Karolinska Institutet
Technical Audiology
KTH, 100 44 Stockholm 70

Report TA No. 86
June 1977

ADJECTIVE RATINGS AND DIMENSION ANALYSES OF PERCEIVED SOUND QUALITY OF HEADPHONES.

Alf Gabrielsson, Sten-Åke Frykholm and Håkan Sjögren

ABSTRACT

Twenty musical subjects made ratings on thirty adjective scales concerning the perceived sound quality of eight headphones reproducing five different programs of music. The ratings were subjected to various forms of analysis of variance and factor analysis. The factor analysis resulted in five factors interpreted as "Hardness/Sharpness/Loudness - Softness", "Clearness/Distinctness", "Brightness - Darkness", "Feeling of space", and a factor related to various "disturbances" in the reproductions. The relations to physical parameters of the headphones are preliminary discussed.

INTRODUCTION

Several papers on perceived sound quality of loudspeakers and of hearing aids have recently appeared in this report series (Gabrielsson, Rosenberg & Sjögren, 1971, 1972, 1973; Gabrielsson & Sjögren, 1974, 1975, 1976, 1977). The present report deals with perceived sound quality of headphones. This is studied by means of adjective ratings and factor analysis as in many of the above-mentioned reports, which also present the general background for this research project. The present experiment also includes a more detailed study of the relations between various physical parameters of the headphones and the perceptual dimensions. Since this part of the investigation is rather extensive and introduces several new techniques it will be described in a separate report to appear.

METHOD

Twenty musical subjects made ratings on 30 adjective scales concerning the perceived sound quality of eight headphones reproducing five different pieces of music. The data were treated by various forms of analysis of variance and factor analysis.

Rating scales (adjectives)

Using information from the earlier studies 30 adjectives were selected assuming that they would cover most of the variations in the perceived sound quality of the present reproductions. As discussed in earlier reports this selection is always somewhat arbitrary, and the relevance of the scales has to be checked in several independent studies. The adjectives are given in Swedish alphabetical order in Table V together with translations into English.

Stimuli and listening conditions

The stimuli were five different music programs presented stereophonically over each of eight headphones. Listening levels were set by the experimenters to correspond approximately to the original sound level at the respective performances. This task was facilitated by having available detailed information about the recordings of three of the programs (program 1, 3, and 5). The reasons for using (approximate) original sound levels in listening tests are presented on an experimental basis in Gabrielsson & Sjögren (1976). The programs were:

Program 1 (P1): The very end of the finale chorale in "St. John Passion" (J.S. Bach) performed by the Bach choir at Adolf Fredrik, Stockholm. Recorded in the empty church of Adolf Fredrik, Stockholm, using two omnidirectional microphones. Sound level: 85 - 97 dB(A). Gramophone record: PROPRIUS 7741.

Program 2 (P2): Oscar Peterson's jazz trio (piano, bass, drums). Recorded in a gramophone studio. Sound level 80 - 90 db(A). Gramophone record: VERVE 2304 062, sample from the tune "Something 's coming".

Program 3 (P3): "Solveig's sang" from the "Peer Gynt" suite by Grieg, sung by Grynet Mollvig accompanied by piano, bass, and choir. Recorded in the auditorium of Ljungskileskolan, Sweden, synthetic reverberation added afterwards. Sound level 80 - 88 dB(A). Gramophone record: PROPRIUS 7739.

Program 4 (P4): Excerpt from the "Summer" concerto in "The Four Seasons" by Vivaldi, performed by the Academy of St. Martin-In-The-Fields. Sound level: 80 - 90 dB(A). Gramophone record: ARGO ZRG 654.

Program 5 (P5): Excerpt from the end of "The Firebird Suite" by Stravinsky, performed by the Stockholm Philharmonic Orchestra. Recorded in the Concert Hall of Stockholm, using two omnidirectional microphones. Sound level: 90 - 97 dB(A). Special recording for the Cullberg ballet.

The program sections lasted for about 30 seconds each and represented "musically homogeneous" sections. The stimulus tapes were copied directly from the master tapes as regards programs 1, 3, and 5, and from gramophone record in the remaining two cases. Tape copies of the program sections are available at self-cost.

The programs were reproduced by eight different headphones. In the following they are labelled H1, H2, ... H8. The headphones represented different technical solutions to the transducer design problem. Electro dynamic, electrostatic, ortho dynamic and piezo electric systems were all included. Also in the way of applying them to the ear they differed, supra aural and circum aural as well as completely open types were represented.
Their frequency responses, as measured by the IEC coupler (Brüel & Kjaer 4153), are given in Fugure 1.

Measurement of the frequency response of headphones meet with many difficulties. The transfer function from the headphone's membrane to the eardrum is different for different individuals due to differences in the outer ear and in the ear canal. To make possible some kind of standardized measurement use is made of a coupler which simulates the acoustical impedance of an average ear. Various couplers have been constructed for various purposes. The present IEC standard for measurement on Hi-Fi headphones recommends a three volume coupler Brüel & Kjaer 4153, in which the sound pressure is measured in a plane just inside the adhesion plane of the headphone. An alternative is the Zwislocki coupler in which the sound pressure is measured in the plane of the ear drum.

The commercially available stereo music is most often intended for reproduction with loudspeakers in a normal living-room. From such recordings a true headphone sound image of the original cannot be recovered. Therefore there is no unanimous solution to the "ideal" headphone transfer function problem. Usually a flat response on the IEC coupler will give a good tonal balance. This of course corresponds to a non flat response on the Zwislocki coupler due to ear canal influence.

Another problem lies in the way of applying the headphone to the coupler. By various types of adapters the IEC coupler can be used for measurement on different types of headphones. Special care must be taken for circumaural headphones since an adapter plate may give rise to resonances. On the other hand "open" types are very sensitive to small variations in the way of applying them to the coupler.

A more detailed discussion on measurement of the frequence response and physical parameters in headphones is given separately in a forthcoming report.

The level of the different headphones was approximately equalized using a white noise signal at the octave 1000 Hz and broadband dB(A), see Table I. The equalization was checked perceptually by the authors and by three trained subjects in a pilot experiment to give approximately the same loudness impression from all headphones when they were listened to in rapid succession for each program section.

The subjects, in general two at a time, were sitting in a laboratory room, turned away from each other. All technical equipment (except for the headphones) was handled in an adjoining room, from which the experimenter could see the subjects. All headphones were kept in the laboratory room and could not be seen by the subjects.Furthermore the subjects were not allowed to put on the headphones or adjust them in any other way. This was instead made for each new trial by the experimenter, who then walked into the room and changed headphones approaching the subject from behind and adjusting the phones until the subject was satisfied.

Subjects

There were 20 subjects, 14 males and 6 females, 19-29 years old (except two subjects who were 46 years old). All subjects had normal hearing, less than 15 dB hearing loss 125-8000 Hz (one subject had a 25 dB loss at 4000 Hz in one ear).

The subjects were recruited by means of advertisements put up at the Royal Academy of Music and in the Concert Hall of Stockholm. As conditions for participation it was requested that the subjects should often listen to live performances of music. Nearly all of them had long experience of own musical performance and in general visited concerts of different kinds at least twice a month. Most subjects listened to moderately advanced high-fidelity equipment at home. On a question if they were used to listen by headphones and/or had compared different headphones 18 subjects answered "No" or "Seldom". Two subjects may be said to have a certain experience of comparing headphones, however, not to the extent exercised by Hi-Fi enthusiasts. All subjects were paid for their participation.

Procedure

In total there were 40 different combinations of program sections and headphones (5 programs x 8 headphones). These are referred to as P x H combinations. Since there were 30 adjectives each subject made 40 x 30 = 1200 judgements. This required two experimental sessions of about two hours each. Each such session comprised all 40 P x H combinations in a randomized order (the randomization was different for different subjects and different sessions), and the subjects made ratings for half of the adjectives on the respective list in each session. At each reproduction of a program section over a certain headphone the section was always played two times in succession (with 1-2 seconds break between the times) and thus lasted for 60-70 seconds.

