation could indicate the best suitable method for assessment of low frequency noise. It could probably also indicate a need for an adjustment or a revision of the Danish assessment method that was published in Information No. 9/1997 from the Danish Environmental Protection Agency, "Low frequency noise, infrasound and vibration in the environment" [1].

The project reported here was made in cooperation between the Danish Environmental Protection Agency and the Department of Acoustic Technology at The Technical University of Denmark. M. Sc. Frank Rysgaard Mortensen prepared, conducted and analysed the listening tests, and Associate Professor Torben Poulsen was responsible for the quality control and for the major part of the report.

Preliminary results from the project was presented at the conference “Low Frequency noise and Vibration” in Aalborg, Denmark, May 2000 [2].

The views that are expressed in this report do not necessarily reflect the attitude of the Danish Environmental Protection Agency. It is found however that the project represents a valuable contribution to the investigation about noise and its effect on people.

Summary

The present report documents a project that was carried out in cooperation between The Department of Acoustic Technology, Technical University of Denmark and the Danish Environmental Protection Agency in year 2000. The purpose of the project was to investigate the subjective annoyance due to a number of low frequency noise examples. The main part of the project consisted of a series of listening tests, where a group of test persons listened to different examples of low frequency noise and gave their evaluation of the loudness of the noises and the annoyance of the noises. The main group of test persons (reference group) was eighteen young persons with normal hearing.

The test persons evaluated eight different noise examples, where one example was road traffic noise and the other examples were realistic samples of environmental low frequency noise. The noise examples included both stationary noise with and without tones, intermittent noise (from a passing ferry), music, and impulsive low frequency noise from a punch press. The noises were presented in an IEC listening room at three different L_{Aeq} levels between 20 and 35 dB, and each presentation was repeated once in a random order.

The subjective assessments made by the test persons were compared to objective results from a number of recent methods for assessment of low frequency noise. A total of seven different assessment methods were used. It was found that the Danish assessment method, published in Information No. 9/1997 from the Danish Environmental Protection Agency, "Low frequency noise, infrasound and vibration in the environment" gave the best relation to the subjective assessments made by the test persons. There was no obvious indication in the present results of how to further improve the Danish method. Apparently one important property of this method is that it includes a 5 dB penalty in case of impulsive noise.

In addition to the main investigation with 18 young normal hearing test persons, listening tests were also made with a small group of four test persons, the special group. The special group consisted of four persons that had reported annoyance due to low frequency noise in their homes. It was found that the special group assessed the noise examples in a different way than the reference group. Generally the annoyance was much higher, especially the annoyance at night. It was also found that the two groups evaluated the noise samples differently. It was not the same types of noise that the two groups found most and least annoying. Due to the limited number of persons in the special group, no explanations of these differences can be indicated within the frame of the present investigation.

Udvidet resumé

Denne rapport dokumenterer et projekt, som blev udført i samarbejde mellem Institut for Akustisk Teknologi på Danmarks Tekniske Universitet og Miljøstyrelsen i år 2000. Projektets formål var at undersøge genevirkningen af en række eksempler på lavfrekvent støj. Baggrunden for projektet er, at der de seneste år er offentliggjort flere forskellige vurderingsmetoder, som alle har til formål at sætte mål på genevirkningen af lavfrekvent støj. Metoderne adskiller sig en del fra hinanden, og de tilhørende kriterie- eller grænseværdier er også temmelig forskellige. Der er således nogen usikkerhed om, hvordan en objektiv vurdering af genevirkningen fra lavfrekvent støj kan udføres.

Resultatet af projektet kunne muligvis pege på en velegnet objektiv metode til vurdering af lavfrekvent støj, eller på behovet for en revision eller en justering af den nuværende vurderingsmetode, som er beskrevet i Orientering fra Miljøstyrelsen nr. 9/1997 "Lavfrekvent støj, infralyd og vibrationer i eksternt miljø" [1].

Projektets hoveddel udgøres af lytteforsøg, hvor forsøgspersoner vurderede en række eksempler på lavfrekvent støj med hensyn til lydstyrke og oplevet gene under forudsætning af, at støjen kunne høres om dagen og aftenen eller om natten. Hovedgruppen af forsøgspersoner bestod af 18 unge mennesker med normal hørelse. Forsøgspersonernes hørelse blev kontrolleret både ved normal audiometri (i frekvensområdet 125 Hz - 8 kHz) og i lavfrekvensområdet ned til 31 Hz.

Der blev benyttet 8 forskellige støjeksempler, hvor et eksempel var vejtrafikstøj, som skulle bruges som reference idet netop genen af vejstøj er godt belyst i litteraturen. De syv andre støjeksempler var realistiske lydoptagelser af lavfrekvent ekstern støj. Der indgik stationær støj både med og uden toner, intermitterende støj (fra en forbisejlende færge), musik, og impulsagtig støj fra en smedehammer, hvor lyden var transmitteret igennem jorden. Alle støjeksemplerne blev præsenteret via højttalere i et IEC lytterum, og de blev afspillet ved tre forskellige L_{Aeq} niveauer mellem 20 dB og 35 dB. Hver præsentation blev gentaget to gange.

Forsøgspersonernes subjektive vurdering af støjens genevirkning blev sammenholdt med objektive målinger på de samme støjtyper. Syv nyere vurderingsmetoder til brug for lavfrekvent støj blev benyttet. Ud over den danske vurderingsmetode blev den svenske (SOSFS 1996:7) og den tyske metode (DIN 45 680) benyttet tillige med et forslag til en polsk og to hollandske vurderingsmetoder. Endelig blev det forsøgt at bruge støjens C-vægtede lydtrykniveau som parameter. Det blev fundet (rapportens afsnit 6.8) at den danske vurderingsmetode (Miljøstyrelsens orientering nr. 9/1997) giver den bedste sammenhæng med den subjektive vurdering. Undersøgelsen har ikke indikeret hvordan metoden eventuelt kan forbedres eller justeres. Det tyder på, at en vigtig egenskab ved metoden er, at grænseværdien skærpes med 5 dB når der er tale om impulsagtig lavfrekvent støj; den skærpede grænseværdi er i undersøgelsen også brugt for den støjtype som bestod af musik.

Ud over hovedundersøgelsen blev der gennemført tilsvarende lytteundersøgelser med en mindre gruppe af forsøgspersoner, specialgruppen. Der er tale om fire personer, som oplever at være generet af lavfrekvent støj i deres hjem. Det var ikke et kriterium, at deltagerne i specialgruppen skulle have normal hørelse; nogle af forsøgspersonerne havde et markant høretab ved høje frekvenser. Specialgruppen havde ikke bedre hørelse ved lave frekvenser end referencegruppen. Undersøgelserne viste, at specialgruppen havde en anderledes vurdering af støjeksemplerne end referencegruppen. Først og fremmest angav specialgruppen en tydeligt kraftigere gene for de samme støjeksempler, især genen om natten var meget højere end referencegruppens vurdering. Forskellen mellem de to gruppers vurdering svarer omtrent til, at støjens styrke blev forøget med 15 dB. Dernæst blev det fundet, at det ikke var de samme støjtyper som de to grupper af forsøgspersoner fandt henholdsvis mest og mindst generende. På grund af det lille antal forsøgspersoner har det ikke været muligt at gå i detaljer for at finde en forklaring på de meget forskellige vurderinger.

.

1 Background

In 1997, The Danish Environmental Protection Agency published Information No. 9/1997 "Low frequency noise, infrasound and vibration in the environment" [1]. Here a general description is given of generation and transmission of low frequency noise and of the properties of hearing in the low frequency and infrasound region. Also recommended measurement and assessment method for the annoyance is described and recommended limit values are stated for environmental infrasound and low frequency noise. Contrary to usual measurement and assessment of environmental noise, such as road traffic noise or industrial noise, measurement of environmental infrasound or low frequency noise shall be made indoors in dwellings.

1.1 Infrasound

Sound in the frequency range below 20 Hz is defined as infrasound. Previously it was believed that infrasound was inaudible, and the three decades 20 Hz – 20 kHz was often called ‘the audible frequency range’.

We now know that infrasound can be heard provided it is strong enough. The threshold of hearing is determined at least down to 4 Hz, and there is a reasonable agreement between the recently measured thresholds. The G-weighting function has been standardised in ISO 7196 [3] and has a close relation to the shape of the hearing threshold in the infrasound region but also includes a sharp cut off at frequencies above 20 Hz. A determination of equal loudness level contours in the infrasound region at 20, 40, and 60 Phones was reported as early as in 1984 [4]. The slope of the lowest loudness level contour was comparable to the slope of the threshold curve, while the higher loudness level contours had slightly less slope. As a first approximation it may thus be expected that the G-weighted infrasound level would predict the strength or loudness of the infrasound, at least when the level is not too high above the threshold.

Infrasound is usually not perceived as a tonal sound but rather as a pulsating sensation, pressure on the ears or chest, or other less specific phenomena. The loudness and annoyance due to infrasound increase very quickly with increasing level. The hearing threshold for single tones is usually about 95 dB(G), and tones with a 20 dB higher level are expected to be sensed as very loud. It can be assumed that infrasound below the hearing threshold is not annoying.

The recommended limit value for infrasound in dwellings in Denmark is 85 dB(G). The Danish Environmental Protection Agency explains that the purpose is to secure that even persons with a very low hearing threshold are not annoyed by infrasound. After the ‘Information No 9’ was published in 1997 a number of measurements of environmental infrasound has been made in dwellings. Only at rare occasions levels have been seen near the recommended limit value. In most situations the level is at least 20 dB lower than the 85 dB limit.

1.2 Low Frequency Noise

Whereas infrasound is a well-defined concept, low frequency noise is not. In the Information No. 9/1997 [1] low frequency noise is defined to comprise the frequency range 10 Hz – 160 Hz, but other references – including other assessment methods – use other frequency ranges usually within the interval 8 – 250 Hz. Low frequency noise is noise with a strong content at low frequencies, giving the noise a rumbling character different to e.g. road traffic noise.

The concept of low frequency noise was introduced as this type of noise was expected to be more annoying than noise without the strong low frequency content. This expectation is mentioned in ‘Guideline on environmental noise from industry’ [5]. Here it is also mentioned that there ‘is no universally accepted assessment method for low frequency noise at present’. Since then, several countries have published measurement and assessment methods for low frequency noise; reference is made to Chapter 4 in the present Report. Many assessment methods use the spectrum of the noise (1/3-octave spectrum measured indoors), and compare this spectrum to a criterion curve, which must not be exceeded in any 1/3-octave band. In the Danish assessment method the noise is evaluated by the A-weighted level of the noise in the frequency range consisting of all the 1/3-octave bands from 10 Hz to 160 Hz.

Depending on the actual conditions, many types of noise can be regarded as low frequency noise. The firing rate of many diesel engines is usually below 100 Hz, so road traffic noise can be regarded as low frequency noise as well as (diesel) train noise or noise from ferries. Similar considerations can be made for engines or compressors in industries or co-production plants. Burners can emit broadband low frequency flame roar. Low frequency noise can be noise or vibration from traffic or from industries, totally or partly transmitted through the ground as vibration and re-radiated from the floor or the walls in the dwelling. By this transmission, frequencies above approximately 20 Hz are attenuated. It is a general observation that indoor noise is perceived as more ‘low-frequency-like’ than the same noise heard out of doors.

1.3 Motivation for the Investigation

The Danish Environmental Protection Agency has the opinion that hearing and assessment of infrasound is reasonably well described. In addition it is extremely rare that infrasound in the environment is anywhere near the recommended limit value.

Contrary to this there is an obvious need for investigations where the subjective annoyance due to typical examples of low frequency noise is compared to different objective measures of the strength of the same noises. There are several different features of the different assessment methods presently in use, and the corresponding limits or criteria values differ. The fundamental assumptions for the assessment methods are, however, largely the same. (I.e. the same investigations and the theories are mostly based on experience).

2 Description of the Listening Tests

The listening tests were made in a standardised listening room at the Department of Acoustic Technology, Technical University of Denmark, where the noise examples were presented for the test persons. The dimensions of the listening room are 7.52 x 4.75 x 2.76 m, and the room fulfils the recommendations in IEC 268-13 [6]. A total of eight different noise examples were used, presented at three different levels. All presentations were made twice and the sequence of the presentations was randomised. Prior to the listening tests, the test persons were trained using four noise examples. Thus each test person has assessed more than fifty presentations. After each presentation the test person gave evaluations of the noise example on a paper form.

2.1 Noise Examples

Eight different noise examples were chosen for the listening tests. Example No. 1 was the noise from a densely trafficked six-lane highway. The traffic noise served as a reference noise, because there is a well-described relation between the level of road traffic noise and the annoyance of this type of noise. The other noise examples had all a strong low frequency content. The noise examples were:

No.	Name	Description	Tones, characteristics
1	Traffic	Road traffic noise from a highway	None – broadband, continuous
2	Drop forge	Isolated blows from a drop forge transmitted through the ground	None – deep, impulsive sound
3	Gas turbine	Gas motor in a power-and-heat plant	25 Hz, continuous
4	Fast ferry	High speed ferry; pulsating tonal noise	57 Hz, pass-by
5	Steel factory	Distant noise from a steel rolling plant	62 Hz, continuous
6	Generator	Generator	75 Hz, continuous
7	Cooling	Cooling compressor	(48 Hz, 95 Hz) 98 Hz, continuous
8	Discotheque	Music, transmitted through a building	None, fluctuating, loud drums

Table 1. Description of the noise examples used for the listening tests

Noise example 1 has a broadband character and it is almost continuous. Since it was filtered to simulate an indoor measurement (see next section), the tonal character of the engine noise from passing heavy vehicles is clearly audible though the tire noise is also obvious. Noise example 2 consists of a series of very deep, rumbling single blows from a drop forge. The noise examples 3, 4, 5, and 6 each have one tonal component. Noise example 7 has three tones but two of them are at a low level, and noise example 8 has a characteristic rhythmical pulsating appearance due to the drums.

2.2 Noise duration, levels and presentation

The duration of all the noises was 2 minutes. The noise examples were either recorded indoor or filtered to simulate indoor noise. They were taken from DAT recordings to the hard disk of a PC where they were edited digitally. The examples are presented to the test persons at nominal levels of 20 dB, 27.5 dB, and 35 dB (A-weighted levels). In the listening room the sound was measured at the listening position and subsequently analysed to obtain the objective measurements of the noises. The noise examples were played directly from the PC via a D/A converter to a crossover filter and via four amplifiers to two broadband loudspeakers (KEF 105) and two subwoofers (Amadeus Sub). A detailed description of the set-up can be found in [7].

2.3 Filtering of the Noise Signals

The noise examples were filtered with two types of filtering prior to the presentation. One type of filter compensates for resonance modes in the listening room. The listening position was chosen so that only one mode at 45.5 Hz influenced the sound level. At this mode the room mode increases the sound pressure level by about 18 dB, which is compensated for by a narrow band notch filter as illustrated in Figure 1. This filter was used in all presentations.

