



TECHNICAL AUDIOLOGY

Report TA No. 98
September 1980

CLINICAL ASSESSMENT OF PERCEIVED SOUND QUALITY IN
HEARING AIDS

Alf Gabrielsson, Björn Hagerman, Caroline Berg,
Anders Ovegård and Lars Änggård

Report TA No. 98
September 1980

CLINICAL ASSESSMENT OF PERCEIVED SOUND QUALITY IN
HEARING AIDS

Alf Gabrielsson	
Björn Hagerman	
Caroline Berg,	Danderyd Hospital
Anders Ovegård,	Danderyd Hospital
Lars Änggård,	Danderyd Hospital

Material from this report may be reproduced provided
the sources are indicated.

From the Department of Technical Audiology
Karolinska Institutet
KTH
S-100 44 STOCKHOLM, Sweden

Tel: 46-8-11 66 60

CLINICAL ASSESSMENT OF PERCEIVED SOUND QUALITY IN HEARING AIDS

Alf Gabrielsson
Björn Hagerman
Caroline Berg, Danderyd Hospital
Anders Ovegård, Danderyd Hospital
Lars Ånggård, Danderyd Hospital

ABSTRACT

12 patients made sound quality judgments for three hearing aids preliminarily chosen to fit their hearing impairment. Ten programs and nine adjective scales were used. The reliability of the patients' judgments was considered as sufficient, but the sensitivity of the scales should be improved in order to make finer discriminations between the hearing aids. Ratings made at home correlated rather well with ratings from the more controlled tests at the clinic.

This work was supported by the National Swedish Board for Technical Development.

Table of Contents

	ABSTRACT	
1	INTRODUCTION.....	1
2	METHODS.....	1
2.1	Subjects.....	1
2.2	Listening room and technical equipment.....	1
2.3	Stimuli.....	2
2.4	Definition of scales.....	2
2.5	Procedure.....	3
2.6	Statistical treatment.....	3
3	RESULTS.....	3
3.1	Reliability.....	3
3.2	Relations between Total impression and the other scales.....	4
3.3	Factor analysis on the judgment scales.....	4
3.4	The patient's final choice of hearing aid.....	4
3.5	Comparison of the judgments made at the clinic and at home.....	6
4	DISCUSSION.....	6
5	REFERENCES.....	7

1 INTRODUCTION

During the last ten years research about perceived sound quality in sound reproduction systems has been made in the department of Technical Audiology (Gabrielsson 1979, Gabrielsson & Sjögren 1979a, 1979b). The aims of this research have mainly been to survey different perceptual dimensions in the perceived sound quality, but also to tentatively relate these dimensions to the physical characteristics of the systems, to determine the positions of different systems in different perceptual dimensions and to study the relations between the perceptual dimensions and overall judgments of the sound quality.

As this work has been mostly explorative, it is now necessary to examine the validity and the reliability of the perceptual dimensions discovered in this earlier research. This is done in different ways. One of them is to use the scales as a tool for hearing aid fitting. This has been tried during 1978-1979 in the Audiological Department, Danderyd Hospital as described in the following.

The purpose of our work was to investigate the possibility of making sound quality judgments of hearing aids within the clinical practice. A report giving all details in methods and results is written in Swedish. The present report gives only a summarizing description to demonstrate the type of work going on.

2 METHODS

2.1 Subjects

12 hearing impaired subjects participated in the investigation. Most of them had a sensory-neural impairment. Mean age, mean speech reception threshold and mean speech discrimination were 63 years, 54 dB and 87 %, respectively. Most of them were experienced hearing aid users.

2.2 Listening room and technical equipment

The sound quality judgments took place in a room normally used for hearing aid fitting, with a reverberation time of 0.50 to 0.25 seconds from 250 to 4000 Hz and a volume of 33 m³. The stimulus tape was played on a Tandberg 10XD tape recorder, which fed the amplifier Quad 303 S over an attenuator. The loudspeaker Sonab OD 11 hang on the shorter wall of the room 1.85 meter above the floor directed towards the listener, who was sitting at 1.5 meter distance facing the loudspeaker. The frequency curve of the room was checked with pink noise in octave bands, and the differences of the sound levels between different bands was less than 3 dB from 250 to 4000 Hz.