Each subject had a number of adjective lists, one for each of the 40 P x H combinations. The order of the 30 adjectives on each list was randomized, different for different lists and different subjects. The instruction conformed to that used in earlier reports. It was given both orally (tape-recorded) and in written form. The main parts were as follows:

"In this experiment we want to investigate how you perceive the sound reproduction of different headphones. You will listen to some different pieces of music. Each presentation lasts for about one minute. For every such case you shall describe how you perceive the sound reproduction of the respective piece by writing a number from 0 to 9 for each of the adjectives on the respective list.

0 means that the reproduction has nothing of the quality denoted by the adjective. 9 means, on the contrary, that the reproduction has a "maximum" of that quality. For levels between these extremes you use values in between. The more of the quality, the higher number (up to 9); the less of the quality, the lower number (down to 0). Only one number shall be written and decimals are not allowed. Try to make use of the whole scale 0-9. The reproductions are so varying that there are opportunities for using the whole scale.

It is very important to observe that the judgements shall refer to the sound reproduction, not to the music as such! It may sometimes be difficult to do this distinction - but do think of it and concentrate yourself on the sound reproduction!"

Further different practical matters were included in the instruction. The adjectives were left undefined except for the expression "True to nature". This was explained as follows: "True to nature" refers to how well the headphone in question reproduces the original sound. Thus you have to imagine how the original music sounded and judge the reproduction in relation to that... 9 denotes a perfectly true reproduction (no difference between original sound and reproduction) and 0 "no fidelity at all".

After the instruction in the first session there were about ten preliminary trials in which the subjects trained to do ratings at different P x H combinations. They were informed about the music and the respective recordings by reading an abbreviated version of the information about programs given under Stimuli above. No information at all was given about the number of headphones or about their characteristics in any form. In the second session the instruction was repeated and four preliminary trials made. There were two short breaks within each session. At the end of the second session the subjects answered some questions related to the experiment.

Data treatment

Analysis of variance was applied for each of the 30 adjectives as dependent variable. The sources of variance in each analysis were the headphones, the programs, and the subjects (plus the possible interactions), and each analysis was thus based on 800 ratings (8 x 5 x 20). Overall F tests were performed for the headphone and the program variables using a mixed, non-additive model (subjects considered as a random variable). Following significant F ratios for headphones pairwise comparisons between headphones were made using Tukey's HSD statistic (*a posteriori* comparisons; Kirk, 1968). The variance analyses were also used to estimate *inter*-individual reliabilities in each of the rating scales in analogue with procedures described by Winer (1962, p. 124; 1971, p. 283).

Considering subjects as replications the ratings were then averaged over subjects for each of the 1200 P x H x A combinations (A = adjective), see Table II, and in a further step also over programs, Table III. The means given in these tables were subjected to various forms of factor analysis (component analysis) according to procedures described in detail in earlier reports (Gabrielsson & Sjögren, 1974, 1975). Suitable references for factor analysis in general are the works by Harman (1967) and Gorsuch (1974).

RESULTS

Differences between headphones

Using the means in Tables II-III and the summarized evidence from the analyses of variance given in Table IV the results of the ratings in the 30 adjective scales may be summarized as follows.

There were highly significant differences between the headphones in almost all adjective scales. (The single exception from this occurs for "Loud", see further below). In this sense the selection of the adjectives seems to have been successful. The values given in the column marked HSD in Table IV give an approximate estimation of how big difference (in any direction) is required between any two headphones to be statistically significant within the respective adjective scale. The HSD values are fairly similar for all scales (most of them lie between 0.8 and 1.1), and so it may roughly be said that a difference of about one unit between any two headphones in any row of Table III is statistically significant.

Significant interactions between headphones and programs appear in 12 of the 30 adjective scales. This means that the differences between the headphones in the respective scale somehow vary from program to program. In such cases it is not enough to look at Table III (in which the ratings are averaged over programs) but the more extensive Table II must be consulted to compare the differences between the headphones at different programs. To take but one example: The difference between H6 and H7 in the rating scale "White noise/ Hissing" is about 1.5 units when averaged over the programs (see in Table III). Inspection of the corresponding data at each of the five programs (Table II) reveals that the difference may be even bigger than so (for P1 and P3) while there is practically no difference at all at P5. The presence of an interaction between the programs and the headphones is thus a signal to look at the data more carefully to interpret the meaning of the interaction. Numerous interactions were found in our earlier investigations.

The single scale in which no significant differences between headphones occurred was the "Loud" scale. Since the intention was to equalize the headphones in perceived loudness (see under Stimuli) the absence of significant differences is an indication that this was in fact accomplished. However, there is a significant interaction between program and headphones for the "Loud" scale. Comparing the data for the headphones in this scale when averaged over programs (Table III) and at each program separately (Table II) show that, although the loudness of the headphones was fairly similar in average over the five programs, there were certain differences of one unit or slightly more at programs P2 - P5. Which headphones sound most different in loudness varies for different programs.

In certain scales there were also significant differences between the program sections (Table IV). These differences are in general easily understood from the facts about sound levels and recording conditions (see Stimuli), the presence of certain musical instruments etc, and are not further commented here.

Reliability of ratings

The many significant differences are in themselves an indication of the reliability of the data. The more formally estimated reliability index for inter-rater reliability given in Table IV varies between 0.60 and 0.91 for most adjectives. For five adjectives this index is lower than 0.60, viz. "Hollow", "Clear", "Pure", "Jarring", and "Whistling/Whizzing". Possible reasons for lower inter-rater reliability are that the subjects disagree about the meaning of the respective adjective as applied to sound reproduction. Another, and more statistical, reason is that this way of estimating reliability (Winer 1962, 1971) will tend to decrease the reliability index if there are no sizable differences between the headphones in the respective scale. The headphones used here all sounded relatively "clear" and "pure" and not especially "jarring". The lower reliability indices for these adjectives may thus be a function of the stimulus context.

Factor analyses

Since apparently many of the adjectives are somehow related in meaning to other adjectives it is usually more efficient to discuss the results in terms of a lower number of "more basic scales" arrived at by studying the correlations between the adjective scales. (For the same reason a multivariate analysis of variance approach may be useful). To accomplish this a number of different factor analyses (principal component analyses) were made starting from the correlations between the adjectives over all P x H combinations (Table II) or over headphones (Table III). Various other approaches, including three mode factor analyses, were also tried in analogy with what is described in many earlier reports. The various solutions were compared with regard to "simple structure", proportion variance accounted for, and above all the interpretability of the resulting factors. Orthogonal ("varimax") as well as oblique ("simple loadings") rotations were tried using the BMD 08M computer program. On the whole the various solutions were very similar to each other as regards number and interpretation of factors. Therefore only one solution is presented in the following, the one which seemed to give most consistent and clear interpretation.

A five factor oblique solution starting from the correlations between adjectives over P x H combinations (Table II) accounted for 85.8% of the total variance. The factor loadings for the adjectives appear in Table V and the factor scores for the P x H combinations in Table VI. The interpretation of the factors and their relation to the physical characteristics of the headphones and programs is made using the evidence in these two tables and the information about headphones and programs given under Stimuli, above all the frequency response of the headphones. The emphasis here is on the identification of perceptual dimensions. Their relations to physical parameters are discussed only in a tentative and restricted manner awaiting a more formal and detailed exposition in a forthcoming report.

Factor 1 (F1) may be labelled "Hardness/Sharpness - Softness", possibly confounded with "Loudness". There are a few high factor loadings on each side of the continuum: for "Loud", "Jarring/Grating", "Hard", and "Sharp/Keen" on one side, for "Soft" on the other side.