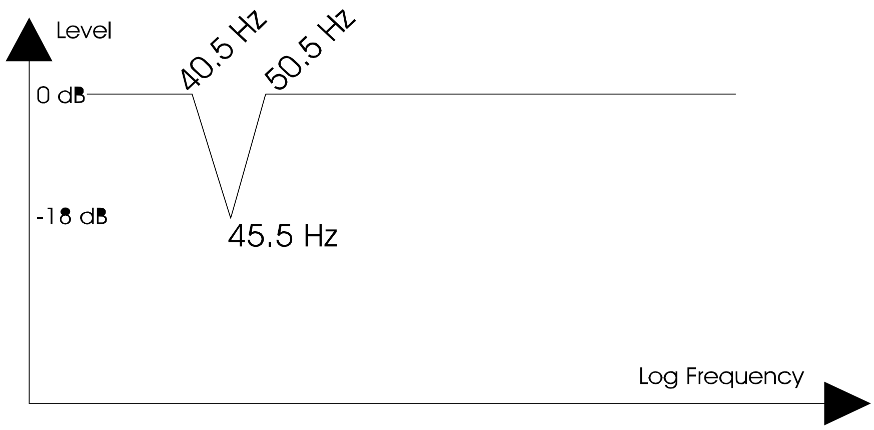


Figure 1. Illustration of the filter that compensates for the resonance modes in the listening room.

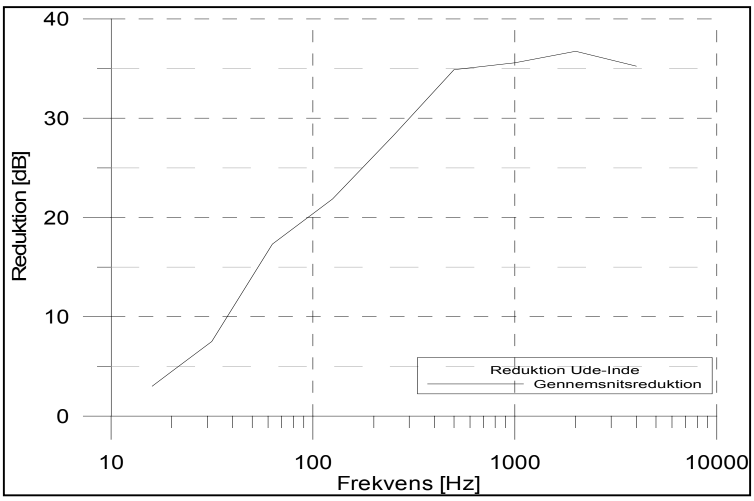


Figure 2. Illustration of the 'outdoor-to-indoor' filter used with the outdoor sound recordings.

The other type of filter was used only with those noise examples that were recorded outdoors. Based on the information in [8], [9] and [10] an ‘outdoor-to-indoor’ filter was created that represents the reduction index of ordinary building materials and construction principles. The filter was defined in the range 16 Hz – 4000 Hz and is illustrated in Figure 2. The level was further rolled off below 15 Hz to suppress wind noise and other ‘errors’ in the recordings that would otherwise be enhanced. Also the level above 4000 Hz was attenuated, as this frequency range was not considered in the project.

From a subjective evaluation, the noise examples sounded ‘natural’ in the listening room. All the examples – including the traffic noise – had a pronounced low frequency characteristic.

The outdoor-to-indoor filtering of the noise examples recorded outdoors changed also the characteristic of the traffic noise. Although the noise was easily recognised as traffic noise, the idea of using this noise as a kind of reference became less obvious. It seems that the test subjects regarded the traffic noise as one of the low frequency noises.

2.4 Test Subjects

Eighteen young persons (9 males and 9 females) with normal hearing were chosen as reference group for the listening tests. The age of the test persons was between 19 and 25 years. This group of listeners is called ‘Reference group’. In addition, listening tests were made with a small group of people that have reported annoyance due to low frequency noise in their homes. This group (called ‘Special group’) consisted of four persons aged between 41 and 57 years, two of each gender. The inclusion of the special group was regarded as a pilot experiment only because of the small number of test persons and the significantly higher age range.

The main results of the project will be based on the listening tests with the reference group. The results from the pilot experiment with the special group are described in Chapter 7.

Pure tone audiometry was made in the frequency range 125 Hz to 8000 Hz with a Madsen Midimate 602 audiometer, equipped with Sennheiser HDA 200 earphones. The calibration of the audiometer was made with the values from [11] which are practically identical to ISO 389-8 [12]. The results of the audiometry are given directly as hearing threshold level in dB HL. Hearing threshold levels at or below 15 dB HL was accepted in the frequency range 125 Hz to 4000 Hz, and a hearing threshold level at 20 dB at a single frequency (including 8 kHz) was also accepted. The average hearing threshold is shown in Figure 3.

In addition to the conventional audiometry, the hearing threshold in the low frequency range was determined. The tests were made with pure tones at 31 Hz, 50 Hz, 80 Hz, and 125 Hz, and the signal length was 800 ms. A computer controlled Tucker Davis system with Sennheiser HDA 200 earphones was used to measure the thresholds according to the Two Alternative Forced Choice method. The present implementation of the procedure determines the 79.4 % point of the psychometric function. A detailed description of the procedure can be found in [13]. The results of these hearing threshold measurements were given as sound pressure level (SPL). No reference values are available at present for the HDA 200 earphone at frequencies below 125 Hz and therefore the measured SPL-thresholds were compared to the standard hearing thresholds given in ISO 389-7 [14]. In this way the deviation from the standardised ISO threshold could express approximate, artificial dB HL values. The average of these artificial hearing levels are also shown in Figure 3. It can be seen in Figure 3 that the two sets of curves differ by 5 dB at 125 Hz. This is due to the artificial reference for the low frequency curves and also because the definition of ‘threshold’ differs in the two methods.

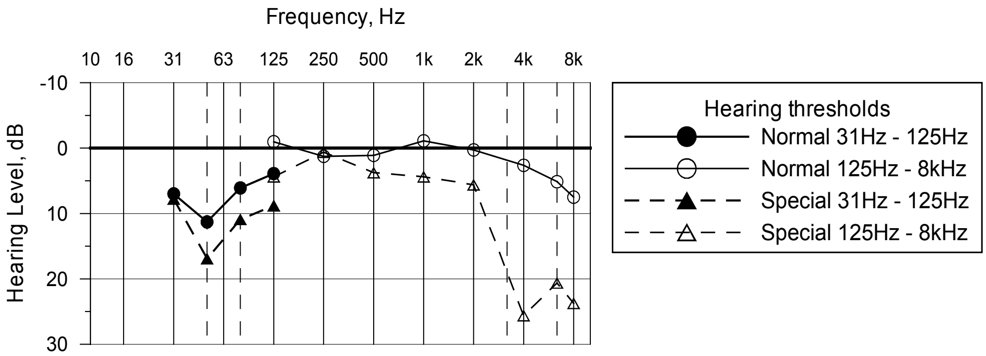


Figure 3. Average pure tone hearing thresholds for the reference group (normal hearing, young subjects) and for the special group. Curves are given separately for the conventional audiometry (125 Hz – 8 kHz) and for the low frequency thresholds (31 Hz - 125 Hz).

From Figure 3 it is seen that the threshold of the reference group has a slight decrease at the highest frequencies (6 kHz and 8 kHz). The special group show a hearing loss at the highest frequencies above 2 kHz. It should be noted though, that the data for the special group are average values over just four persons. Some of the persons had a considerable hearing loss, partly due to a higher age. It is assumed that the high frequency hearing loss does not influence the subjective evaluations of the low frequency noises used in the present investigation.

The low frequency part of the hearing thresholds (Figure 3, 31 Hz – 125 Hz) shows that the special group is about 5 dB less sensitive than the reference group. At 31 Hz the (average) thresholds are almost the same. The threshold data show that the special group of people – who have reported annoyance due to low frequency noise in their homes – do not have higher (better) hearing sensitivity at the low frequencies. The hypothesis that the special group should be able to hear the low frequency sounds more easily than the reference group is not supported by these hearing threshold measurements.

2.5 Subject’s Task

The test persons were given a written introduction to the tests, and they could ask about the procedure throughout the tests. A full training session was made prior to the listening tests. Information about the sound examples was given after all the tests were finalized.

The tests persons were asked to answer four questions after each presentation:

- ‘How loud is the sound?’ (on a scale labelled “not audible” in one end and “very loud” in the other end)
- ‘How annoying do you find the sound if it was heard in your home during the day and the evening?’ (on a scale labelled “not annoying” in one end and “very annoying” in the other)

- ‘How annoying do you find the sound if it was heard in your home during the night?’ (on a scale labelled “not annoying” in one end and “very annoying” in the other)
- ‘Is the noise annoying?’ (answer yes or no).

All the scales were 10 cm long, and the response was measured in cm with a ruler and thus given as a figure between 0 and 10.

An example of one of the annoyance scales is given here:

not annoying very annoying

3 Primary Results of the Listening Tests

The primary results of the listening tests are the scalings made by the test subjects. The scalings are given as values between zero and ten as measured from the test subject's indications on the response lines. All the results are shown in Annex A of this report. In chapter 6 it is shown that the annoyance during day/evening and the annoyance at night are very closely related and thus it suffice at this point to look at only one type of annoyance recordings. Table 2 below shows the average subjective evaluation - made by the reference group of listeners - of the annoyance during night from the various sound examples.

Nominal presentation level	20 dB	27.5 dB	35 dB
Noise example	**Subjective annoyance Night**	**Subjective annoyance Night**	**Subjective annoyance Night**
Traffic noise	1.6	3.4	5.2
Drop forge	4.3	5.9	6.9
Gas turbine	0.9	2.5	5.2
Fast ferry	0.9	3.2	5.4
Steel factory	1.0	2.7	4.9
Generator	1.7	3.2	5.0
Cooling compressor	2.7	4.4	6.0
Discotheque	3.0	5.4	6.7

Table 2. Subjective assessment by the reference group of the annoyance from the noise examples if the noise was heard at night. Annoyance rating is given on a scale from 0 (not annoying) to 10 (very annoying).

It can be seen from the results in this Table 2 that the subjectively assessed annoyance increases when the same type of noise is played louder; this is a general as well as an expected result. It can also be seen that the different types of noise are not assessed equally annoying; apparently the noises from the drop forge, the discotheque and the cooling compressor are evaluated as more annoying than the other types of noise. This gives some promise for interesting results from a closer inspection of the different objective assessment methods – they should be able to give some form of explanation why these types of noise are considered more annoying than the others. On the other hand it can be seen that the traffic noise is just as annoying as the main part of the low frequency noise examples. It was the intention that this type of noise should serve as a reference noise (which was not a particularly low frequency noise), and the listening experiments should then indicate how much more annoying a number of different low frequency noises would appear. But it seems that the filtering applied to the traffic noise has turned this into another low frequency noise example.

3.1 Statistical Analysis of Primary Results

In order to investigate the structure lying behind the average data, the raw data from each listening test were typed into a spreadsheet. For each data the following information was also recorded: the age and gender of the test person, the repetition number (round 1 or 2), the sound example number, the nominal presentation level, the measured A-weighted level (dB(A)), and the A-weighted level of the sound in the frequency range 10 Hz – 160 Hz, $L_{pA,LF}$.

All data were subsequently transformed to a statistical analysis program (Statgraphics 4.0). It was found that if one disregarded data near the endpoints of the scales, the responses almost followed a normal distribution curve. However since many of the responses were near one or the other of the endpoints, the primary results were not normal distributed. Despite this lack of normality in the distribution of the data it was decided to perform an analysis of variance. The analysis was made for each parameter separately.

Table 3 shows the significance levels of the influence from a number of different factors upon the evaluations of the reference group. If the number is less than 0.05, this factor has a significant effect on the evaluation on a 95% level or above (this means less than 5 % probability for drawing a wrong conclusion). If the number is above 0.05 it cannot be proved that this factor has a significant effect upon the relevant evaluation.

	Loudness	**Annoyance day**	**Annoyance night**	**Annoying? (Y/N)**
Noise example	0.0000	0.0000	0.0000	0.0000
Nominal level	0.0000	0.0000	0.0000	0.0000
dB (A)	0.0000	0.0000	0.0000	0.0000
$L_{pA,LF}$	0.0000	0.0000	0.0000	0.0000
Repetition no.	0.5814	0.6123	0.6804	0.1533
Gender	0.1888	0.0001	0.0001	0.0654

Table 3. The significance level of different factors that may influence the evaluation by the reference group.

It is seen from Table 3 that – as expected – the noise example, the nominal level, the dB(A) level and the low-frequency level ($L_{pA,LF}$), all have a significant influence upon the evaluations from the test persons.

The repetition number (round 1 or round 2 with the same presentation) has no significant influence, which shows the absence of a training effect. The gender of the test persons has influence on the evaluation of annoyance during the day and during the night but not on the evaluation of loudness and on the yes/no question about whether the noise is annoying or not.

3.2 Statistical Analysis of the Results from the Special Group

A corresponding analysis was made with the data from the special group. Since this group has only four persons the data are very uncertain and highly dependent on random variations. The result of the analysis is shown in Annex A. It is found that the noise level influences the evaluations. The influence from noise example on the annoyance evaluations is just at the limit of being significant.

4 Assessment Methods for Low Frequency Noise

A number of different methods have been suggested for the assessment of low frequency noise. In this investigation the Danish method [1] is compared to the standardised German method [15], to the Swedish method [16], to a recent Polish method [18], and to two different methods from the Netherlands [20] and [19]. These methods are used in different situations to assess the annoyance due to low frequency noise, based on the indoor noise level. They give different guidelines or criteria for the allowed noise level. Furthermore, the administrative procedures used in the individual countries to enforce the criteria for low frequency noise are very different. This question, however, has not been regarded in this project but is the object of another investigation initiated by the Danish Environmental Protection Agency [21].

4.1 Danish Method

The Danish method [1] is described in ”Information No. 9/1997 from the Danish Environmental Protection Agency”. It also gives the recommended limit values for low frequency noise and infrasound. The recommended measurement method is specified. The noise is measured in several positions indoor, and is analysed in 1/3-octave bands. The nominal A-weighting corrections are added to the spectra, and the weighted spectrum is summed to form the A-weighted level of the noise in the frequency range 10 Hz – 160 Hz. The resulting level is called $L_{pA,LF}$.

It would not be possible make a direct measurement of the A-weighted level, $L_{pA,LF}$, since the minimum limit of the tolerance for the A-weighting filter is undefined (i.e. minus infinity) below 20 Hz. The nominal A-weighting is shown in Table 4.