2.3 Stimuli

The stimuli consisted of 10 tape-recorded programs of about one minute each. Shorter samples of the first five programs were initially presented several times for adjusting the loudness to a comfortable level. The programs were:

1. Male voice
2. Female voice
3. Male voice in cafeteria noise
4. Song (female voice and organ)
5. Piano music
6. March music
7. Washing up the dishes
8. Typewriter
9. Closing a door several times
10. Traffic noise

The sound levels of the programs were intended to be the same as when they were recorded. This was checked both by careful listening of the authors and by measuring with a sound level meter. The three music pieces were copied from records, and the other programs were recorded with an 1" B&K condenser microphone and a Nagra tape recorder.

2.4 Definition of scales

The sound quality of the respective hearing aids was judged for each of the ten programs on nine different rating scales, each graded from 1 to 5. The scales 1-8 refer to the eight perceptual dimensions found in analyses of earlier experiments. The scales were labeled as follows:

<u>English</u>	<u>Swedish</u>
1. Softness	Mjukhet
2. Distinctness/Clearness	Tydlighet
3. Fullness	Fyllighet
4. Brightness	Ljushet
5. Nearness	Närhet
6. Feeling of space	Rymdkänsla/Rymlighet
7. Loudness	Ljudstyrka
8. Disturbances/Hissings	Störningar/Brus
9. Total impression	Sammanfattande intryck

Definitions of the scale steps were given according to the principle that 5 represented very much of the quality in question (for instance, very soft), 4 rather much (rather soft), 3 a midway position, 2 rather much of the opposite quality (for instance, rather sharp) and 1 very much of the opposite quality (very sharp). These definitions were explicitly given for each of the scales on the lists that the patients used for doing their ratings. The judgments were made for each of the scales during each program and always in the same order. The patient marked the answers on one list for each program.

2.5 Procedure

The patient visited the clinic 4 or 5 times with 2-3 weeks between each visit. The first visit was used for making tone and speech audiogram and then a preliminary choice of a hearing aid called A was made. The sound quality judgments started at the first or the second visit using the patient's old aid and the hearing aid A. Then the patient left the clinic with hearing aid A, in order to perform a practical evaluation at home, and with a detailed questionnaire to answer regarding this aid. The questionnaire asked for the patient's hearing ability in different listening situations but also for some practical details regarding, for example, the easiness to handle the hearing aid. The same procedure was repeated at the next visit, when a second hearing aid (here called B) was selected with respect to the patient's report regarding hearing aid A. The sound quality judgments were now made for the hearing aids A and B, and the patient went home with the aid B and a new (but equal) questionnaire. On the fourth visit a third hearing aid was selected considering the patient's report on the previous two hearing aids. Quality judgments were made for all three aids A, B and C, and the patient went home with a questionnaire and the aid C. Finally the last visit was used for quality judgments of all three aids A, B and C again, and the patient decided which of them he preferred.

2.6 Statistical treatment

The data from the 12 patients were thoroughly analysed by several statistical methods, for example, calculation of means, standard deviations and correlations, analysis of variance, regression analysis and factor analysis. The analyses were made mainly on an individual basis since the hearing impairments and the selected hearing aids differed between different patients.

3 RESULTS

The individual results not unexpectedly reveal rather varying patterns and are not shown here (except in Table 1). Only the most important general tendencies are mentioned below.

3.1 Reliability

The reliability of the patients' judgments is generally considered as satisfying as shown by various indices of reliability and also more directly by the fact that test-retest judgments seldom differ more than one scale unit. The judgments on the scales Softness and Loudness seem to be somewhat more reliable than the judgments on the other scales. Some patients used constant or almost constant judgments in a certain scale or for a certain hearing aid, for example, 1 for the Disturbance or 5 for the Total impression. This may give an impression of high reliability but it may as well reflect memory influences, reluctance for differentiating etc.