Studying the factor scores for P x H combinations in F1 indicates that this factor partly reflects differences between the different programs. The highest factor scores on the "Soft" side occur for most reproductions of P3, which has the lowest sound level among the programs, probably also less energy in the treble region. The highest factor scores on the "Hard/Sharp/Loud" side occur for various reproductions of P5, which has the highest sound level among the programs and a lot of brass instruments and percussion playing in *fortissimo*.

There is also a considerable variation between the factor scores of the headphones. Within each program H6 is the most "Soft" one, followed by H5 or H4 (and H8 for P3), while H3 is fairly outstanding on the "Hard/Sharp/Loud" side. A look at the frequency responses (Figure 1) suggests that this perceptual character of H3 is related to the prominent peak around 3000-4000 Hz occurring in its frequency curve.

Factor 2 (F2) may be interpreted a "Distinctness/Clearness". The highest factor loadings on one side of the continuum appear for "Clear" and "Pure/Clean" followed by "True to nature" and "Feeling of presence". On the opposite side an outstanding high loading occurs for "Diffuse".

The evidence of the factor scores in F2 shows that H1 lies highest on the "Clear/Pure" side within each of the programs, while H5 and H3 lie utmost on the other side (except for H5 at P3). The order of the remaining headphones varies from program to program. It may be suggested that the relatively more "Diffuse" character of H5 is related to the bass boost in its frequence curve. As regards H3 one reason is again probably its pronounced peak around 3000-4000 Hz in the sense that other frequency regions than this are suppressed in its reproduction.

Factor 3 (F3) seems related to various unwanted "disturbances" in the reproduction. High positive factor loadings occur for "Crackling/Crunching", "White noise/Hissing", "Hissing", and "Whistling/Whizzing".

In the factor scores for F3 there are indications that this factor partly reflects characteristics of the programs (or recordings). The reproductions for P3 lie more towards the "disturbance" side that what is the case for the other programs, especially for P5. P3 represents the lowest sound level among the programs and thus a somewhat lower signal-to-noise ratio. On the other hand P5 has the highest sound level among the programs and this fortissimo music effectively masks the tape noise and related phenomena.

As regards the headphones H3 lies most towards the "disturbance" side within all programs except P3, followed by H1 and H7, while H6 lies most away from the "disturbance" side (except at P2) together with H5 or H4. It may be suggested that this perceptual dimension is related to the presence of resonance peaks at higher frequency regions as in the frequency responses of H3, H1, and H7.

Factor 4 (F4) may be labelled "Brightness-Darkness", possibly with a touch of "Fullness". The adjectives "Bright" and "Thin" have the highest factor loadings on one side, while "Emphasized bass", "Rumbling", and "Dull" dominate the opposite side (note also "Full/toned/" with a moderate high loading).

In the factor scores for F4 the "Bright/Thin" side is represented above all by H3, followed by H8, while the opposite "Dark/Bass" side has H5 as rather outstanding example within each program. It seems fairly evident that these perceptual characteristics reflect differences in the frequency responses: the peak around 3000-4000 Hz in H3 versus the bass boost in H5.

Factor 5 (F5) is aptly described by its single high loading on the positive side occurring for "Feeling of space". In contrast to this the highest loading on the other side appears for "Shut up/ Closed".

No doubt this factor reflects much of the recording conditions for the different programs. In the factor scores for F5 it is striking that reproductions of P2, which was recorded in a studio, give less impression of space than the other programs which were recorded in big rooms, for instance, P1 in a church and P5 in a concert hall.

There are, however, also recurring differences between the headphones in this factor. H1 gives most "feeling of space" within each of the programs (except for P3), while H3 is extreme on the "Shut up/Closed" side, followed by H8 or H5. It is noted that H1 is the single headphone which is not directly applied against the outer ear (Figure 1).

There were four moderate intercorrelations between the factors in this five factor oblique solution. The "Hardness/Sharpness - Softness" factor (F1) correlated negatively (-.41) with "Brightness-Darkness" (F4). As seen in the factor scores headphones which belong to the "soft" side of F1 for the most part also belong to the "darker" side of F4. The "Distinctness/Clearness" factor (F2) correlated positively (.45) with the "Feeling of space" factor (F5). It is seen in the factor scores that the more "distinct/clear" headphones in F2 in general lie high in the "feeling of space" dimension. The "disturbance" factor (F3) correlated positively (.39) with "Brightness-Darkness" (F4): there is a certain tendency for headphones with high positions in the "disturbance" factor to belong to the "bright" side of F4. Finally "Brightness - Darkness" correlated negatively (-.33) with "Feeling of space": there is a slight tendency for headphones on the "bright" side to give less "feeling of space". The remaining intercorrelations between the factors were negligibly low.

Evaluative ratings

Although most of the adjectives somehow reflect various positive or negative characteristics in sound reproduction, two adjectives may be said to represent a kind of "overall" evaluation, viz."Pleasant" and "True to nature". It is interesting to note their factor loadings in the five factors described above. In the "Hardness - Softness" dimension (F1) "Pleasant" has a moderately high loading on the "Soft" side, while "True to nature" lies in a "neutral" middle position. In the "Clearness/Distinctness" dimension (F2) "True to nature" has a high loading on the "Clear/Distinct" side, while "Pleasant" is only moderately loaded on the same side. In the remaining three factors their positions are fairly close to each other: far away from the "disturbance" side in F3, fairly low loadings on the "Dark/Bass" side in F4 and on the "Feeling of space" side in F5.

Another interesting point is to see which adjectives show the highest correlations, positive and negative, with "Pleasant" and "True to nature". This information is given in Table VII. The results conform well to what may be expected. The high correlations between certain of the adjectives and "Pleasant" or "True to nature" suggest that it might be possible to "predict" with high accuracy what the "overall" evaluation of an headphone would be given the ratings in certain adjective scales (for instance, by multiple regression procedures). It would be even more interesting, of course, if a prediction of the "overall" evaluation could be made from the physical characteristics of the headphone (that is, from frequency responses, distortion etc.).

The mean ratings for the headphones in the "Pleasant" and "True to nature" scales appear in Tables II - III.

DISCUSSION

When considering the results of this experiment several limitations should be noted. The results are valid only for the music programs and headphones used here as judged by a limited sample of musical subjects on a limited number of adjective scales etc.

Listening tests on headphones present certain problems not present when testing, for instance, loudspeakers. The subject may get cues from the tactile sensations on his head of various headphones, resulting from different weights of headphones, different ways of attachment to the ears and so on. After the experiment was completed, the subjects were asked how many headphones had been used and if they could recognize any of them from earlier experience or from the way they were put on to the head etc. The guesses regarding the number of headphones were fairly equally distributed in the range from five to ten headphones. Two subjects correctly identified H1 by name, otherwise there were few and wrong guesses as to names. Many subjects stated that they recognized some headphones from the way they were put on to the head, especially that headphone "that was not attached to the ears" (that is H1), but also stated that such recognitions had not affected their judgements or only slightly so.

As in other listening tests, the setting of the sound level for the systems to be tested presents many problems. It may be generally recommended that the systems should be equalized in a way that the subjective loudness is about the same for all systems. In the present case this was done by equalizing the sound level from the different headphones for the octave around 1000 Hz and for broadband condition, supplemented with listening (see Stimuli). However, the measurements of the sound level are made by means of a coupler and are thus very sensitive to variations in the way the headphone is adapted to the coupler (see Stimuli). As regards subjective listening there is a risk of confounding "loudness" and "hardness/sharpness", and it may be difficult to compare loudness from headphones with very different frequency responses. In the present experiment the attempts at loudness equalization seemed to be successful when considered in average over the five programs (see under Results). However, within certain programs there may have been some differences in loudness between some headphones. Whether this has affected judgements in other scales is not possible to evaluate.