Frequency, Hz	**8**	**10**	**12.5**	**16**	**20**	**25**	**31.5**	**40**
A-correction, dB	-77.8	-70.4	-63.4	-56.7	-50.5	-44.7	-39.4	-34.6
Frequency, Hz	**50**	**63**	**80**	**100**	**125**	**160**	**200**	**250**
A-correction, dB	-30.2	-26.2	-22.5	-19.1	-16.1	-13.4	-10.9	-8.6

Table 4. The nominal A-weighting corrections

In the Danish method a table of recommended limit values is used for assessment of the noise. In dwellings the A-weighted equivalent level (averaged over 10 minutes) shall not exceed 20 dB $L_{pA,LF}$ in the evening and the night (18 – 07) or 25 dB $L_{pA,LF}$ in the day period (07 – 18). In offices, teaching rooms etc the A-weighted level shall not exceed 30 dB, and in other rooms in enterprises the limit is 35 dB. If the noise has an impulsive character, the limits are reduced by 5 dB.

4.2 German Method

In the German method [15, DIN 45 680 / 1997] low frequency noise is defined as noise where the C weighted noise level is at least 20 dB higher than the A-weighted level, based on either equivalent levels or maximum levels.

If the noise is evaluated to be 'low frequency', a 1/3-octave frequency analysis is made. The method considers the frequency range 10 Hz – 80 Hz, but in special situations the 8 Hz and / or the 100 Hz band can be included. The method applies to rooms in dwellings where people stay or rests. In an Annex to the method a range of limit or criteria values are given for the day period (06 – 22) and for the night period (22 – 06).

In the German method, distinction is made between tonal noise and noise without tones. If the level in a particular 1/3-octave band is 5 dB or more above the level in the two neighbouring bands, the noise is said to be tonal.

- For tonal noise, the level of the frequency band with the tone is compared to the hearing threshold (L_{HS}) in the same band. It is then found how much the tone is above the threshold. The levels in the other frequency bands are not taken into account. The limit value for the equivalent level of the tone in the day period is: 5 dB in the 8 Hz – 63 Hz bands, 10 dB in the 80 Hz band, and 15 dB in the 100 Hz band. The same assessment method applies to the maximum level of the noise; here the limit values in the same three frequency ranges are 10, 15, and 20 dB. In the night period all the limits are reduced by 5 dB, and thus the limits for the equivalent level of the tones are 0 dB, 5 dB, and 10 dB.
- If the noise is not tonal, the limit for the A-weighted equivalent level (10 Hz – 80 Hz) is 35 dB during daytime and 25 dB during the night. The A-weighted level is calculated by adding the A-weighting corrections to only those levels that are above the hearing threshold. As opposed to the Danish method, the contributions from levels below the threshold are disregarded. The corresponding levels for the maximum levels are 45 dB and 35 dB.

The hearing threshold, L_{HS}, used in the German assessment method is given in Table 5:

Frequency, Hz	**8**	**10**	**12.5**	**16**	**20**	**25**	**31.5**	**40**	**50**	**63**	**80**	**100**
Hearing Threshold (L_{HS}), in Sound Pressure Level, dB	103	95	87	79	71	63	55.5	48	40.5	33.5	28	23.5

Table 5. One-third octave values for the hearing threshold given in the German standard DIN 45 680

4.3 Swedish method

The recommendations from the Swedish National Board of Health and Welfare (SOSFS 1996:7) [16] give guidance to an assessment of whether noise under different conditions may have health effects. The recommendation comprise a criteria curve of recommended maximum levels of low frequency noise in rooms used for living (see Table 6).

The curve covers the frequency range 31.5 Hz – 200 Hz and applies to the equivalent level of the noise. A measurement method is specified and is described in a report from the Swedish Testing Institute [17]. If the noise level exceeds the criteria curve in any 1/3-octave band, the health and environmental authorities may characterise the noise as a sanitary nuisance.

Frequency, Hz	31.5	40	50	63	80	100	125	160	200
Criteria curve, dB	56	49	43	41.5	40	38	36	34	32

Table 6. Values of the criteria curve for low frequency noise, applied to the equivalent noise level per 1/3.octave band in the Swedish assessment method.

4.4 Polish Method

Also the Polish method applies a threshold curve. This is defined in the frequency range 10 Hz – 250 Hz, and corresponds to 1/3-octave levels each giving an A-weighted level of 10 dB (i.e. 10 dB above the inverse A-weighting correction). The criterion curve is called LA10 [18].

The noise is considered annoying if both of these conditions are met:

- The spectrum of the noise exceeds the criterion curve LA10 in one or more 1/3-octave bands, and
- The spectrum of the noise exceeds the spectrum of the background noise.

It is mentioned in [18] that usually the background noise is somewhat higher than the criterion curve at the highest frequencies, above 100 Hz.

4.5 Dutch Proposed Method, Sloven

This proposed method [19] is intended for use along with granting environmental permission to industries and enterprises. Also this method uses a criterion curve. It is defined in the frequency range 10 Hz – 200 Hz. In the upper part of the frequency range the criterion curve agree well with the Swedish criterion curve. At the lowest frequencies, where the Swedish curve is not defined, it corresponds with the hearing threshold as specified in the German method.

It is expected that annoying low frequency noise will occur if the criterion curve is exceeded in one or more 1/3-octave bands.

4.6 Dutch Criterion for Audibility

This method is described in [20]. It is intended for use in cases where people complain about low frequency noise to decide if audible low frequency noise occurs. The aim of the method is not to verify whether the noise is annoying or not.

The method employs a hearing threshold based on the best 10 % of a non-selected population aged 50 – 60 years. The threshold curve is used in the frequency range 20 Hz – 100 Hz.

4.7 C-weighted Sound Pressure Level

In the German method the difference between the C-weighted and the A-weighted sound pressure level is used to determine if low frequency noise is present. Similar rules of thumb have regularly been mentioned in the literature.

4.8 Comparison of Criterion Curves

In Figure 4 the criterion curves from the different assessment methods are compared. The order is arranged after the criterion value at 100 Hz.

The curves Sloven, Swedish, and Polish are criterion curves directly aimed at assessing if the noise is annoying. The curves Sloven and Swedish differ only in the frequency range 50 Hz – 80 Hz, where the Swedish curve is clearly lower than Sloven. In the entire frequency range the Polish curve is lower than the two other curves. Here it must be remembered that the background noise is also part of the Polish criterion, which will often have a relieving influence on the criterion curve at high frequencies, but this part of the method is not considered here.

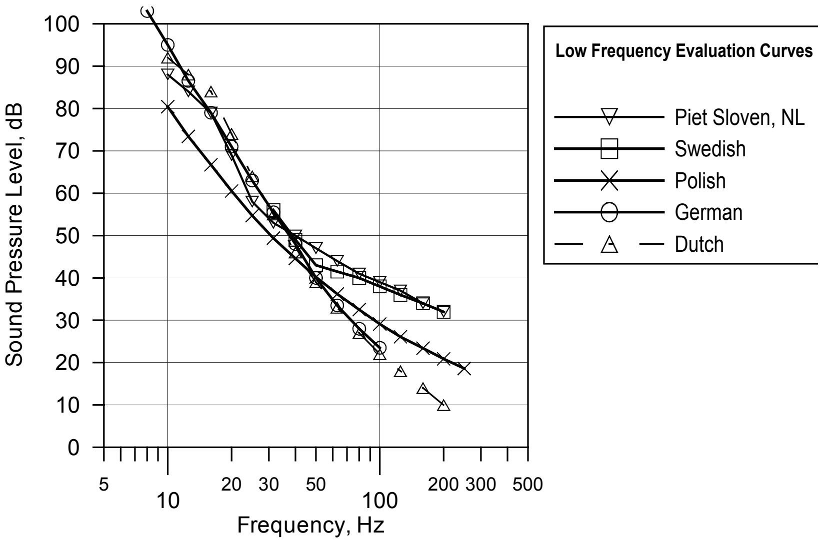


Figure 4. Comparison of criteria curves from the different assessment methods

The curve 'German' is a hearing threshold curve and is used as a criterion for tones in the noise. It is allowed for tones to exceed the curve by 5 dB during daytime, and higher a higher exceeding is allowed at higher frequencies. The curve 'Dutch' is used in cases with complaints in order to decide whether there is audible noise in the relevant frequency range. The curve is not used to determine if the noise is annoying. It can be seen that the German and the Dutch threshold curves are almost identical, and that they almost coincide with the curve 'Sloven' below 40 Hz.

The Polish criterion curve is clearly lower than the threshold curves in the frequency range below ca. 30 Hz.

5 Objective Analysis of the Noise Examples

5.1 One-third octave spectra

Based on the 1/3-octave analysis of the noise at the listening position in the listening room, a range of objective measurements have been made for each noise example. All the levels and spectra are given in Annex B. As an example, the spectrum of the traffic noise is shown in Figure 5 below. The spectrum is shown for the three nominal presentation levels: 35 dB(A), 27,5 dB(A) and 20 dB(A). The spectrum is a typical low frequency spectrum with a maximum around 20 Hz. The spectra of the other noise signals are shown in Annex B.

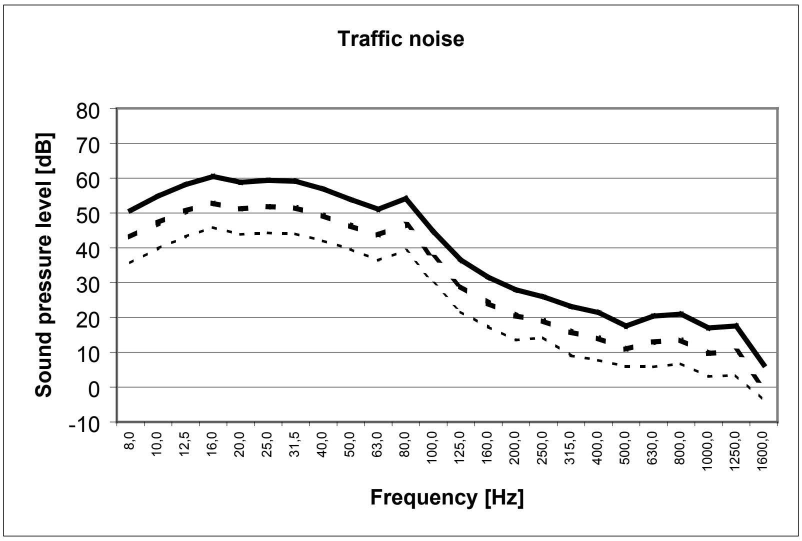


Figure 5. One-third-octave spectrum of the traffic noise shown for the three presentation levels: Solid line: 35 dB(A); thick, dashed line: 27,5 dB(A); thin, dashed line: 20 dB(A).

5.2 Comparison to criterion curves

In this section the different objective measures, resulting from the different assessment methods are discussed. The discussion is limited to the results from the presentations at the highest level 35 dB, since there are no systematic differences between these results and the results from the two presentations of the same noises at lower levels (apart from the change in level). The objective results are used as parameters in the comparison with the subjective evaluation, Chapter 6.

In Table 7 it is shown for each noise example how much each of the criterion curves is exceeded, and at which frequency the highest excess occurred. Also the Danish measure $L_{pA,LF}$ is given, and the German A-weighted level (of the frequency bands exceeding the hearing threshold)

	Traffic noise		Drop forge		Gas motor		Fast ferry	
	dB	***Hz***	**dB**	***Hz***	**dB**	***Hz***	**dB**	***Hz***
Sloven	13.1	*80*	15.7	*50*	19.8	*25*	14.3	*63*
Swedish	14.1	*80*	19.7	*50*	15.9	*50*	16.8	*63*
Polish	21.6	*80*	22.5	*50*	23.1	*25*	22.1	*63*
German	26.1	*80*	23.4	*80*	18.9	*50*	24.8	*63*
Dutch	27.1	*80*	24.4	*80*	19.9	*50*	25.3	*63*
Danish $L_{pA,LF}$	34.5	-	36.5	-	34.8	-	35.0	-
German A	33.5	-	35.9	-	34.7	-	34.5	-

	Steel factory		Generator		Cooling compressor		Discotheque	
	dB	***Hz***	**dB**	***Hz***	**dB**	***Hz***	**dB**	***Hz***
Sloven	13.0	*63*	17.3	*80*	12.0	*50*	9.0	*80*
Swedish	15.5	*63*	18.3	*80*	16.0	*50*	10.0	*80*
Polish	20.8	*63*	25.8	*80*	19.2	*100*	18.4	*125*
German	23.5	*63*	30.3	*80*	24.8	*100*	22.0	*80 (100)*
Dutch	24.0	*63*	31.3	*80*	26.3	*100*	23.5	*100*
Danish $L_{pA,LF}$	33.1	-	36.1	-	34.0	-	33.1	-
German A	32.8	-	36.0	-	29.4	-	28.9	-

Table 7. Comparison of the results from the different assessment methods, exceedence of the criterion curves (in dB) at frequency (Hz).

Comments to the individual noise examples:

- Traffic noise. The noise has no tones. All excess occur at 80 Hz
- Drop forge. The noise has a faint tone at 20 dB, below the threshold. Only the equivalent noise level is considered even though the noise has a pronounced impulsive character. All excess occur at 50 Hz apart from German (80 Hz) and Dutch (80 Hz)
- Gas motor. According to the German method, the noise has three tones, at 16, 25, and 50 Hz. The 16 Hz tone is below the threshold; excess occur at 25 Hz or at 50 Hz.
- Fast ferry. According to the German method, there are no tones in this noise. All excess occur at 63 Hz.
- Steel factory. No tones in this noise, all excess at 63 Hz
- Generator. This noise has a tone at 80 Hz, and all excess occur at this frequency
- Cooling compressor. It has tones at 50 Hz and 100 Hz, and the excess are found at these frequencies.
- Discotheque. The noise has no tones but has an impulsive character

It is generally seen that the Swedish and / or the Sloven method gives less excess than the German method, the Polish method, or the Dutch method (for audibility).

With a view to the Polish method, it has to be remembered that the background noise criterion, which normally will relieve the criterion curve at higher frequencies, has not been considered here. Furthermore the German criterion curve has not been used correctly in this comparison: firstly it only applies to tones in the noise, and secondly the German method specifically mentions a relief of the criterion curve in the 80 and 100 Hz bands which was not applied here. Finally, the Dutch method (for audibility) is not a method for assessing annoyance, so it is not surprising that it gives a stronger assessment than some of the other methods.

6 Comparison of Subjective Evaluation of Annoyance with Objective Measures

This Chapter describes the assessments made by the reference group and the relation between these assessments and various objective measures.

In every evaluation the test person was asked to assess the annoyance for two different situations: if the noise was heard in the day and the evening, and if the noise was heard at night. It has been found that there is a very close relation between these two assessments, as is illustrated in Figure 6. Generally the annoyance at night is slightly larger than annoyance in the day / evening.