3.2 Relations between Total impression and the other scales

The relations between the judgments of the Total impression and the judgments of the other scales vary considerably among the patients. A review of the whole material suggests that Clearness is the scale most highly correlated to the Total impression, followed by Feeling of space, Fullness, Brightness and Softness. These correlations are positive, while Disturbances and sometimes also Nearness and Loudness show negative correlations to the Total impression.

3.3 Factor analysis on the judgment scales

Factor analysis has been made in order to examine if some of the scales covary to such a degree that they may be based on a common factor. The results differ between the patients also here, but some groupings return rather often and are verified in the factor analysis of the whole material (that is, made over all the 12 patients). Thus Clearness, Fullness, Feeling of space and Brightness are often grouped together to one factor, while another factor consists of the scales Softness and Loudness in opposition to each other. Sometimes Brightness appears as a certain factor as well as Disturbances. Nearness occurs sometimes together with Clearness, Fullness etc., but sometimes with Loudness.

3.4 The patient's final choice of hearing aid

In the final choice of hearing aid five patients (no. 1,4,7,10 and 12) chose the aid ranked highest in the Total impression scale. Six other patients (no. 3, 5, 6, 8, 9 and 11) chose the aid ranked as the next highest in Total impression. (Patient no. 2 made no differentiation at all in the Total impression). For these six patients the difference between the two aids ranked highest in Total impression is less than 0.27, which is nonsignificant, see Table 1.

A comparison between the chosen aid and the patient's old hearing aid indicates that the new one is judged as better in most cases. However, the old aid was judged only once, that is, the first time the patient tried the procedure, and this judgment of the old aid should thus be more or less regarded as a training session.

Pat no	Aid	Soft	Clear	Full	Bright	Near	Space	Loud	Distu	Tot.imp
1	A	2.55	4.30	4.30	3.32	4.22	4.45	3.52	1.05	4.13
	B	2.33	4.40	4.37	3.27	4.40	4.47	3.63	1.00	4.40 *
	C	2.40	4.00	4.05	2.75	4.15	4.10	3.60	1.00	3.80
2	A	2.85	5.00	4.77	4.32	5.00	4.60	4.00	1.07	5.00 *
	B	2.80	4.97	4.80	4.30	5.00	4.57	4.00	1.00	5.00
	C	2.60	5.00	4.85	3.65	5.00	4.45	4.05	1.00	5.00
3	A	2.38	4.82	4.65	3.63	4.85	4.57	4.22	3.75	4.67
	B	2.43	4.80	4.67	3.60	4.80	4.60	4.17	3.73	4.73 *
	C	2.50	4.95	4.90	3.55	4.75	4.60	4.35	4.00	5.00
4	A	3.00	4.90	4.60	3.88	4.32	4.70	3.30	1.00	4.70
	B	3.17	4.97	4.90	4.10	4.50	4.83	3.33	1.00	4.87
	C	3.55	5.00	4.95	4.40	4.25	4.95	3.20	1.00	5.00 *
5	A	2.70	4.60	4.40	3.92	4.57	4.20	3.82	2.30	4.13
	B	2.80	4.67	4.43	3.87	4.60	4.27	3.93	2.17	3.97 *
	C	2.55	4.65	4.20	4.00	4.60	4.25	3.95	2.30	3.70
6	A	3.15	4.63	4.67	4.10	4.80	4.47	4.02	1.32	4.38 *
	B	2.90	4.77	4.70	3.83	4.93	4.47	4.30	1.20	4.33
	C	3.45	4.70	4.70	4.20	4.90	4.45	3.85	1.25	4.40
7	A	2.42	4.30	3.27	2.92	4.70	3.50	3.85	1.10	4.82
	B	2.53	4.20	2.80	3.20	4.60	3.50	3.70	1.00	5.00 *
	C	2.60	4.45	2.65	3.20	4.50	3.40	3.85	1.00	5.00
8	A	3.27	4.27	3.97	3.40	4.67	4.02	3.40	1.02	4.20 *
	B	3.07	4.23	3.67	3.23	4.67	3.93	3.40	1.10	3.73
	C	3.60	4.80	4.45	3.45	4.95	4.30	3.35	1.05	4.25
9	A	2.60	3.57	3.10	3.17	3.80	3.22	3.42	2.22	3.20
	B	2.83	3.83	3.17	3.33	3.90	3.40	3.63	1.13	3.90
	C	3.20	3.90	3.30	3.00	3.65	3.65	3.20	1.10	3.80 *
10	A	2.90	4.67	3.05	3.88	4.25	3.92	3.92	1.00	3.72
	B	2.97	4.70	2.90	3.77	4.20	3.90	3.93	1.00	3.73
	C	3.05	4.60	2.80	3.60	4.20	3.80	3.80	1.00	3.75 *
11	A	2.67	4.70	4.65	3.85	4.47	4.50	3.80	1.07	4.72
	B	2.83	4.73	4.70	4.13	4.33	4.43	3.30	1.03	4.87 *
	C	2.90	4.70	4.55	4.10	4.50	4.60	3.65	1.00	5.00
12	A	3.02	4.45	4.30	3.10	3.65	3.77	3.17	1.05	4.82
	B	2.93	4.70	4.33	3.00	3.87	3.97	3.07	1.00	5.00 *
	C	3.05	4.60	4.00	3.15	3.50	3.55	3.10	1.00	4.90