The purpose with this investigation was not to design a test of eight selected headphones. In fact eight other headphones could have been selected just as well as those here. One aim was to gain some insight into specific problems associated with testing of headphones as discussed above. The primary aims, however, were to see if the dimensions for describing perceived sound quality obtained in earlier studies would hold also for headphones, and further to try to understand in more detail the relations between various physical parameters of the headphones and the perceptual dimensions. In fact the perceptual dimensions found by factor analysis agree very well with dimensions obtained in our earlier investigations on loudspeakers and hearing aids. Furthermore there are apparent agreements with earlier results concerning the reliabilities, the pattern of correlations between various adjectives and the overall evaluating scales ("Pleasant", "True to nature") etc.

The discussion and results about relationships between physical parameters and perceptual dimensions proved to be so complex and extensive that it will be saved for separate presentation in a following paper in this report series.

ACKNOWLEDGEMENTS

This investigation was initiated and supported by the Swedish Philips Group.

REFERENCES

Gabrielsson, A., Rosenberg, U. & Sjögren, H. (1971, 1972). Judgements and dimension analyses of perceived sound quality of sound-reproducing systems. I and II. Reports from Technical Audiology, Karolinska Institutet, Stockholm, No. 66 and 68.

Gabrielsson, A., Rosenberg, U. & Sjögren, H. (1973). Adjective ratings and dimension analyses of perceived sound quality of sound-reproducing systems. Reports from Technical Audiology, Karolinska Institutet, Stockholm, No. 72.

Gabrielsson, A. & Sjögren, H. (1974, 1975b, 1977). Adjective ratings and dimension analyses of perceived sound quality of hearing aids. I, II, and III. Reports from Technical Audiology, Karolinska Institutet, Stockholm, No. 75, 77, and 85.

Gabrielsson, A. & Sjögren, H. (1975a). Similarity ratings and dimension analyses of perceived sound quality of hearing aids. Report from Technical Audiology, Karolinska Institutet, Stockholm, No. 76.

Gabrielsson, A. & Sjögren, H. (1976). Preferred listening levels and perceived sound quality at different sound levels in "high fidelity" sound reproduction. Report from Technical Audiology, Karolinska Institutet, Stockholm, No. 82.

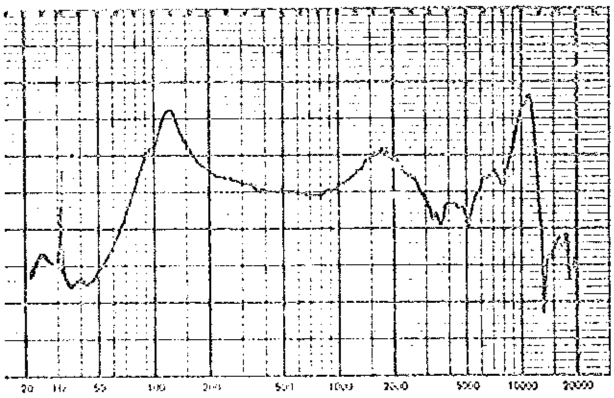
Gorsuch, R.L. (1974). Factor Analysis. W.B. Saunders Comp., Philadelphia.

Harman, H.H. (1967). Modern factor analysis (rev. ed.) University of Chicago Press.

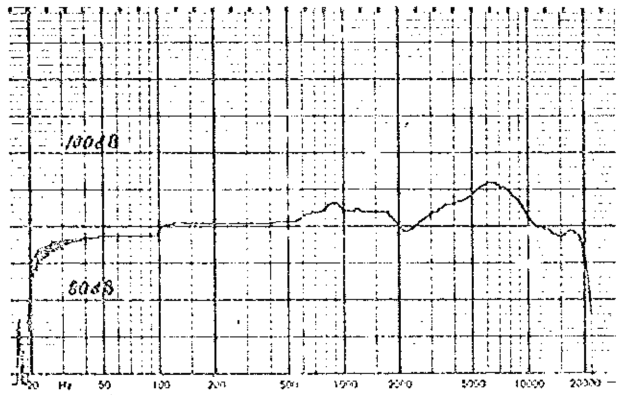
Kirk, R.E. (1968). Experimental design. Procedures for the behavioral sciences. Belmont, California: Brooks/Cole.

Winer, B.J. (1962, 1971). Statistical principles in experimental design. (1st and 2nd ed.). New York: McGraw Hill.

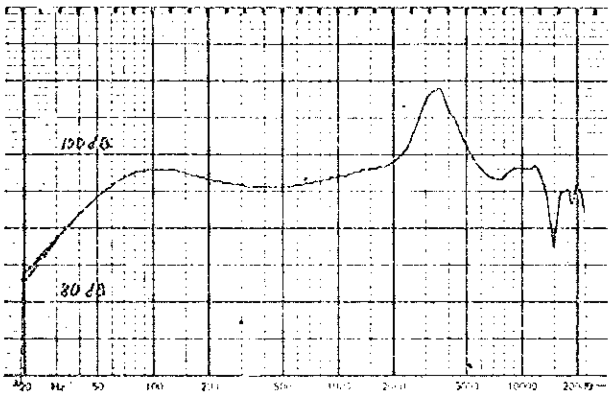
FIGURE 1 FREQUENCY RESPONSE OF 8 HI FI HEADPHONES MEASURED ON IEC COUPLER (TYPE B&K 4153)