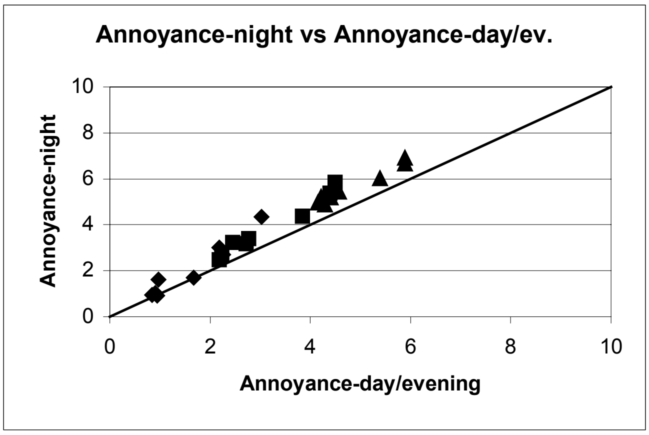


Figure 6. Comparison between the assessment made by the reference group of the annoyance in the day and evening period and the annoyance at night for the same noise examples.

The relation between the pair of evaluations can be described by the correlation coefficient, which is as high as 0.9885. This indicates that it is not necessary to perform a complete analysis on both the assessments, as they are not independent. In the following only the subjective assessment of the annoyance in the night period is considered.

6.1 Danish Method

In this section a thorough description of the analysis process is given. The same analysis method is applied to the other assessment methods, but for the other assessment methods only the results and some comments are given.

In Table 8 below all the noise examples are shown, the names are given in the first column. In the second column the excess of the Danish limit is shown. For the night period the limit value is $L_{pA,LF}$ = 20 dB, but since the drop forge as well as the discotheque are considered as impulsive noises, the limit for these noises is 15 dB. The third column gives the average assessment made by the test persons. Column 2 and 3 are from the presentation with a nominal level of 20 dB, the lowest presentation level. Column 4 and 5 are form the presentations with the middle presentation level, nominally 27.5 dB, and column 6 and 7 are from the loudest presentation with 35 dB nominal level.

Nominal level	20 dB		27,5 dB		35 dB	
Noise example	Excess	Subjective annoyance	Excess	Subjective annoyance	Excess	Subjective annoyance
Traffic noise	-0,3	1,6	7,0	3,4	14,5	5,2
Drop forge	6,9	4,3	14,2	5,9	21,5	6,9
Gas motor	-0,2	0,9	7,3	2,5	14,8	5,2
Fast ferry	0,1	0,9	7,5	3,2	15,0	5,4
Steel factory	-1,0	1,0	5,6	2,7	13,1	4,9
Generator	1,3	1,7	8,6	3,2	16,1	5,0
Cooling compressor	-0,4	2,7	6,5	4,4	14,0	6,0
Discotheque	3,7	3,0	10,7	5,4	18,1	6,7

Table 8. Subjective evaluation of the various noise examples shown together with the objective 'assessment' by use of the Danish method ($L_{pA,LF}$).

These results are also shown in Figure 7 below. The diamond shaped points are from the examples with the low level, the square points are from the intermediate level, and the triangular points are from the highest level.

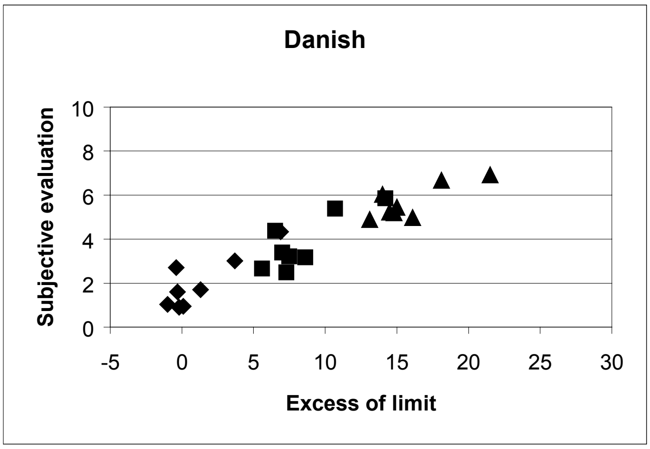


Figure 7. Illustration of the relation between the Danish assessment method and the subjective evaluation. Diamonds: low presentation level; squares: intermediate presentation level; triangles: high presentation level.

The Figure illustrates the subjective evaluation as a function of the excess of the Danish criteria. It is seen that a straight line (not shown in the figure) can represent the group of points. This line is found by linear regression (least squares method, in practise the function LINEST of an Excel spreadsheet is used). The regression line has the formula:

$$y = 1.61 + 0.26 * x$$

Where 'y' represents the (average) subjective evaluation made by the test subjects and 'x' represents the excess of the limit in dB.

The regression line does not explain how well it represents the group of points. Here we can use the residuals, which are the vertical distance between the points and the line. We also use the average of all y-values and the y-values determined by the regression method for each x-value. The determination coefficient or degree of explanation (r^2) is defined as:

$$r^2 = S_{Se} / (S_{Se} + S_{Sr})$$

where S_{Se} is the residual sum of squares, and S_{Sr} is the regression sum of squares. In practice r^2 is calculated by the Excel function LINEST.

If r^2 equals 1.00 there is a perfect linear relationship between the points. If r^2 is close to zero, the regression line cannot be used to explain the relation between x and y. In other words, the r^2 value indicates how well the points can be described by a straight line. The closer the value is to 1, the better the description.

The relation between x and y can also be described by the correlation coefficient, ρ. This is calculated as the ratio between the covariance of x- and y-values, and the product of the x-variance and the y-variance:

$$\rho = \text{covariance}(x,y) / (\sigma_x * \sigma_y)$$

The covariance is calculated as the deviation between the x-value and the x-average, multiplied by the deviation between the y-value and the y-average. The x-variance, σ_x, is the deviation between the x-value and the x-average squared. Similarly the y-variance, σ_y, is the deviation between the y-value and the y-average squared. The correlation coefficient is calculated by use of the Excel function CORREL. It explains the degree of relation between the x and the y values; it gives a coarse indication of the shape of the swarm of points. If ρ is close to 1 (or –1) the shape of the group of points has the shape of a 'cigar' around the regression line. If ρ is close to zero, the points lie in a diffuse cloud and there is no obvious relation between x and y.

There is an important difference between the degree of explanation and the correlation coefficient. The degree of explanation, r^2, assumes a functional relationship between y and x (e.g. the subjective evaluation is 'caused by' the noise). The correlation coefficient does not assume such causality and can be calculated from any two datasets. E.g. in chapter 7 a correlation coefficient is calculated for the annoyance evaluations made by the reference test subject group and the special test subject group. In this situation it is obvious that the ratings made by one of the groups are not caused by the ratings made by the other group.

The values calculated for the Danish method are summarized in the following table.

Slope	Intersection (x = 0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.26	1.61	0.88	0.94

In the next sessions these statistical calculations are made for each of the different assessment methods. In every analysis the objective metric is chosen as the x-parameter, and the subjective annoyance evaluation (which is the same in all cases) is the y-values. The same scaling is used in all Figures in order to make comparisons easier.

6.2 German A-level

Figure 8 shows the A-weighted level of the noise examples, calculated according to the German standard DIN 45 680. The level is calculated as the sum of the A-weighted levels of those 1/3-octave bands that exceed the hearing threshold. All the noise examples are used in this calculation, including those examples where the noise contains tones.

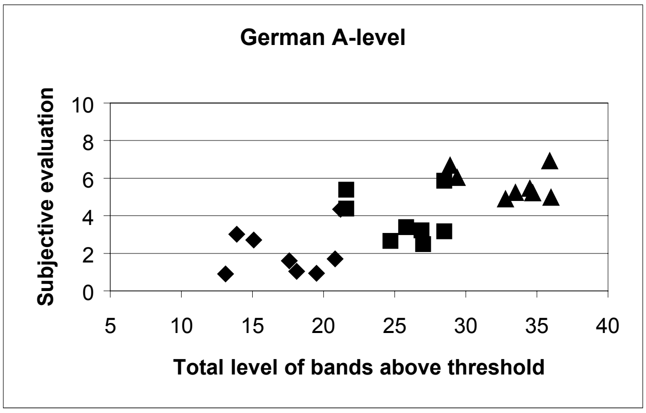


Figure 8. Illustration of the relation between the German assessment method using the A-weighted level and the subjective evaluation

It is seen that the points fall in two groups (two lines); the upper points are the noise examples from drop forge, discotheque, and compressor. The German method (in the present interpretation) obviously cannot give a sufficient assessment of impulsive noise (drop forge and discotheque). The degree of explanation is not too convincing, r^2 being 0.54, though ρ is 0.73.

Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.19	-0.98	0.54	0.73

6.3 German Tonal Method

Figure 9 shows the relation between the subjective evaluation and the tone level above the hearing threshold for those noises that contain tones. The ‘tone/no tone’ decision was made according to the German assessment method. In the figure only one point per noise is shown. This point corresponds to the tone with the greatest excess of the threshold.

The Gas Turbine has two tones above threshold at the highest presentation level (triangle) and one at the intermediate level (square). There is no tone above threshold at the lowest presentation level.

The Generator has one tone above threshold at the highest level (triangle), at the intermediate level (square) and at the lowest level (diamond). The tone is at 80 Hz and thus the level has been reduced by 5 dB according to the German assessment method.

The Cooling Compressor has two tones (50 Hz and 100 Hz) above threshold at the highest presentation level (triangle) and at the intermediate level (squares). As one of the tones is at 100 Hz the level for this tone must be reduced by 10 dB according to the German assessment method. This makes the other tone (at 50 Hz) the greatest. There is only one tone above threshold at the lowest presentation level (diamond).

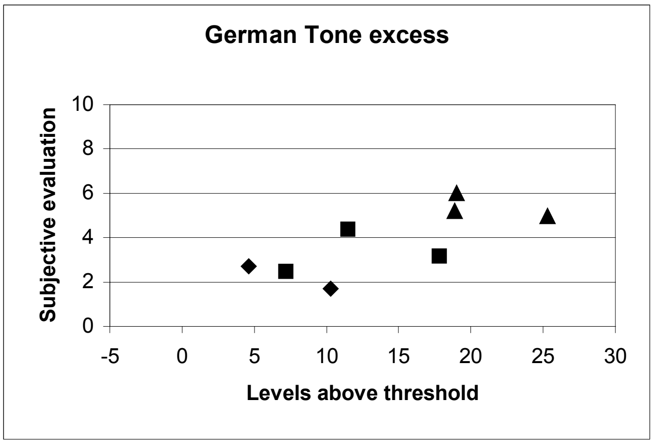


Figure 9. Illustration of the relation between the German assessment method for tonal noise and the subjective evaluation

The calculations gave the following results that are very similar to the results obtained from the German A-level method (see section 6.2).

Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.16	1.58	0.52	0.72

6.4 Swedish Method

This method uses a criterion curve that must not be exceeded in any 1/3-octave band. Figure 10 shows the subjective assessment as a function of the greatest excess.

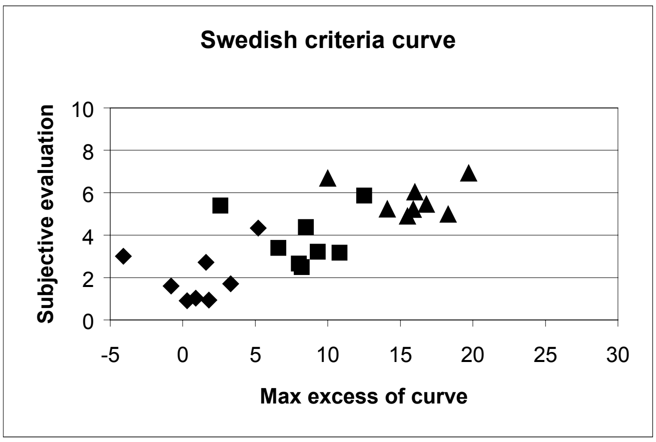


Figure 10. Illustration of the relation between the Swedish assessment method and the subjective evaluation

The result of the regression and the correlation analysis is

Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.22	2.10	0.57	0.76

The relation between the subjective assessment and the objective criteria is not too convincing.

It may be seen from Figure 10 that three points fall ‘to the left’ of the rest of the points. These three points are from the discotheque. Obviously this type of noise should have been assessed about 10 dB ‘higher’ for the points to fit into the rest of the points. The relative low values of r^2 and ρ are mainly caused by these points.

An analysis where the discotheque points have been discarded gave the following result

Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.25	1.34	0.81	0.90

When the discotheque points are removed, the slope of the regression line is almost the same and consequently the intersection will change. But more importantly it is seen that the degree of explanation, r^2, and the correlation coefficient, ρ, becomes much greater.

6.5 Polish Method

Figure 11 shows the excess of the Polish criterion curve, which is a curve of 1/3-octave band levels each of which corresponds to an A-weighted level of 10 dB. The other part of the Polish method, which deals with the excess of the background noise level, has not been considered.

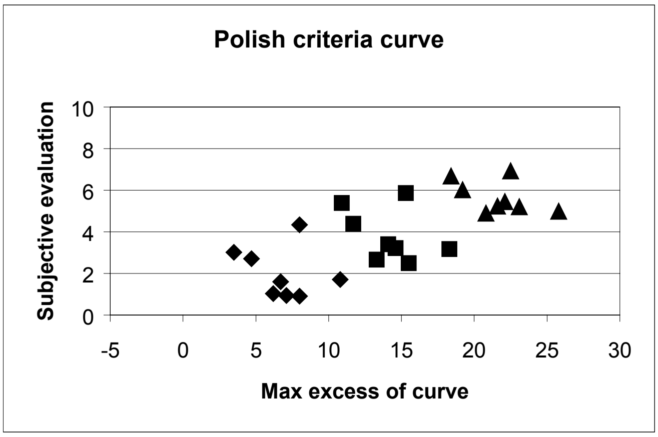


Figure 11. Illustration of the relation between the Polish assessment method and the subjective evaluation

It may be seen that the noise examples fall in two rather distinct groups, where (again) the discotheque, the drop forge and the cooling compressor are ‘to the left’ of the remaining points. They should have been assessed at least 5 dB ‘higher’ by the method to align with the remaining points. If one looks at the three groups of points from each nominal level separately (diamonds, squares and triangles), it appears as if the points in each group shows a tendency to a downwards slope to the right; that is, the subjective evaluation decreases as the max excess increases within each group.

The degree of explanation is not convincing, r^2 being 0.50.

Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.20	1.00	0.50	0.71

6.6 Dutch Proposal, Sloven

The proposed assessment for use with environmental approval of industries also employs a criterion curve. As is seen in Figure 12 the points spread a lot, and no clear picture can be seen. If one again focuses at the three groups of points from the same nominal level, it appears as if these points slope 'the wrong way' as it was seen with the Polish method. As an example (filled squares) it can be seen that when the excess increases from 1 dB to 12 dB, the subjective evaluation decreases from 5.5 to 2.5.

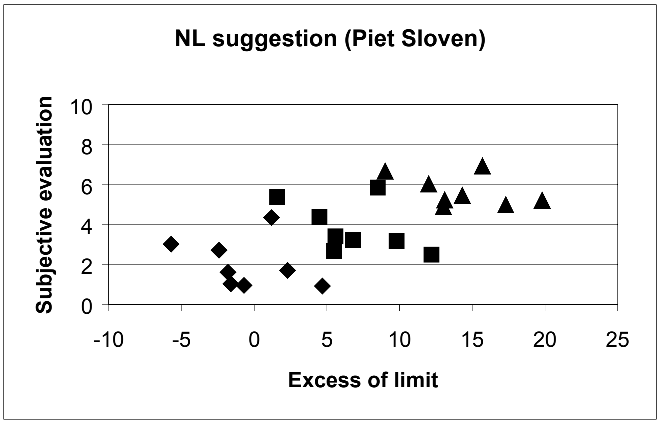


Figure 12. Illustration of the relation between the Dutch proposed assessment method by Sloven and the subjective evaluation

Both the degree of explanation and the correlation are rather poor.

Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.17	2.67	0.40	0.64

6.7 C-weighted Level

Figure 13 shows the relation between the C-weighted level of the noise examples and the subjective evaluation. It can be seen that the spread of the points is very large and the correlation is poor. Only the frequency range 10 Hz to 160 Hz is included in the calculation of the C-weighted level.

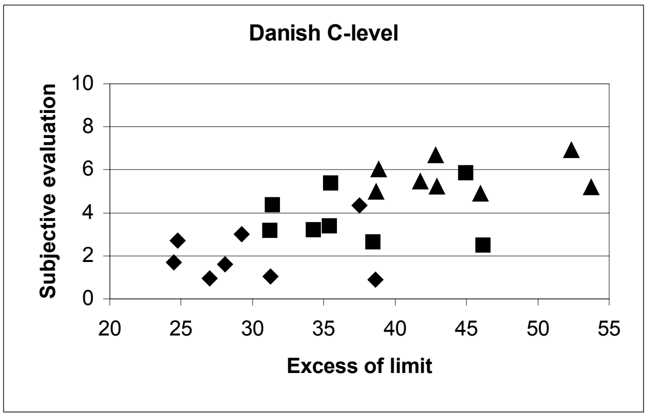


Figure 13. Illustration of the relation between the C-weighted sound pressure level and the subjective evaluation

Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ
0.15	-1.82	0.44	0.66

6.8 Overview of the Results

The results of the analysis of the relation between the subjective evaluation and the various objective measures are summarized in table 9 below.

Assessment method	Slope	Intersection (x=0)	Degree of explanation, r^2	Correlation coefficient, ρ	Annoyance = 1	Annoyance = 2
Danish	0.26	1.61	0.88	0.94	-2.3 dB	1.5 dB
German A-level	0.19	-0.98	0.54	0.73	10.4 dB	15.7 dB
German tonal	0.16	1.58	0.52	0.72	-3.6 dB	2.6 dB
Swedish	0.21	2.10	0.57	0.76	-5.2 dB	-0.5 dB
Polish	0.20	1.00	0.50	0.71	0.0 dB	5.0 dB
Sloven	0.17	2.67	0.40	0.64	-9.8 dB	-3.9 dB
C-level	0.15	-1.82	0.44	0.66	18.8 dB	25.5 dB

Table 9. Overview of the results from regression analysis of the relation between the subjective evaluations made by the reference group and the different assessment methods.

In addition a set of calculations was made on the results of the statistical analysis with the purpose of comparing the different assessment methods to each other. The rightmost two columns shows the excess of the criteria from each of the methods that would result from the same subjective annoyance, respectively annoyance = 1 (on the chosen scale from 0 to 10) and annoyance = 2. These results can be used in a discussion on ‘calibration’ of the assessment methods.

It can be seen from the results that the assessment method with far the best correlation between the subjective and the objective assessment is the Danish method, using the A-weighted level in the frequency range from 10 Hz to 160 Hz. The degree of explanation is high, and there is no obvious indication in the results obtained here of ways to further improve the method. The second best method is either the Swedish method, based on a criterion curve, or the German method using the A-weighted level.

It is seen from the rightmost column that a noise that would exceed the Danish criterion by 1.5 dB will give rise to a subjective annoyance of 2 (on a scale from 0 to 10). The same noise would be 0.5 dB below the Swedish criterion and would give rise to a German A-weighted level as low as 15.7 dB. Thus it can be stated in broad terms, that when the assessment methods are compared in this way with representative noise examples and 'calibrated' to the same subjective annoyance

- the Swedish method is a little less restrictive than the Danish method (that is, the Danish limit value - for the evening and night period - corresponds to a lower annoyance than the Swedish method)
- the German method for broadband noise - using a limit value of 25 dB for the night period - is considerably less restrictive than the Danish method. On the other hand, the German method for tonal noise is a little more strict than the Danish method. Thus the German method change significantly with the appearance of tones in the noise
- the Polish method appears more restrictive than the Danish method, though in this comparison the 'release' of the Polish criterion at higher frequencies that results from considering the background noise has not been implemented, and
- the Dutch method proposed by Sloven is somewhat less restrictive than the Danish method.

These results cannot be used to decide which of the methods that are unnecessarily strict or too relaxed, since the results are based on a limited number of noise examples only.

7 Results obtained with the Special Group

Apart from the reference group of eighteen young normal hearing test persons, also a special group of four persons that have reported annoyance due to low frequency noise in their homes was included in the listening tests. The results obtained with this special group of subjects have not been used in the analysis in the previous Chapter. Instead they are analysed and discussed in the present Chapter.

7.1 Comparison of the two groups of test persons

In the first series of analysis, the evaluations made by the two groups of test persons are compared. This is illustrated in the following four Figures (14-17), showing how the reference group and the special group evaluates the same noise examples with respect to:

- Loudness
- Annoyance, if heard at day or in the evening
- Annoyance, if heard at night
- Annoying? (yes / no)

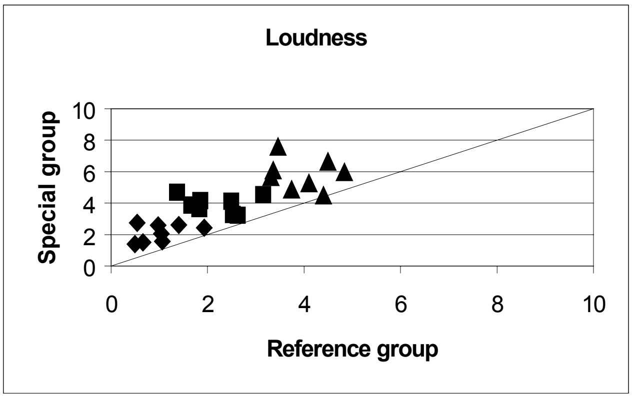


Figure 14. Illustration of the relation between the assessments of loudness made by the reference group (young people with normal hearing) and the special subject group.

There is a good correlation between the assessments of loudness between the two groups, and the correlation coefficient is calculated to be 0.82. The special group generally finds the noise examples somewhat louder than the reference group does. The points are rather close to a line that would be offset from but parallel to the line indicated in Figure 14 (showing a fictive 1:1 relationship).

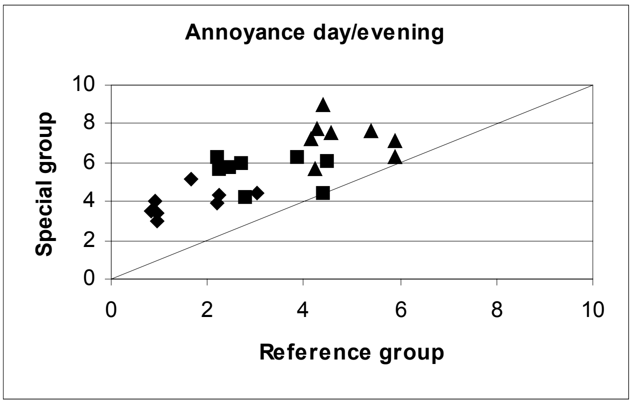


Figure 15. Illustration of the relation between the assessments of annoyance during day/evening made by the reference group (young people with normal hearing) and the special group.

The relation between the assessments of annoyance (day / evening) of the two groups, Figure 15, is less clear. The correlation coefficient drops to 0.75, and especially the group of points from the highest nominal level (triangles) shows a considerable scatter. In every case the special group finds the noise example more annoying than the reference group does.

On the average the special group rate the annoyance at day/evening about 2 to 3 scale units higher that the reference group. An increase in the rating of 2 to 3 units corresponds roughly to an increase in level of about 10 dB.

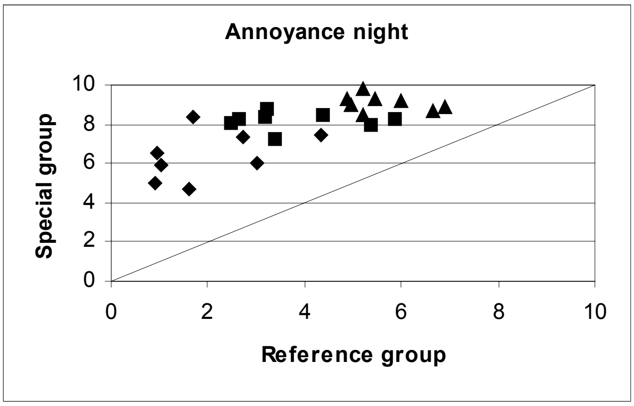


Figure 16. Illustration of the relation between the assessments of annoyance at night made by the reference group (young people with normal hearing) and the special group.

For the assessment of annoyance at night, Figure 16, the picture is shifted. The special group finds the noises much more annoying at night than at day (or evening), and the difference between the assessments of the two groups increases significantly. Figure 16 shows a ‘saturation’ phenomenon, that is, one or more of the test persons in the special group uses the maximum indication of the annoyance scale, and this will break the linear relation between the points. The correlation coefficient drops to 0.73.

On the average the special group rate the annoyance at night about 4 to 5 scale units higher that the reference group. Such an increase in the rating of 4 to 5 units corresponds roughly to an increase in the level of about 17 dB.

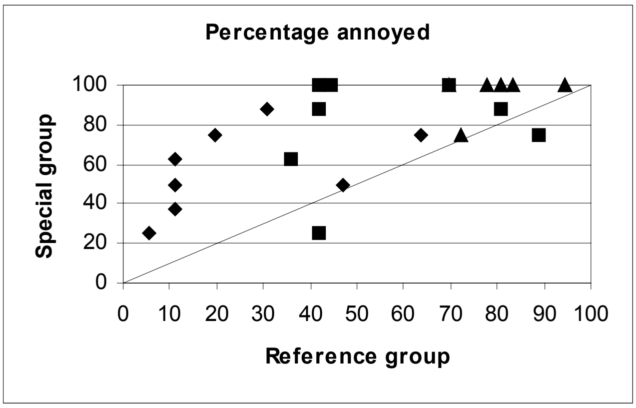


Figure 17. Illustration of the relation between the assessments of annoyance (yes/no) made by the reference group (young people with normal hearing) and the special group.

Finally the assessments ‘annoying yes / no’ of the two groups are shown in Figure 17. The scales show how many percent of the group that have marked the noise as annoying. Here the saturation is obvious, where all (four) persons in the special group have marked several noise examples as annoying. The Yes/No parameter has not been used in the previous analysis of the data from the reference group.

7.2 Comparison of the assessments of loudness and annoyance by the special group

The relation between the assessments by the special group of loudness and annoyance (day / evening) of the same noise examples is illustrated in Figure 18. It is seen that the relation is almost linear, and the correlation coefficient is as high as 0.96. The loudness scalings are less than the annoyance scalings and thus the noises are perceived more annoying than loud.

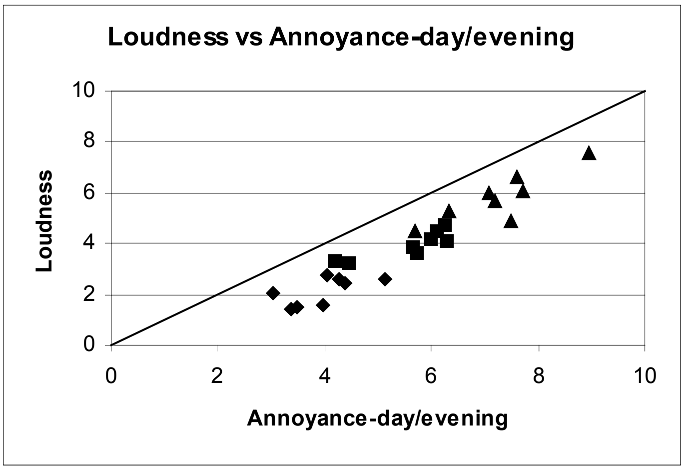


Figure 18. Relation between the assessment made by the special group of annoyance (day/evening) and of loudness of the same noise examples.

The relation between the assessments of annoyance at day / evening and at night is illustrated in Figure 19. The non-linear relation due to saturation is clearly seen.

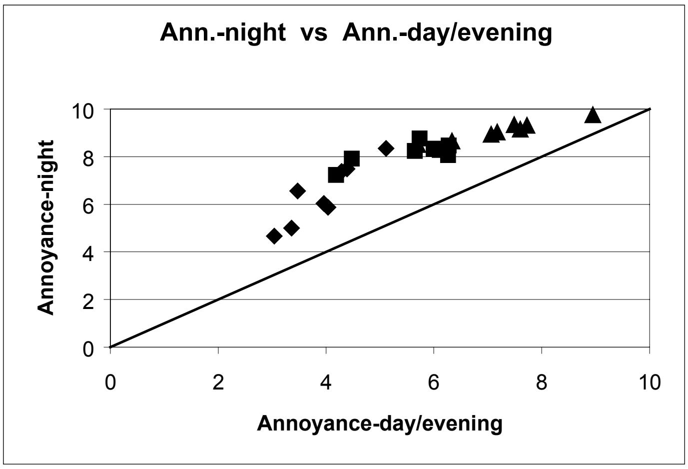


Figure 19. Relation between the assessment made by the special group of annoyance (night) and of annoyance (day) of the same noise examples.

This indicates the need (in this case) for a longer response line – or rather the need for a stronger assessment than ‘very annoying’. The group of points from the middle level (filled squares) is evaluated 2 – 3 ‘units’ more annoying when they occur at night than at daytime, but the points from the loudest presentation are only indicated 1 – 2 ‘units’ more annoying.

For the special group the annoyance generally increase by two ‘units’ from day to night corresponding roughly to a 10 dB change in the noise level. This can be compared to the results from the reference group, Figure 6 (in the beginning of Chapter 6) where the annoyance at night generally was rated about one ‘unit’ higher than at day – at all presentation levels. Such a one-unit change in the rating corresponds approximately to a 5 dB change in the noise level and supports thus the 5 dB penalty in the noise limits at night.

7.3 Comparison of subjective evaluation of annoyance with objective measures

In the same way as in Chapter 6 the subjective evaluation of annoyance in the night period was compared to a number of objective measures, but in this case only the subjective evaluations from the special group were used. Table 10 below shows the subjective evaluation by the special group, and Table 11 shows the results of the statistical analysis made for each of the different objective assessment methods.