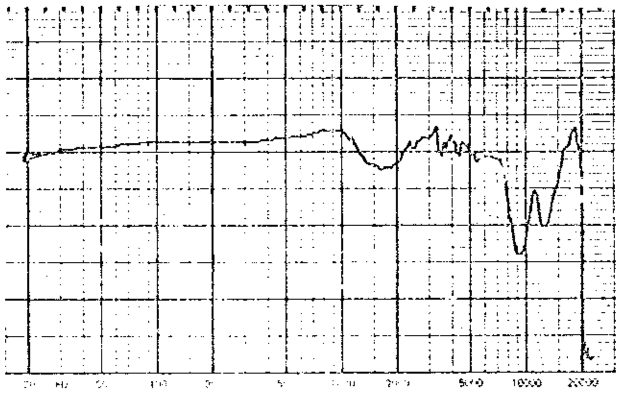
H 1



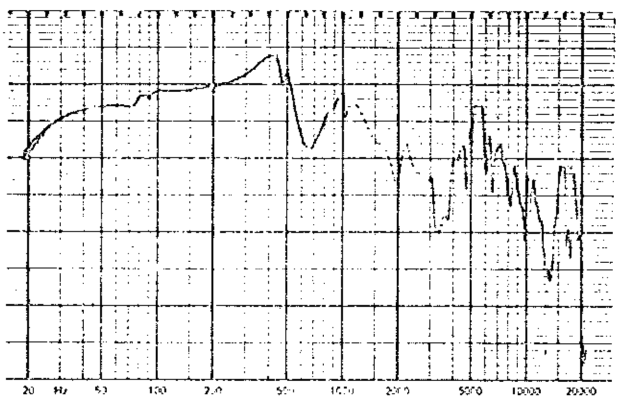
H 2



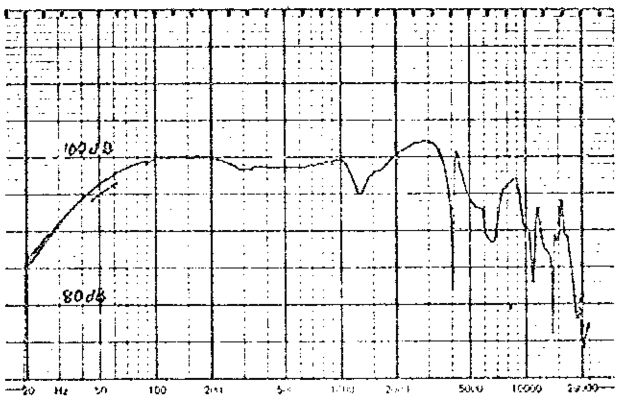
H 3



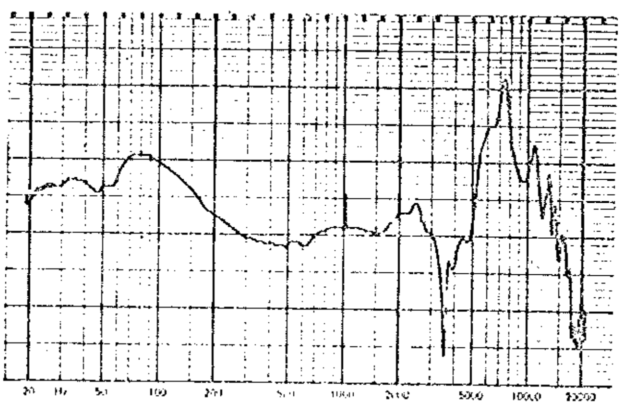
H 4



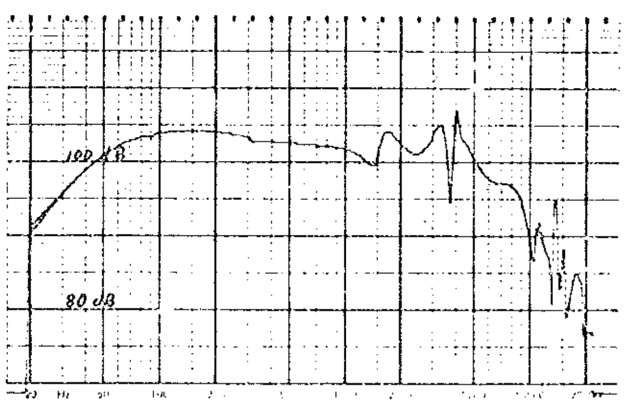
H 5



H 6



H 7



H 8

TABLE I. Relative level (dB) at the octave 1000 Hz and broadband (dB, A-weighted) for eight headphones using a white noise signal and with headphone 5 as reference. Differences between right and left headphones are within 1 dB.

Headphone	1000 Hz	Broadband
1	-2.0	1.0
2	1.0	2.0
3	-2.0	6.0
4	2.0	2.0
5	0.0	0.0
6	0.0	2.0
7	-4.5	3.5
8	-2.5	5.0

TABLE II

Mean ratings for 8 headphones, averaged over 20 subjects, in 30 adjective scales at five music programs.

* standard deviation between 1.00 and 1.99

** standard deviation between 2.00 and 2.50

PROGRAM 1 ADJECTIVES	HEADPHONES H1	H2	H3	H4	H5	H6	H7	H8
BALANSERAD	6.35 **	5.00 *	4.25 **	5.40 *	4.90 **	5.60 *	4.70 *	4.35 **
BEHAGLIG	5.50 **	4.60 *	3.15 **	5.70 *	5.00 *	6.15 *	4.25 *	3.70 **
BRUSIG	4.45 **	4.20 **	3.60 **	2.60 *	2.85 **	2.20 *	4.20 **	2.65 **
DIFFUS	2.95 **	3.90 **	3.90 **	3.10 *	4.50 **	3.05 **	4.05 **	3.30 **
DOV	2.95 *	3.30 **	2.15 **	3.00 *	5.80 **	3.15 **	3.10 *	2.55 **
FRAMHAVD BAS.	4.00 **	2.60 *	1.65 *	2.90 **	5.35 *	2.85 *	2.70 *	1.35 *
FRAMHAVD DISKANT	4.30 **	4.40 **	6.40 *	4.50 **	2.85 **	3.45 **	4.55 *	5.25 **
FRASANDE	3.50 **	3.00 **	4.25 **	2.10 *	2.50 **	2.20 **	3.85 **	2.95 **
FYLLIG	5.85 **	5.30 *	2.10 *	5.50 *	5.65 *	5.30 *	4.10 *	3.30 *
HARD	3.00 *	3.60 *	5.50 **	3.95 **	2.40 **	3.42 **	3.50 *	4.65 **
IHALIG	2.65 **	2.75 *	4.60 **	2.75 **	3.35 **	2.40 *	3.25 *	3.85 **
INSTANGD	2.40 **	3.00 **	4.60 **	3.15 **	5.10 **	2.60 **	2.30 *	4.10 **
KLAR	5.05 **	4.25 **	4.10 *	5.25 *	3.00 *	5.05 *	4.40 **	5.00 **
KNASTRANDE	2.35 **	2.65 **	3.25 **	1.79 *	1.85 *	1.80 *	3.05 *	2.05 *
LJUS	4.75 *	4.45 **	6.15 *	4.10 *	2.90 *	4.80 *	5.00 *	5.90 *
MATT	2.20 **	3.10 **	3.20 **	3.10 **	3.80 **	2.80 **	3.30 **	3.40 **
MJUK	4.45 **	3.85 **	2.05 *	4.25 *	4.40 *	4.95 *	3.40 *	2.60 *
MULLRANDE	2.80 *	2.15 *	1.70 *	3.00 *	4.20 **	2.50 *	2.80 *	2.30 *
NASAL	2.70 *	2.90 **	4.85 **	2.80 **	3.30 **	2.63 **	3.65 **	4.20 **
NATURTROGEN	5.70 **	4.75 **	3.10 **	5.25 *	3.74 *	5.47 *	4.75 **	3.90 **
GER NARVAROKANSLA	5.75 **	5.20 *	3.50 **	4.95 *	3.60 **	5.45 *	4.55 **	4.00 **
REN	5.15 **	4.75 *	4.00 **	4.37 *	3.65 **	5.55 *	5.05 *	4.65 **
GER RYMDKANSLA	6.32 *	5.10 **	4.25 **	6.15 *	5.30 **	6.00 *	5.40 *	4.80 **
SKORRANDE	2.70 **	2.70 **	3.00 **	2.32 **	2.50 *	2.00 *	2.63 **	2.45 **
STARK	5.20 *	5.20 *	4.95 *	4.95 *	4.80 *	5.20 *	4.95 *	5.20 *
STRAV	2.95 **	3.40 *	4.65 **	3.05 **	2.95 **	2.45 *	3.85 **	3.25 **
TORR	2.90 **	3.65 *	4.80 **	2.45 **	3.05 **	2.80 **	3.55 **	4.00 **
TUNN	1.55 *	2.50 *	6.00 **	2.40 **	1.95 **	2.65 **	3.80 **	4.75 **
VASS	3.25 *	3.00 **	6.40 **	3.60 **	2.21 *	2.95 **	4.15 **	4.60 **
VINANDE	3.90 **	3.35 **	4.90 **	3.20 *	2.95 *	3.35 **	3.45 **	2.75 **