Nominal presentation level	20 dB	27,5 dB	35 dB
Noise example	Subjective annoyance at night	Subjective annoyance at night	Subjective annoyance at night
Traffic noise	4.7	7.2	8.5
Drop forge	7.5	8.3	8.9
Gas motor	5.0	8.1	9.8
Fast ferry	6.6	8.8	9.3
Steel factory	5.9	8.2	9.3
Generator	8.4	8.3	9.0
Cooling compressor	7.4	8.5	9.1
Discotheque	6.0	7.9	8.6

Table 10 Results of the subjective evaluation of annoyance during the night time made by the special group.

The general result from Table 10 is – again – that the special group assesses the noise examples as much more annoying than the reference group does. This is seen more clearly if Table 10 (for the special group) is compared to Table 8 in Section 6.1 (for the reference group). It is also seen that the subjectively assessed annoyance increases with increasing level (apart from the noise from the generator, which apparently is equally annoying at both a nominal level of 20 dB and at 27.5 dB). The annoyance found by the special group at a nominal level of 20 dB corresponds almost to the annoyance reported by the reference group at a level of 35 dB.

A very interesting result that is obtained by comparing Table 8 and 10 is that it is *not* the same noise examples that are evaluated as most annoying by the two groups. The reference group clearly found the drop forge, the discotheque, and the cooling compressor the most annoying. This rank would hold at any of the three presentation levels. In contrast, the special group found the generator the most annoying (at the lowest presentation level) and the discotheque as one of the lesser annoying sounds.

Assessment method	Slope	Intersection (x=0)	Degree of explanation, r2	Correlation coefficient, ρ
Danish	0,16	6,52	0,60	0,78
German A-level	0,16	3,83	0,69	0,83
German tonal	0,05	7,99	0,39	0,54
Swedish	0,17	6,44	0,72	0,85
Polish	0,17	5,47	0,66	0,81
Sloven	0,15	6,84	0,59	0,77
C-level	0,09	4,40	0,31	0,55

Table 11. Summary of results of regression analysis of the relation between the assessments made by the special group and the different assessment methods.

It can be seen from Table 11 that none of the assessment methods give any particularly successful correlation to the subjective assessment made by the special group. Two examples are illustrated below in Figure 20 and 21, the Swedish and the Danish method. The groups of points from the intermediate and the highest presentation level both line up reasonably well with a slightly sloping line in the upper part of the Figures, while the group of points from the low presentation level appears very different in the two Figures. In Figure 20 showing the Swedish method these points have a curved tail-like appearance, while they in Figure 21 showing the Danish method appear as a diffuse cloud.

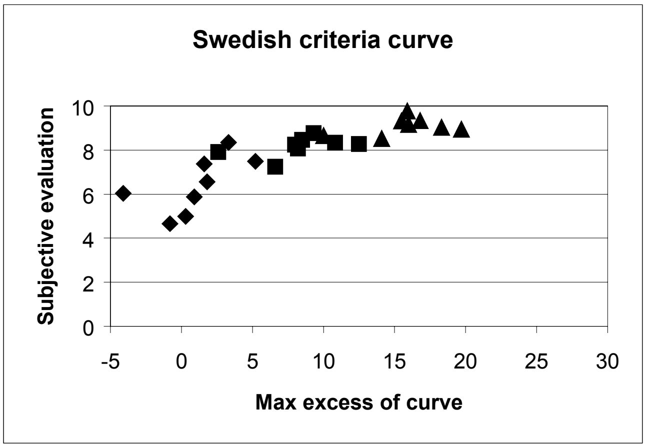


Figure 20. Illustration of the relation between the Swedish assessment method and the subjective evaluation made by the special group

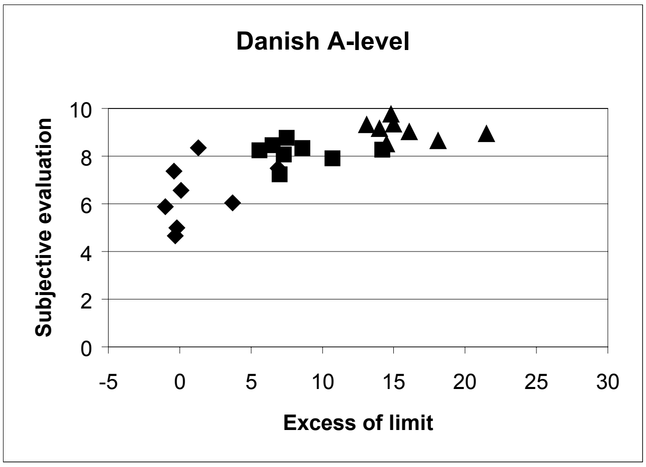


Figure 21. Illustration of the relation between the Danish assessment method and the subjective evaluation made by the special group

The other assessment methods show results without any particular trend like it is seen with the Danish assessment method. Obviously there is no strong connection between the subjective assessment made by the special group and the objective results found by the objective measuring methods.

Since only four persons have made these assessments, it was evaluated that no more effort should be paid to optimise an objective assessment method to fit the special group's evaluations. However, the results obviously give rise to a number of questions about how low frequency noise in the environment is experienced and how it can be assessed.

7.4 Discussion of results from the special group

The various scaling results show clearly that the special group made the annoyance evaluations differently from the reference group. The overall scaling value (averaged over all *annoyance* evaluations, presentation levels and noises) was 3,5 for the reference group and 6,7 for the special group. In other words, the special group evaluated the noises to be almost double as annoying as the reference group did.

This may also be illustrated by ordering the noises from the most annoying to the least annoying. This is done in Table 12.

Order, reference gr.	**Average Scaling**	**Order, special gr.**	**Average Scaling**
Drop forge	5,1	Generator	7,3
Discotheque	4,6	Cooling compressor	7,2
Cooling compressor	4,1	Drop forge	7,0
Generator	3,1	Gas turbine	6,9
Traffic noise	3,0	Fast ferry	6,9
Fast ferry	2,9	Steel factory	6,8
Steel factory	2,7	Discotheque	6,2
Gas turbine	2,7	Traffic noise	5,6
Grand average	*3,5*	*Grand average*	*6,7*

Table 12. Ordering of the noises from most annoying to least annoying for the two groups of test subjects.

The reference group has the Drop forge and the Discotheque on top of the list. These two noises are heard as ‘impulsive’ and thus a 5 dB penalty is added to the calculated level in the Danish evaluation method. The Generator and the Traffic noise are in the middle of the list and the Gas turbine is evaluated as the least annoying.

For the special group the Generator is on top of the list whereas the Discotheque and the Traffic noise are evaluated as the least annoying. It is interesting that Traffic noise gives the lowest overall scaling. The value 5,6 is well below the next one (Discotheque) at 6,2. The order of the noise signals could indicate that the special group put more attention to those noises, which resemble the typical low frequency noises that they are complaining about.

8 General discussion

8.1 The Experimental Method

The present investigation has been performed as a typical laboratory experiment in contrast to a field investigation. The advantage of a laboratory experiment is that it is possible to control almost all the experimental conditions (noises, levels, duration, presentation sequence, test subjects, etc). The disadvantage is that the presentations may not be realistic enough.

The test subjects made the scaling by a mark on a 10 cm long horizontal line. The results from the special group of subjects showed in some cases a saturation effect that made the results more difficult to interpret. A way to alleviate the saturation effect could be to change the layout of the response line. An example of one of the annoyance scales was

not annoying   very annoying

but the saturation effect might have been less pronounced if the scale was made like this:

not annoying   very annoying

Another way of reducing the saturation effect could be to exchange the word ‘very’ with a stronger adjective like e.g. ‘extremely’.

The response sheet contained all the questions on one sheet. It has been argued that the different questions should have been on separate sheets in order to avoid a smudging effect from one question to the next. Also the wording of the questions might be revised even though no test subjects reported any difficulties with the questions.

The amount of subjects (18 in the reference group; 4 in the special group) could be increased in order to obtain more certainty in the results. For the reference group it is believed that an increase in the amount of test subjects would not change the general results dramatically. For the special group, an increase in the amount of test subjects would certainly improve the validity of the findings as group results. On the other hand it may be difficult to handle the persons with low frequency problems as a homogeneous group of subjects. The problems they report are very different and thus it might be more relevant to handle the results from these test persons individually. Such an individual analysis has not been done in the present report.

The noises constitute a reasonable broad selection of low frequency sounds. The noises were selected to represent typical low frequency noise known to produce community claims. In retrospect it would have been an improvement to include more noises with an impulsive character in order to better 'test' the impulse penalty in the Danish method. All noises had clearly a low frequency character partly because of an outdoor-to-indoor filtering of the noises recorded out of doors. Traffic noise was included in order to serve as a reference noise, but due to the outdoor-to-indoor filtering the traffic noise was converted into another low frequency noise. In possible future research a real reference noise should be included.

8.2 Low Frequency Hearing Thresholds

The low frequency pure-tone hearing threshold was measured for both subject groups. The result showed that the special group was less sensitive to low frequency sounds than the reference group. Although this is an interesting result, it might be more informative to measure the loudness growth curve for the test subjects. It is believed that the loudness growth curve would be a much better predictor for annoyance than just the hearing threshold. Measurement of loudness growth is very time consuming and is certainly not straightforward at low frequencies.

8.3 Criteria and Measurement Methods

The results from the two groups seem to be different. All the subjects in the reference group evaluate the noises in almost the same way whereas more variance is seen in the special group. This could be caused by a bias in the special group who tend to put more emphasis on the type of noises that they complain about.

There is a good agreement between the annoyance evaluations from the reference group and the Danish measurement/calculation method (including the impulse noise penalty). The same good agreement is not found for the special group. This raises a question about the aim or objectives of a measuring method. Shall such an evaluation method be made for the average person (the general population) or shall a method be made with special emphasis on the persons who react more pronounced to low frequency noise?

The criteria and evaluation methods used in this investigation are all based on some kind of measurement of the noise level. For the reference group there is a clear connection between the noise level and the experienced annoyance and thus it makes sense to use such criteria and evaluation methods. For the special group this connection between level and annoyance is less clear and thus an evaluation method based on noise level measurements may be of little value for this subject group.

8.4 Statistical analysis

In the statistical analysis it was seen that the data deviated somewhat from a normal distribution partly caused by the saturation effect from the fixed endpoints of the scale. Despite this deviation in the distribution of the data the statistical analysis showed the expected effects and thus no attempt was made to correct for the saturation effect in the data.

For the special group of subjects the saturation effect was pronounced for the annoyance ratings at night. In this case a simple linear regression may not be suited for a description of the data. A logarithmic or polynomial model would probably give a better fit to the data. Due to the relatively small amount of data in this subject group, no attempt has been made to do this.

8.5 Noise limits, Criteria curves and equal loudness level contours

In chapter 5 and in chapter 6 a comparison with the noise limits and the criteria curves was made and as input data for the analysis the maximum excess of a criterion curve or the excess of a noise limit was used. It was decided to do it in this way because this is the way the criteria curves and the noise limits are used in practise. For the use of noise limits there is no problem because the level of the noise is calculated according to some rule and compared to the limit. The calculation of the level is based on the low frequency spectrum of the noise within a certain frequency range.

For the criteria curves, on the other hand, the procedure may constitute a problem as only a single frequency band of the noise is used in the comparison and not the whole spectrum. Only the band where the maximum excess occur is taken into account and the excess at other frequency bands are neglected. From Figure 4 (in chapter 4) it is seen that the course of the different criteria curves differ somewhat above 40 Hz and this means that the various criteria curves will give very different results if the excess occur in this frequency range. It also means that the excess decision will be very dependent on the inherent measurement uncertainty in the measurement of the spectrum. The calculation of a level – based on a spectrum – is much less sensitive to the measurement uncertainty as the uncertainties are ‘averaged’ in the calculation process.

The measurement uncertainty is inversely proportional to the bandwidth of the analysing filter and also inversely proportional to the duration of the measurement (i.e. the integration time). This means that a one-third-octave analysis of a low frequency noise must be extended over a long period of time in order to keep the uncertainty below a certain limit. It is common practise to require that the standard deviation of repeated measurements shall be less then 0,2 dB. This corresponds to an integration time (in seconds) greater than 471/B where B is the bandwidth in Hertz of the analysing filter. For the one-third-octave filter at 10 Hz this means an integration time of almost five minutes. At 40 Hz a one-minute integration time is necessary and at 1000 Hz two seconds are needed. The noise signal shall be stable over this period of time but this is not always the case in practise.

Uncritical use of criteria curves may be misleading. Some of the curves (e.g. the German and the Dutch) are hearing threshold curves and can therefore only be used to predict whether a nose is audible or not. If the curve is exceeded, the amount of excess cannot predict neither the loudness of the noise or the annoyance. This will depend on the shape of the spectrum of the noise.

The use of weighting functions (such as G- and A-weighting) will not automatically give a loudness or annoyance measure. In the conventional audible frequency range it is well known that neither the loudness level contours nor the A-weighting can predict the loudness of *complex* sounds. Loudness can only be predicted by the loudness level contours when a pure tone (no harmonics) is heard in isolation. The A-weighting resemble the hearing sensitivity at low levels but should not be used for loudness ratings.

It is believed that loudness is a major component of annoyance. Loudness is related to the level and the spectrum of the noise. Annoyance is therefore also dependent on level and spectrum but annoyance is also influenced by (or dependent on) many other factors and these factors cannot be described by physical measurements of the noise.

9 Conclusion

A laboratory investigation of the annoyance of low frequency noises has been performed. Eighteen normal hearing test subjects listened to eight different noises and evaluated the loudness, the annoyance at day/evening and the annoyance at night. All noises had a considerable low frequency content. The evaluations were compared to the noise limits and criteria curves for low frequency noise used in the European countries.

As a pilot project, a group of four persons – who were known to experience problems with low frequency noises – participated in the same listening test.

The results show that the Danish measuring method describes the subjectively experienced annoyance better than the measuring methods used in other countries. This result relies on the 5 dB impulse noise penalty included in the Danish method. The decision about whether or not a 5 dB penalty shall be applied to a specific noise is based on a purely subjective judgment and therefore the Danish method could be improved at this point.

The Swedish method is almost as good as the Danish method if the (impulsive) discotheque sound is omitted from the analysis. The Swedish method is based on a specified criterion curve (in contrast to the Danish noise level calculation) and as such more sensitive to random measurement uncertainties.

Because the traffic noise was outdoor-to-indoor filtered, the traffic noise did not serve as a reference noise as originally intended. This also prevented a comparison of the present findings to ordinary traffic noise annoyance investigations from the literature.

An almost perfect correlation was found between the annoyance at day/evening and the annoyance at night. The annoyance at night is slightly lager than the annoyance at day/evening. The difference in the annoyance ratings between day and night corresponds to a level change of about 5 dB.

The low frequency hearing threshold of the four special test subjects was not found to be better than the hearing threshold of the ordinary test subjects.