PROGRAM 2 ADJECTIVES	HEADPHONES H1	H2	H3	H4	H5	H6	H7	H8
BALANSERAD	5.75 *	5.00 *	3.80 **	5.20 *	5.05 *	5.45 *	5.20 *	5.10 *
BEHAGLIG	4.65 **	4.10 *	2.90 *	4.90 *	4.25 **	4.95 *	4.80 *	4.45 **
BRUSIG	4.50 **	4.10 **	5.10 **	3.00 **	3.20 **	2.65 **	4.10 **	3.55 **
DIFFUS	2.20 *	3.10 **	2.70 *	3.35 *	3.45 **	2.85 **	3.15 **	3.05 **
DOV	2.70 **	2.65 **	1.90 *	3.00 *	4.85 **	4.20 **	4.25 **	3.05 **
FRAMHAVD BAS	4.35 **	2.70 *	2.20 *	4.05 **	5.65 **	4.65 **	4.80 **	3.35 **
FRAMHAVD DISKANT	5.30 **	5.75 *	5.95 **	4.25 **	4.00 **	4.10 **	5.25 **	5.70 **
FRASANDE	4.10 **	4.10 **	4.65 **	3.15 *	2.60 *	3.15 **	4.20 **	3.50 **
FYLLIG	4.75 **	3.95 *	2.00 *	4.90 *	4.20 **	4.95 *	4.25 *	3.30 **
HARD	5.00 **	4.80 **	4.90 **	4.65 **	3.80 **	3.75 **	3.90 **	5.25 **
IHALIG	2.40 *	4.15 **	4.35 **	3.65 **	3.30 **	3.75 **	3.20 **	2.80 **
INSTÄNGD	3.45 **	4.80 **	5.35 **	3.65 **	4.20 **	3.45 **	3.80 *	5.00 **
KLAR	6.00 **	4.60 *	4.85 *	5.25 *	4.60 **	4.70 **	5.35 *	4.45 **
KNASTRANDE	3.40 **	3.10 **	3.80 *	2.30 *	2.25 **	2.65 *	2.70 *	2.90 **
LJUS	4.90 **	5.00 *	6.05 *	4.10 *	3.10 **	3.80 **	4.90 **	5.10 **
MATT	2.30 *	3.20 **	3.65 **	3.15 **	4.20 **	3.95 **	2.95 **	3.40 **
MJUK	3.15 **	2.75 *	2.40 *	3.60 **	3.35 *	4.60 *	3.70 *	3.05 **
MULLRANDE	2.95 *	2.15 **	2.00 *	2.85 *	3.90 *	2.95 *	3.45 *	2.10 *
NASAL	2.80 *	3.25 **	4.25 **	2.60 **	3.20 **	2.80 **	2.90 *	3.53 *
NATURTROGEN	5.65 **	4.70 **	2.95 *	5.25 *	4.85 **	5.35 **	5.20 **	4.30 **
GER NARVAROKANSLA	5.35 **	4.10 **	2.60 **	5.40 **	4.25 **	5.40 **	4.60 **	4.40 **
REN	5.15 *	4.70 *	3.90 **	4.95 *	4.50 *	4.75 *	5.65 *	4.70 **
GER RYMDKANSLA	4.55 **	3.40 **	2.35 **	4.20 **	3.65 **	4.00 **	4.30 **	3.45 **
SKORRANDE	2.60 **	2.40 **	3.40 **	2.60 **	2.30 **	2.25 *	2.80 **	2.85 **
STARK	4.95 **	4.85 *	4.20 *	5.05 **	5.10 *	4.20 *	4.30 *	3.95 *
STRÄV	3.90 **	4.10 *	3.90 *	3.80 **	3.15 **	3.35 *	3.60 **	4.10 *
TORR	3.60 **	4.35 **	5.20 **	3.65 **	2.95 *	3.50 **	3.35 **	4.60 **
TUNN	2.25 **	3.60 **	5.10 **	2.50 *	2.40 **	3.35 **	2.90 **	4.85 *
VASS	4.55 **	4.90 **	6.10 **	3.45 **	2.80 **	2.70 *	3.85 **	4.15 **
VINANDE	4.15 **	3.30 **	4.65 *	3.10 **	3.05 *	2.85 **	4.20 **	3.90 **

TABLE II

Mean ratings for 8 headphones, averaged over 20 subjects, in 30 adjective scales at five music programs.

* standard deviation between 1.00 and 1.99

** standard deviation between 2.00 and 2.50

PROGRAM 3 ADJECTIVES	H1	H2	H3	H4	H5	H6	H7	H8
			HEADPHONES					
BALANSERAD	5.75 *	4.75 *	4.40 **	5.40 *	5.65 *	5.89 *	4.95 **	5.20 *
BEHAGLIG	5.75 **	4.84 **	3.90 *	5.40 **	5.30 **	5.30 **	4.45 **	4.75 **
BRUSIG	6.40 *	5.35 **	6.30 **	4.80 *	4.55 **	4.30 **	6.70 *	5.25 **
DIFFUS	2.95 **	3.75 **	3.74 **	3.05 **	3.70 **	2.90 **	3.75 **	3.50 **
DOV	3.90 *	3.55 **	2.80 **	3.90 **	4.75 **	4.05 **	3.00 **	3.55 *
FRAMHAVD BAS	4.75 **	2.95 *	2.45 *	3.35 *	5.70 *	4.00 **	4.40 **	2.90 *
FRAMHAVD DISKANT	4.50 **	4.50 **	5.30 *	3.15 *	2.80 **	3.40 **	4.40 **	4.35 **
FRASANDE	4.35 **	4.30 **	4.65 **	3.05 **	3.25 **	3.30 **	5.00 **	3.37 **
FYLLIG	5.65 **	4.95 *	3.10 *	5.10 *	6.05 *	5.60 **	4.45 *	4.30 **
HARD	2.95 **	2.40 *	3.80 **	2.15 **	2.15 **	2.00 *	1.95 **	2.37 **
IHALIG	2.60 *	3.00 *	4.20 **	2.25 *	3.65 **	3.80 **	3.25 **	3.75 **
INSTANGD	2.70 **	3.15 **	3.95 **	2.80 **	3.90 **	2.75 **	3.70 **	4.30 **
KLAR	4.75 **	4.50 *	4.05 **	5.05 *	4.50 **	4.90 **	4.30 **	3.84 **
KNASTRANDE	4.35 **	3.55 *	3.75 **	2.50 *	2.55 *	2.00 *	3.45 **	2.90 **
LJUS	3.80 **	4.10 *	5.30 *	4.10 *	2.80 *	3.45 *	4.30 **	4.25 *
MATT	2.40 *	3.75 **	4.00 **	2.80 **	3.80 **	3.20 **	3.90 **	4.50 **
MJUK	4.95 **	5.05 **	3.50 *	5.45 *	5.65 *	5.35 *	4.80 *	4.65 *
MULLRANDE	3.40 **	2.90 *	2.25 **	3.40 *	3.55 **	2.84 *	2.90 *	2.55 **
NASAL	2.45 *	2.60 *	4.25 **	2.80 **	2.70 **	2.70 *	3.40 **	3.25 **
NATURTROGEN	5.15 **	4.55 **	3.80 **	4.89 *	5.15 **	4.50 **	4.15 **	4.20 **
GER NARVAROKANSLA	5.45 **	5.16 *	3.60 *	4.90 **	4.45 **	5.05 **	4.85 **	5.05 **
REN	4.55 **	4.70 *	4.00 **	4.70 *	4.65 *	4.60 **	3.95 **	4.65 **
GER RYMDKANSLA	4.35 **	5.10 **	3.50 **	5.40 **	6.00 *	5.60 *	4.65 *	4.40 **
SKORRANDE	2.15 **	2.00 *	2.50 **	1.60 *	2.05 **	1.85 *	1.60 *	2.25 *
STARK	3.70 **	4.25 *	3.90 *	3.90 **	4.05 **	3.90 *	3.75 **	3.05 *
STRÄV	2.60 *	3.05 *	4.50 *	3.30 *	2.80 **	2.80 **	3.65 **	3.10 **
TORR	3.45 *	2.90 *	4.55 **	3.30 **	2.75 **	2.75 **	4.35 **	4.40 **
TUNN	1.90 *	3.10 *	5.50 *	2.40 **	1.95 *	2.50 **	3.50 **	3.37 **
VASS	3.35 **	2.70 **	5.10 **	2.65 **	1.70 **	2.00 *	2.55 *	2.70 **
VINANDE	4.10 **	4.35 **	3.90 **	3.30 **	3.05 **	2.80 **	4.35 **	3.30 **

TABLE II

Mean ratings for 8 headphones, averaged over 20 subjects, in 30 adjective scales at five music programs.

* standard deviation between 1.00 and 1.99

** standard deviation between 2.00 and 2.50

PROGRAM 4 ADJECTIVES	HEADPHONES H1	H2	H3	H4	H5	H6	H7	H8
BALANSERAD	5.65 *	4.60 *	4.00 *	5.40 *	4.25 **	5.00 *	4.85 *	5.20 **
BEHAGLIG	5.50 **	4.25 *	3.05 *	4.55 **	4.15 **	5.35 *	4.55 *	4.80 *
BRUSIG	4.25 **	3.95 **	4.25 **	2.80 **	2.90 *	2.15 **	3.50 **	2.55 *
DIFFUS	2.60 *	2.75 **	2.60 **	3.50 **	4.95 **	3.25 *	3.15 **	2.50 **
DOV	3.25 *	3.35 *	1.75 **	3.65 **	6.00 *	4.60 **	2.95 **	3.20 **
FRAMHAVD BAS	3.85 **	2.75 **	2.10 *	3.65 **	5.00 **	3.95 **	2.95 *	2.85 **
FRAMHAVD DISKANT	4.65 **	4.40 **	6.90 *	3.80 **	2.85 **	3.35 **	5.20 **	4.30 **
FRASANDE	3.80 **	3.45 **	4.90 **	2.65 **	2.45 **	2.70 **	3.60 **	2.95 *
FYLLIG	5.60 *	4.50 *	2.84 *	5.10 *	4.70 **	5.20 **	4.50 *	4.10 *
HARD	3.60 **	3.65 **	5.15 **	3.65 **	3.15 *	3.60 **	3.10 **	4.55 **
IHALIG	2.50 *	3.65 **	4.00 **	2.75 **	4.10 **	3.15 **	3.00 **	3.20 **
INSTANGD	1.90 *	3.80 **	3.95 **	3.65 **	5.00 **	3.10 **	2.70 **	3.60 **
KLAR	5.35 **	4.30 **	4.40 *	4.80 **	3.30 *	4.95 *	5.00 *	4.70 **
KNASTRANDE	2.95 *	2.80 *	2.68 **	2.75 *	2.05 *	2.00 *	3.15 **	2.10 *
LJUS	4.25 **	4.60 *	5.65 *	3.60 **	3.05 **	3.95 *	4.60 **	4.50 *
MATT	2.75 *	2.70 **	3.55 **	3.15 *	4.70 **	2.90 **	3.40 **	3.25 **
MJUK	3.25 **	3.55 *	1.90 *	4.55 *	4.20 *	4.30 *	3.45 **	3.75 **
MULLRANDE	3.50 **	2.75 **	2.50 **	3.50 **	5.20 **	3.05 *	3.00 **	2.65 **
NASAL	3.00 **	2.90 *	4.40 **	2.85 **	2.90 **	3.25 **	3.05 **	3.50 **
NATURTROGEN	5.65 **	5.10 *	3.90 **	4.65 **	3.45 **	4.45 **	4.10 **	4.30 **
GER NARVAROKANSLA	5.60 **	4.75 *	3.45 **	4.75 **	3.50 **	4.85 **	4.50 **	4.10 **
REN	4.95 **	4.65 **	4.25 *	4.55 **	3.30 *	4.30 *	4.45 **	4.95 **
GER RYMDKANSLA	5.70 *	4.70 *	3.70 **	4.35 **	4.00 **	4.65 **	5.15 **	4.20 **
SKORRANDE	2.95 **	2.50 *	3.47 **	2.10 **	2.35 **	2.10 **	2.35 **	2.75 **
STARK	5.35 **	4.75 *	5.50 *	4.85 **	5.00 *	4.80 **	4.65 **	4.45 *
STRÄV	3.35 *	4.21 **	4.70 **	3.50 **	3.15 **	3.15 **	4.21 **	3.75 *
TORR	3.60 **	3.90 **	4.70 **	3.00 **	3.90 **	2.80 **	4.10 **	4.10 **
TUNN	2.55 *	2.90 **	5.20 **	3.60 **	2.85 **	2.75 **	3.35 **	3.95 **
VASS	3.70 **	4.20 **	6.55 *	3.85 **	2.20 **	3.25 **	4.60 **	4.45 **
VINANDE	4.20 **	4.11 **	4.35 **	3.65 **	3.30 *	3.15 **	3.95 **	3.60 **

TABLE II

Mean ratings for 8 headphones, averaged over 20 subjects, in 30 adjective scales at five music programs.

* standard deviation between 1.00 and 1.99

** standard deviation between 2.00 and 2.50

PROGRAM 5 ADJECTIVES	H1	H2	H3	H4	H5	H6	H7	H8
	HEADPHONES							
BALANSERAD	6.35 *	5.85 *	3.95 **	5.05 *	5.50 **	5.55 *	4.84 **	4.75 *
BEHAGLIG	5.50 **	5.05 **	3.05 **	5.30 *	4.45 **	5.20 *	5.05 *	4.15 **
BRUSIG	3.30 **	2.40 **	3.45 **	2.85 **	2.55 **	2.30 *	2.35 *	2.30 **
DIFFUS	2.00 **	2.90 **	3.05 **	2.85 **	4.20 **	3.05 **	2.20 *	2.75 **
DOV	2.32 *	2.95 **	2.10 **	3.30 *	4.80 **	4.35 **	3.10 **	2.80 **
FRAMHAVD BAS	4.60 *	2.90 *	2.80 *	3.85 *	4.45 **	3.60 **	3.53 **	3.20 *
FRAMHAVD DISKANT	5.40 **	4.00 **	6.40 *	4.15 **	3.35 **	3.50 **	5.60 *	5.65 *
FRÄSANDE	3.30 **	3.05 **	4.20 **	2.70 **	2.70 **	3.10 **	3.60 **	3.35 **
FYLLIG	5.80 *	5.15 *	3.25 **	5.26 *	4.85 *	5.05 *	4.65 *	4.40 *
HARD	4.85 *	4.50 *	6.35 *	3.80 **	4.10 **	3.85 **	5.00 **	5.95 *
IHALIG	1.80 *	3.26 **	3.35 **	3.05 **	3.50 **	3.30 **	2.75 *	3.35 **
INSTANGD	1.70 *	2.40 *	4.60 **	2.80 **	4.30 **	3.50 **	2.85 *	3.85 **
KLAR	5.80 **	5.35 **	4.85 **	4.75 **	4.75 **	5.05 *	5.60 *	5.40 **
KNASTRANDE	2.50 *	2.05 *	3.05 **	2.20 **	2.50 **	1.80 *	3.00 **	2.45 **
LJUS	5.85 **	5.00 *	5.80 *	4.30 *	3.20 **	3.65 *	4.35 *	4.60 *
MATT	1.35 *	2.80 **	2.50 **	3.00 *	3.89 **	3.10 **	2.45 **	2.85 **
MJUK	2.75 *	2.30 *	1.30 *	3.35 **	4.00 **	3.55 *	2.45 *	2.70 *
MULLRANDE	3.55 **	3.70 **	1.95 *	3.50 **	4.75 **	3.65 **	3.75 **	2.35 *
NASAL	1.95 *	2.70 *	4.45 **	2.65 **	2.90 *	2.80 **	3.35 **	3.25 **
NATURTROGEN	6.00 **	5.35 *	3.75 **	5.20 **	3.95 **	5.15 *	5.15 *	4.85 **
GER NARVAROKÄNSLA	6.20 **	4.85 *	3.75 **	5.25 **	4.70 **	5.25 *	5.10 **	4.85 **
REN	5.75 **	5.35 **	4.50 **	4.80 *	4.20 *	4.95 *	5.20 *	4.95 **
GER RYMDKÄNSLA	6.21 **	5.75 *	4.60 **	4.95 **	4.40 **	4.90 **	5.53 *	5.45 **
SKORRANDE	2.47 **	2.84 **	3.21 **	2.75 **	3.45 **	2.55 *	2.95 **	2.90 **
STARK	6.45 *	5.80 *	6.45 *	5.25 *	5.05 *	5.45 **	5.80 *	5.80 *
STRAV	3.55 **	3.28 **	5.15 **	3.00 *	2.89 *	3.15 **	4.00 *	4.15 **
TORR	2.50 **	3.20 **	5.05 **	3.50 **	3.32 **	3.26 **	3.60 **	4.00 **
TUNN	1.80 **	2.85 **	4.20 **	2.75 **	2.30 **	2.55 **	2.80 **	3.45 **
VASS	4.90 **	4.55 **	6.80 *	3.80 **	3.30 **	3.65 **	5.05 **	5.05 **
VINANDE	3.65 *	4.00 **	3.90 **	2.95 **	3.30 **	2.84 **	3.15 **	3.85 **