The annoyance evaluations made by the four special test subjects were clearly different from the evaluations made by the ordinary test subjects. The ratings were systematically higher. Especially at night the annoyance was rated as close to maximum and thus not dependent on the level of the noise. The four special test subjects were not annoyed by the impulsive noises to the same degree as the ordinary test subjects were.

The Danish evaluation method seems to be suitable for predicting the annoyance experienced by the general population. It seems to be is less suitable for predicting the annoyance experienced by the special subject group where the influence from the noise level is less pronounced and where individual reactions to the quality of the sounds becomes more important.

10 References

1. “Low frequency noise, infrasound and vibration in the environment” (In Danish). Information from the Danish Environmental Protection Agency no. 9/1997

2. Mortensen, F.R and Poulsen, T.: Annoyance of low frequency noise and traffic noise. Proc., 9th Int. Meeting on Low Frequency Noise and Vibration, pp 125-128, Aalborg, Dep. of Acoustics, Aalborg University, 2000. ISBN 87-90834-06-2

3. ISO 7196/1993 “Acoustics - Frequency weighting characteristic for infrasound measurements”. International Organization for Standardization. Geneva, Switzerland.

4. H. Møller, J. Andresen: ”Loudness of Pure Tones at Low and Infrasonic Frequencies”; Journal of Low frequency Noise and Vibration 3, p. 78 - 87 (1984)

5. “Environmental noise from industry” (In Danish) Guidelines from Danish Environmental Protection Agency no. 5/1984

6. IEC 268-13 (1985): Sound system equipment. Part 13: Listening tests on loudspeakers.

7. Mortensen, F.R. (1999) Subjective evaluation of noise from neighbours with focus on low frequencies. Main project. Publication no. 53, Department of Acoustic Technology, Technical University of Denmark, DK-2800 Lyngby. ISSN 1395-5985

8. Støjprojektet Lydisolering, Tekniske Løsninger. DSB, Januar 1987. ISBN 87-982432-1-7

9. Wyle Research Report WR 83-26. Preliminary Evaluation of low Frequency Noise and Vibration. Reduction Retrofit Concepts for Wood Frame Structures.

10. “Assessment of low frequency noise from ferries” (In Danish) Working Report no. 10/1997 from the Danish Environmental Protection Agency

11. Han, L. A and Poulsen, T.: (1998): Equivalent threshold Sound Pressure Levels (ETSPL) for Sennheiser HDA 200 earphone and Etymotic Research ER-2 insert earphone in the frequency range 125 Hz to 16 kHz. Scand. Audiol. vol. 27, p. 105-112

12. ISO 389-8: (2001) "Acoustics - Reference zero for the calibration of audiometric equipment - Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones" International Organization for Standardization. Geneva, Switzerland.

13. Buus, S., Florentine, M. and Poulsen, T.: (1997): Temporal integration of loudness, loudness discrimination, and the form of the loudness function. J. Acoust. Soc. Am., vol 101, p. 669-680

14. ISO 389-7 (1996): Acoustics–Reference zero for the calibration of audiometric equipment– Part 7: Reference threshold of hearing under free-field and diffuse-field listening conditions. International Organization for Standardization. Geneva, Switzerland.

15. Deutsche Norm DIN 45680 (1997) “Messungen und Bewertung tieffrequenter Geräuschimmissionen in der Nachtbarschaft” - Beiblatt 1 “Hinweise zur Beurteilung bei gewerblichen Anlagen”

16. SOSFS 1996:7/E: Indoor Noise and High Sound Levels, Socialstyrelsen, Sverige, 15 Maj 1996

17. ”Recommendation for measurement of sound levels in rooms with low frequencies” (In Swedish) SP-INFO 1996:17

18. Mirowska, M.: Evaluation of Low Frequency Noise in Dwellings. New Polish Recommendation. Proc., 9th Int. Meeting on Low Frequency Noise and Vibration, pp 117-124, Aalborg, Dep. of Acoustics, Aalborg University, 2000. ISBN 87-90834-06-2

19. Sloven, P.: Structured approach of lfn-complaints in the Rotterdam region. Proc., 9th Int. Meeting on Low Frequency Noise and Vibration, pp 167-176, Aalborg, Dep. of Acoustics, Aalborg University, 2000. ISBN 87-90834-06-2

20. Berg, G.P and Passchier-Vermeer, W.: Assessment of low frequency noise complaints. Proceedings Inter-Noise 99, p. 1993-1996, December 1999.

21. “Low frequency noise in the environment. Comparison of practise internationally” (In Danish) Working report from Danish Environmental Protection Agency - to be published in 2001

-ooOoo-

1 Data from the Subjective Evaluations

1.1 Subjective Data from the Reference Group

The following tables show the evaluation of the various noises averaged over the eighteen test subjects in the reference group. Each test subject indicated his/her response by setting a mark on a scale consisting of a 10 cm long line.

An example of one of the annoyance scales is given here:

not annoying very annoying

Average rating of loudness

Loudness	**Presentation level, dB(A)**		
Ref. Group	**20**	**27,5**	**35**
Traffic noise	1,0	2,5	4,4
Drop forge	1,9	3,2	4,8
Gas turbine	0,5	1,4	3,5
Fast ferry	0,7	1,8	3,7
Steel factory	0,5	1,7	3,4
Generator	1,0	1,9	3,3
Cooling compr.	1,4	2,5	4,5
Discotheque	1,1	2,6	4,1
Avg. loudness	*1,0*	*2,2*	*4,0*

Average rating of annoyance day/evening

Annoyance Day/evening	**Presentation level, dB(A)**		
Ref. Group	**20**	**27,5**	**35**
Traffic noise	1,0	2,8	4,2
Drop forge	3,0	4,5	5,9
Gas turbine	1,0	2,2	4,4
Fast ferry	0,8	2,5	4,6
Steel factory	0,9	2,2	4,3
Generator	1,7	2,7	4,1
Cooling compr.	2,3	3,8	5,4
Discotheque	2,2	4,4	5,9
Avg. day	*1,6*	*3,1*	*4,9*

Annex A

Average rating of annoyance at night

Annoyance Night	**Presentation level, dB(A)**		
Ref. Group	**20**	**27,5**	**35**
Traffic noise	1,6	3,4	5,2
Drop forge	4,3	5,9	6,9
Gas Turbine	0,9	2,5	5,2
Fast Ferry	0,9	3,2	5,4
Steel factory	1,0	2,7	4,9
Generator	1,7	3,2	5,0
Cooling compr.	2,7	4,4	6,0
Discotheque	3,0	5,4	6,7
Avg. night	*2,0*	*3,8*	*5,7*

Average of percentage of 'yes' responses

Percentage Annoyed	**Presentation level, dB(A)**		
Ref. Group	**20**	**27,5**	**35**
Traffic noise	5,6	41,7	72,2
Drop forge	63,9	88,9	94,4
Gas turbine	11,1	36,1	77,8
Fast ferry	11,1	44,4	80,6
Steel factory	11,1	41,7	69,4
Generator	19,4	41,7	80,6
Cooling compr.	30,6	69,4	83,3
Discotheque	47,2	80,6	94,4
Avg.	*25,0*	*55,6*	*81,6*

1.2 Statistical analysis of the subjective data from the reference group

All data (i.e. subjective evaluation, sound example number, nominal presentation level, measured A-weighted level (dB(A)), A-weighted level in the frequency range 10 Hz – 160 Hz ($L_{pA,LF}$), repetition number, gender and age) were used as input to a statistical analysis program (Statgraphics 4.0). The analysis was made for each parameter separately.

The table below shows the significance levels for the different factors. If the number is less than 0.05, this factor has a significant effect on the evaluation on a 95% level or above (this means less than 5 % probability for drawing a wrong conclusion). If the number is above 0.05 it cannot be proved that this factor has a significant effect upon the relevant evaluation.

	Loudness	Annoyance day	Annoyance night	Annoying ? (Y/N)
Noise example	0.0000	0.0000	0.0000	0.0000
Nominal level	0.0000	0.0000	0.0000	0.0000
dB (A)	0.0000	0.0000	0.0000	0.0000
$L_{pA,LF}$	0.0000	0.0000	0.0000	0.0000
Repetition no.	0.5814	0.6123	0.6804	0.1533
Gender	0.1888	0.0001	0.0001	0.0654
Age	0.0000	0.0000	0.0000	0.0000

Table A-1. The significance level of different factors that may influence the evaluation made by the reference group of listeners.

As expected, the noise example, the nominal level, the dB(A) level and the low-frequency level ($L_{pA,LF}$), all have a significant influence upon the evaluations from the test persons.

The repetition number (round 1 or round 2) is not significant and thus no training effect is seen in the data.

The gender of the test persons has influence on the evaluation of annoyance during the day and during the night but not on the evaluation of loudness and on the yes/no question about whether the noise is annoying or not.

The age of the test subjects is a significant effect. An inspection of the analysis shows that this is caused by one subject, age 21, with responses in the low end of the scale and another subject, age 24, with responses at the high end of the scale. The remaining subjects, age 19 to 25, gave responses in between.

1.3 Subjective Data from the Special Group

The same 10 cm long scale was used in the pilot experiment with the special group. Each of the four test subject indicated his/her response by setting a mark on a scale consisting of a 10 cm long line. The following tables show the evaluation of the various noises averaged over the four test subjects in the special group.

Average rating of loudness

Loudness	**Presentation level, dB(A)**		
Special group	**20**	**27,5**	**35**
Traffic noise	2,1	3,3	4,5
Drop forge	2,4	4,5	6,0
Gas turbine	1,4	4,7	7,6
Fast ferry	1,5	3,7	4,9
Steel factory	2,7	3,9	6,1
Generator	2,6	4,2	5,7
Cooling compr.	2,6	4,1	6,6
Discotheque	1,6	3,2	5,3
Avg.	*2,1*	*3,9*	*5,8*

Average rating of annoyance day/evening

Annoyance Day/evening	**Presentation level, dB(A)**		
Special group	**20**	**27,5**	**35**
Traffic noise	3,0	4,2	5,7
Drop forge	4,4	6,1	7,1
Gas turbine	3,4	6,3	9,0
Fast ferry	3,5	5,7	7,5
Steel factory	4,0	5,7	7,7
Generator	5,1	6,0	7,2
Cooling compr.	4,3	6,3	7,6
Discotheque	4,0	4,5	6,3
Avg. day	*4,0*	*5,6*	*7,3*

Average rating of annoyance at night

Annoyance Night	**Presentation level, dB(A)**		
Special group	**20**	**27,5**	**35**
Traffic noise	4,7	7,2	8,5
Drop forge	7,5	8,3	8,9
Gas turbine	5,0	8,1	9,8
Fast ferry	6,6	8,8	9,3
Steel factory	5,9	8,2	9,3
Generator	8,4	8,3	9,0
Cooling compr.	7,4	8,5	9,1
Discotheque	6,0	7,9	8,6
Avg. night	*6,4*	*8,2*	*9,1*

Average of percentage of 'yes' responses

Percentage Annoyed	**Presentation level, dB(A)**		
Special group	**20**	**27,5**	**35**
Traffic noise	25,0	25,0	75,0
Drop forge	75,0	75,0	100,0
Gas turbine	37,5	62,5	100,0
Fast ferry	62,5	100,0	100,0
Steel factory	50,0	100,0	100,0
Generator	75,0	87,5	100,0
Cooling compr.	87,5	100,0	100,0
Discotheque	50,0	87,5	100,0
Avg.	*57,8*	*79,7*	*96,9*

1.4 Statistical analysis of the subjective data from the special group

All data from the special group (i.e. subjective evaluation, sound example number, nominal presentation level, measured A-weighted level (dB(A)), A-weighted level in the frequency range 10 Hz – 160 Hz ($L_{pA,LF}$), repetition number, gender and age) were used as input to a statistical analysis program (Statgraphics 4.0).

As the special group of listeners consists of only four persons, it is almost irrelevant to test for normal distribution of the data. Despite of this it was decided to perform an analysis of variance on the data from the special group. The analysis was made for each parameter separately.

The table below shows the significance levels for the different factors. If the number is less than 0.05, this factor has a significant effect on the evaluation on a

95% level or above (this means less than 5 % probability for drawing a wrong conclusion). If the number is above 0.05 it cannot be proved that this factor has a significant effect upon the relevant evaluation.

	Loudness	Annoyance day	Annoyance night	Annoying? (Y/N)
Noise example	0.2118	0.0364	0.0592	0.0005
Nominal level	0.0000	0.0000	0.0000	0.0000
dB (A)	0.0000	0.0000	0.0000	0.0000
$L_{pA,LF}$	0.0000	0.0000	0.0000	0.0000
Repetition no.	0.7107	0.4968	0.6794	0.9984
Gender	0.1749	0.3651	0.3651	0.2722
Age	0.0018	0.0625	0.4447	0.0003

Table A-2. The significance level of different factors that may influence the evaluation by the special group.

Noise examples: The annoyance evaluations are found to be close to the limit for being significantly different. The Annoyance day is just below the limit (i.e. the noises are evaluated as different) whereas Annoyance night is just above the limit. This may mean that all the noises are evaluated almost equally annoying by the special group. The evaluation of loudness is not significant. The result of the yes/no question is a significant effect, i.e. some noises are found to be annoying others are not. This is somewhat in contradiction to the findings from the annoyance scaling. The limited number of test subjects in the special group probably causes the contradiction.

The nominal level, the dB(A) level and the low-frequency level ($L_{pA,LF}$) are significant in all cases.

The repetition number (round 1 or round 2 with the same presentation) has no significant influence, which shows the absence of a training effect.

The gender of the test persons has no influence on the evaluations. The age is significant for loudness evaluation and for the response to the yes/no question but not for the annoyance evaluations.

1 Data from Objective Measurements

1.1 Levels of the Noises

The 1/3-octabe levels of the noises are shown in the table below. The levels are equivalent levels averaged over the duration of the noise examples, 2 minutes, and are given in dB re 20 µPa. The data are from a nominal presentation level of 35 dB(A).

Freq.	Traffic Noise	Drop Forge	Gas turbine	Fast Ferry	Steel Factory	Generator	Cooling Compressor	Discotheque
Hz	1/3-oct level	1/3-oct level	1/3-oct level	1/3-oct level	1/3-oct level	1/3-oct level	1/3-oct level	1/3-oct level
8	50,7	23,8	62,1	41,0	59,5	46,4	27,5	50,3
10	54,8	25,0	57,5	50,3	63,9	48,4	30,7	62,5
12,5	58,1	34,8	58,8	54,9	67,4	46,2	35,9	63,6
16	60,5	60,1	65,5	55,3	61,5	44,5	42,4	55,5
20	58,8	67,5	59,1	53,2	63,2	44,3	38,1	45,9
25	59,4	58,0	77,8	47,8	67,1	44,3	41,5	44,6
31,5	59,1	59,6	56,5	51,3	57,2	38,0	45,2	40,6
40	56,8	64,0	53,1	52,4	49,6	46,1	49,2	46,6
50	53,8	62,7	58,9	58,4	52,1	45,1	59,0	49,1
63	51,1	52,0	41,8	58,3	57,0	47,5	38,2	47,7
80	54,1	51,4	37,4	49,1	47,0	58,3	40,2	50,0
100	44,6	44,6	27,0	43,4	35,3	34,0	48,3	45,5
125	36,5	36,7	25,9	33,2	28,8	30,4	38,8	44,5
160	31,5	29,1	28,4	25,9	28,9	28,7	41,3	33,1

The A-weighted levels (in dB re 20 µPa) calculated in the frequency range 10 Hz to 160 Hz are given in the following table for the three presentation levels.