TABLE II

Mean ratings for 8 headphones, averaged over 20 subjects, in 30 adjective scales at five music programs.

* standard deviation between 1.00 and 1.99

** standard deviation between 2.00 and 2.50

TABLE III MEAN RATINGS FOR EIGHT HEADPHONES IN 30 ADJECTIVE SCALES, AVERAGED OVER SUBJECTS AND PROGRAMS

ADJECTIVES	HEADPHONES H1	H2	H3	H4	H5	H6	H7	H8
BALANSERAD	5.97	5.04	4.08	5.29	5.07	5.50	4.91	4.92
BEHAGLIG	5.38	4.57	3.21	5.17	4.63	5.39	4.62	4.37
BRUSIG	4.58	4.00	4.54	3.21	3.21	2.72	4.17	3.26
DIFFUS	2.54	3.28	3.20	3.17	4.16	3.02	3.26	3.02
DOV	3.02	3.16	2.14	3.37	5.24	4.07	3.28	3.03
FRAMHÄVD BAS	4.31	2.78	2.24	3.56	5.23	3.81	3.68	2.83
FRAMHÄVD DISKANT	4.83	4.61	6.19	3.97	3.17	3.56	5.00	5.05
FRÄSANDE	3.81	3.58	4.53	2.73	2.70	2.89	4.05	3.22
FYLLIG	5.53	4.77	2.66	5.17	5.09	5.22	4.39	3.88
HÅRD	3.88	3.79	5.14	3.64	3.12	3.32	3.49	4.55
IHÅLIG	2.39	3.36	4.10	2.89	3.58	3.28	3.09	3.39
INSTÄNGD	2.43	3.43	4.49	3.21	4.50	3.12	3.07	4.17
KLAR	5.39	4.60	4.45	5.02	4.03	4.93	4.93	4.68
KNASTRANDE	3.21	2.83	3.31	2.31	2.24	2.05	3.07	2.48
LJUS	4.31	4.63	5.79	4.04	3.01	3.93	4.63	4.87
MATT	2.20	3.11	3.38	3.04	4.08	3.19	3.20	3.48
MJUK	3.71	3.50	2.23	4.24	4.32	4.55	3.56	3.35
MULLRANDE	3.24	2.73	2.08	3.25	4.32	3.00	3.18	2.39
NASAL	2.58	2.87	4.44	2.74	3.00	2.84	3.27	3.55
NATURTROGEN	5.63	4.89	3.50	5.05	4.23	4.98	4.67	4.31
GER NÄRVAROKÄNSLA	5.77	4.81	3.38	5.05	4.10	5.20	4.72	4.48
REN	5.11	4.83	4.13	4.67	4.06	4.83	4.86	4.78
GER RYMDKÄNSLA	5.52	4.81	3.68	5.01	4.67	5.03	5.01	4.46
SKORRANDE	2.57	2.49	3.12	2.27	2.53	2.15	2.48	2.64
STARK	5.13	4.97	5.00	4.80	4.80	4.71	4.69	4.49
STRÄV	3.27	3.61	4.58	3.33	2.99	2.98	3.86	3.67
TORR	3.21	3.60	4.86	3.18	3.19	3.02	3.89	4.22
TUNN	2.01	2.99	5.20	2.73	2.29	2.76	3.27	4.07
VASS	3.95	3.87	6.19	3.47	2.44	2.91	4.04	4.19
VINANDE	4.00	3.82	4.34	3.24	3.13	3.00	3.82	3.48

TABLE IV. Results from significance tests in each adjective scale (x means significance at 5% level, xx at 1% level, and xxx at 0.1% level), inter-rater reliability, and "Honestly Significant Difference" (HSD) at 5% level for pairwise comparisons between headphones.

Adjective	Significances: Programs	Significances: Headphones	Significances: Programs x Headphones	Reliability	HSD
Balanserad ("Balanced")		xxx		.67	.84
Behaglig ("Pleasant")		xxx		.74	1.04
Brusig ("White noise/Hissing")	xxx	xxx	x	.91	.84
Diffus ("Diffuse")	x	xxx		.60	.98
Dov ("Dull")		xxx	x	.83	.93
Framhävd bas ("Emphasized bass")	xxx	xxx	x	.88	.79
Framhävd diskant ("Emphasized treble")	xx	xxx		.86	.99
Fräsande ("Hissing")	x	xxx		.77	.86
Fyllig ("Full/-toned")	xx	xxx		.85	.92
Hård ("Hard")	xxx	xxx	x	.86	.94
Ihålig ("Hollow")		xxx	x	.59	1.06
Instängd ("Shut up/Closed")	x	xxx	x	.76	1.10
Klar ("Clear")	x	xx	x	.56	1.02
Knastrande ("Crackling/Crunching")	xx	xxx	x	.72	.75
Ljus ("Bright/Light")	x	xxx		.81	.90
Matt ("Faint/Feeble")	x	xxx		.60	.98
Mjuk ("Soft")	xxx	xxx		.86	.95
Mullrande ("Rumbling")		xxx		.78	.70
Nasal ("Nasal")		xxx		.67	.82
Naturtrogen ("True to nature")		xxx	x	.70	.96
Närvarokänsla ("Feeling of presence")		xxx		.67	1.14
Ren ("Pure/Clean")		xx		.46	.87
Rymdkänsla ("Feeling of space")	xxx	xxx	x	.74	.92
Skorrande ("Jarring/Grating")	x	x		.39	.81
Stark ("Loud")	xxx		xx	.85	.69
Sträv ("Harsh")		xxx		.63	.91
Torr ("Dry")		xxx		.65	1.02
Tunn ("Thin")		xxx	x	.85	1.19
Vass ("Sharp/Keen")	xxx	xxx		.88	1.03
Vinande ("Whistling/Whizzing")		xxx		.53	.81

TABLE VII. Adjectives showing highest correlations with the "Pleasant" and "True to nature" scales.

Pleasant		True to nature	
.89	Full/-toned/	.87	Feeling of presence
.83	Balanced	.80	Pure/Clean
-.83	Nasal	.78	Balanced
-.80	Dry	-.76	Nasal
-.79	Thin	-.72	Shut up/Closed
.77	Feeling of presence	.71	Full/-toned
-.75	Harsh	-.69	Hollow
-.68	Hollow	-.69	Thin
.68	Feeling of space	-.64	Dry
-.66	Shut up/Closed	.63	Clear
-.66	Sharp/Keen	-.60	Faint/Feeble
.65	Soft	.59	Feeling of space
-.63	Emphasized treble	-.51	Diffuse
-.59	Bright		
-.52	Hard		
-.52	Jarring/Grating		
-.50	Hissing		

Correlation between "Pleasant" and "True to nature" = .74