Presentation level	Traffic noise	Drop Forge	Gas turbine	Fast Ferry	Steel Factory	Generator	Cooling Compressor	Discotheque
35	34,5	36,5	34,8	35,0	33,1	36,1	34,0	33,1
27,5	27,0	29,2	27,3	27,5	25,6	28,6	26,5	25,7
20	19,7	21,9	19,8	20,1	19,0	21,3	19,6	18,7

The C-weighted levels (in dB re 20 µPa) calculated in the frequency range 10 Hz to 160 Hz are given in the following table for the three presentation levels.

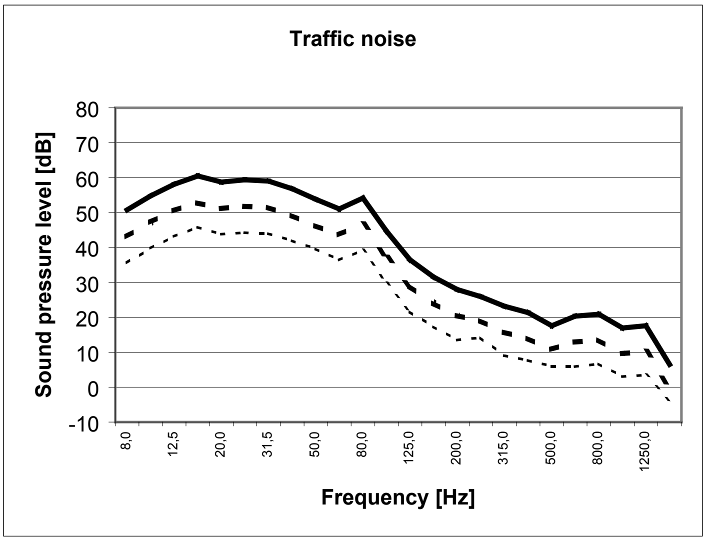
Presentation level	Traffic noise	Drop Forge	Gas turbine	Fast Ferry	Steel Factory	Generator	Cooling Compressor	Discotheque
35	63,0	67,4	73,7	61,8	66,0	58,7	58,8	57,8
27,5	55,4	60,0	66,1	54,3	58,4	51,2	51,4	50,5
20	48,1	52,5	58,6	47,0	51,3	44,5	44,8	44,2

Annex B

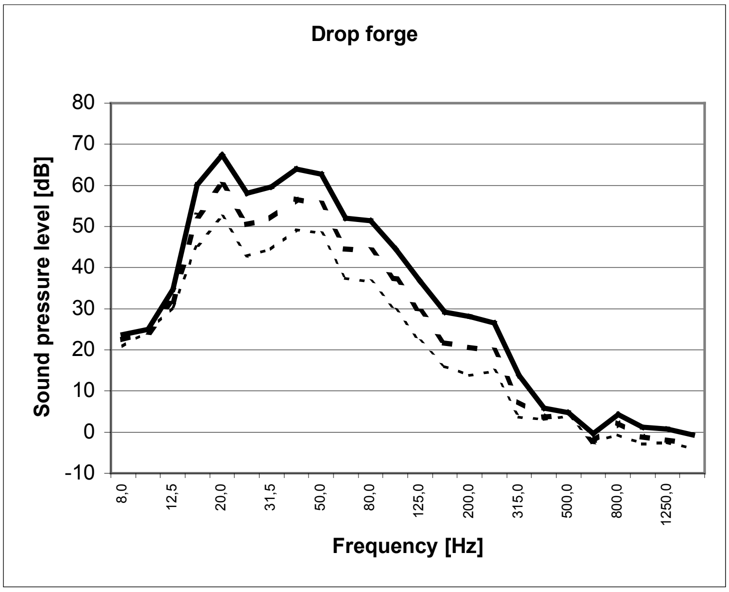
1.2 Spectra of the noises

The 1/3-octave spectra of the noises are shown in the following figures. All measurements were made acoustically at the test subjects position. The spectra are show for the three nominal presentation levels: 35 dB(A) [solid line], 27,5 dB(A) [thick dashed line] and 20 dB(A) [thin dashed line].

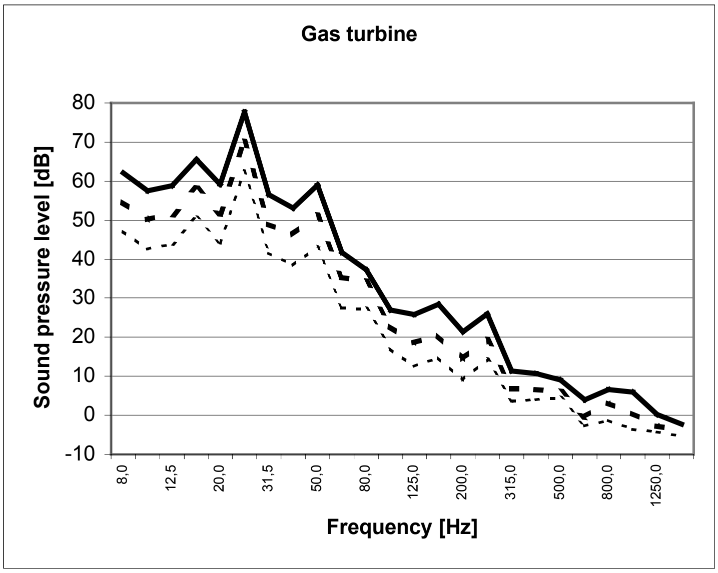
Traffic noise:



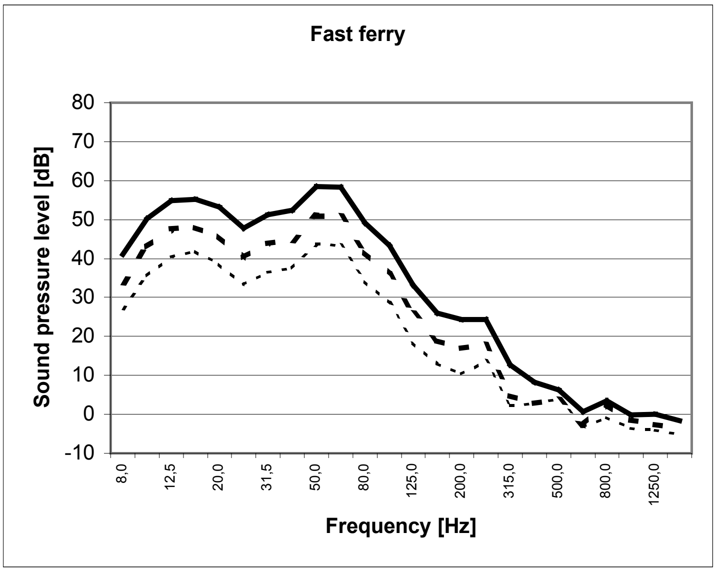
Drop forge:



Gas turbine:

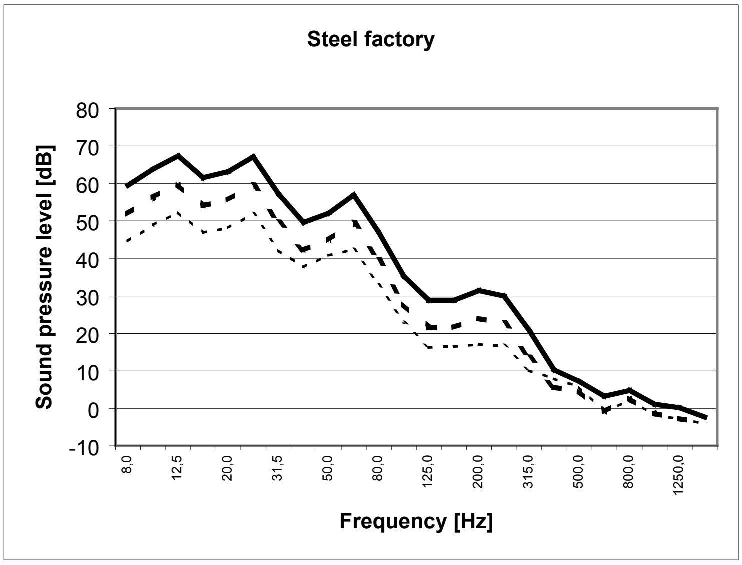


Fast ferry:

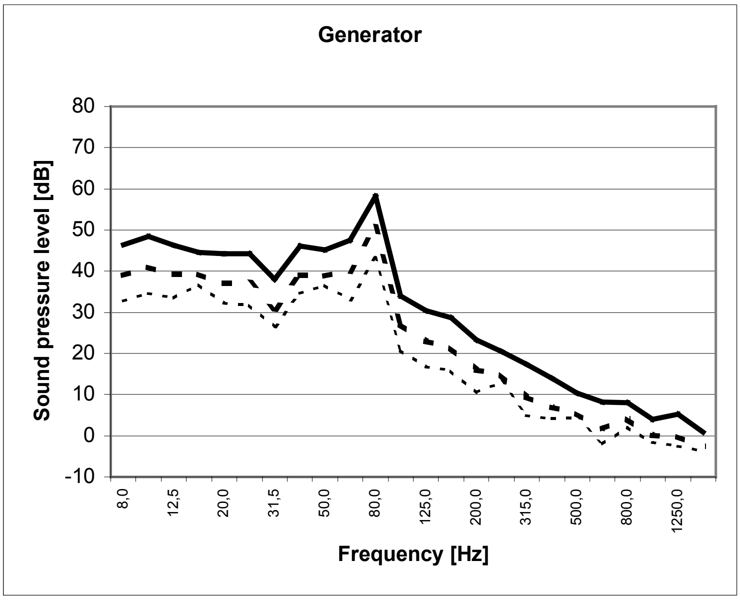


Annex B

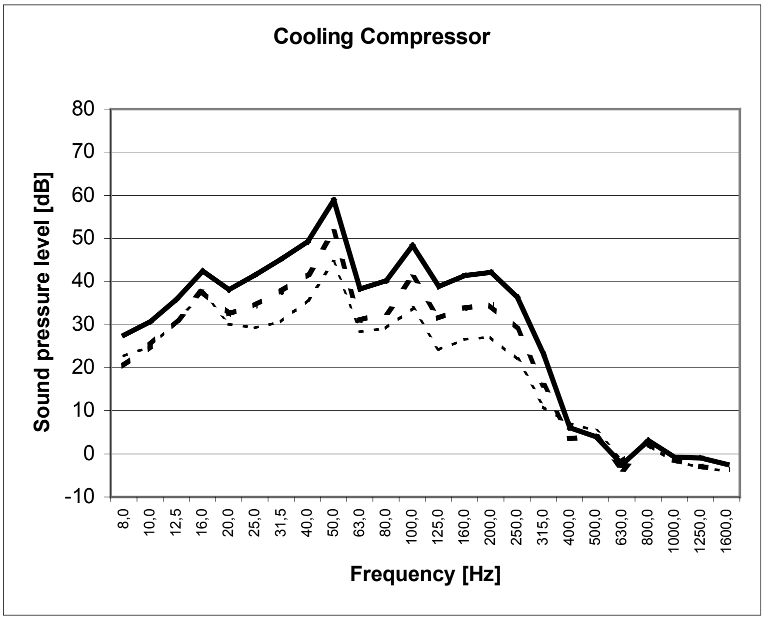
Steel factory:



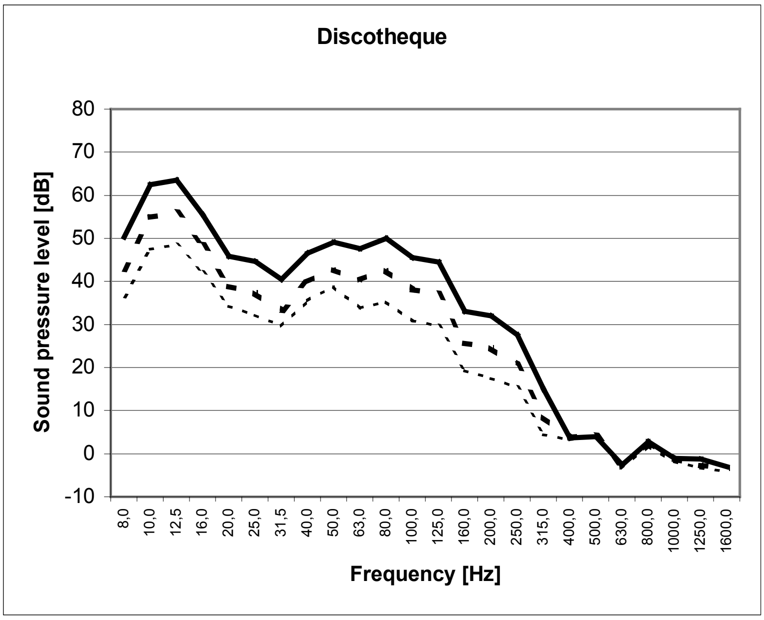
Generator:



Cooling compressor



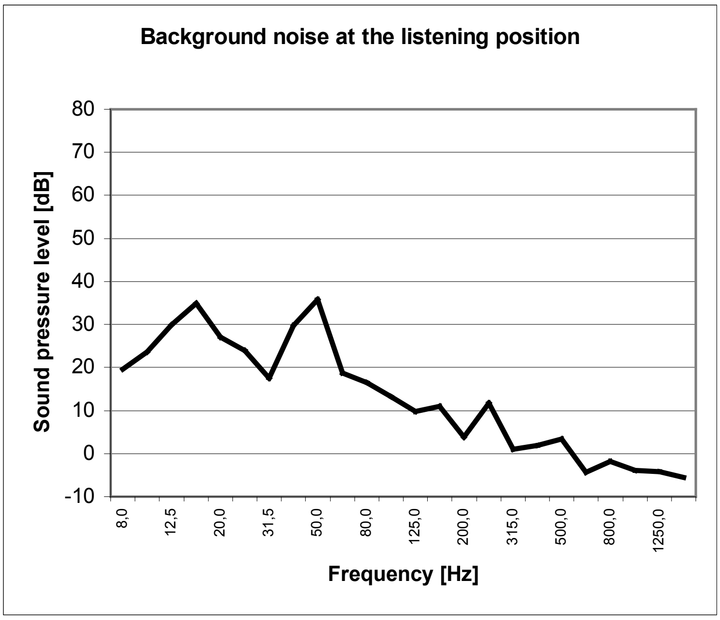
Discotheque:



Annex B

1.3 Background noise

The background noise at the listening position is shown in the figure below.



-ooOoo-