Table 1. Results of the judgments in the different scales averaged over the different programs and over the different test-sessions. Note that the hearing aids A,B and C are different for different patients.

* Indicates the aid chosen by the patient.

3.5 Comparison of the judgments made at the clinic and at home

Comparisons were made between the judgments made at home and at the clinic regarding the scales Clearness and Pleasantness (Pleasantness judged at home was compared with Total impression and with Softness at the clinic). Since the stimulus situations are very different at the clinic and at home, it could be expected that there should be considerable differences between the corresponding judgments. However, the difference seldom exceeds one scale unit which means that there is a fair correspondence between the ratings at the clinic and ratings made at home.

4 DISCUSSION

The validity of the chosen scales was confirmed in the investigation. All eight perceptual scales show (positive or negative) correlations with the scale Total impression. The fact that these correlations differ between different patients is probably due to the different hearing impairments and the different hearing aids used by the patients. Expressions as "clear sound", "shrill", "soft", "bright" etc. return spontaneously in the patients' own descriptions of the sound quality.

The whole procedure described here is rather time consuming. It is therefore desirable to reduce the number of programs and/or the number of scales. The possibility of reducing the number of programs was studied by counting the number of significant differences between hearing aids within each single program and patient. However, the number of significances did not differ considerably between the programs, implying that no program may be removed solely on this basis. It might be mentioned that the male voice in cafeteria noise ranked highest in this comparison.

The possibility of reducing the number of scales is presently left open. The factor analysis indicates some groupings of the scales, for example Clearness together with Fullness and Feeling of space, or Softness versus Loudness. These results, however, do not necessarily imply that these qualities are inter-related. They probably represent different perceptual dimensions, but the programs and hearing aids used perhaps have such characteristics that a system which sounds clear also sounds full etc. More data are needed to justify the possible elimination of one or more scales.

The main problem with the judgment scales, however, apply to their reliability and sensitivity. The reliability seems to be satisfying when conventional measures are used. However, it often seems that the reliability rather refers to differences between the programs than to the differences between the hearing aids. The differences between the programs are thus more obvious than the differences between the hearing aids. One reason for this is probably the fact that the preliminary choice of the three hearing aids is made according to the patient's hearing impairment, which means that the differences between the aids may be restricted from the beginning. It is now necessary to get more sensitive scales which are able to capture smaller differences in the different perceptual dimensions. This may be done by including more scale steps and by giving more detailed definitions for each step etc.

Another way is to change the presentation of the stimuli. It is known from other listening tests that our memory of sound quality is rather poor. To detect small differences it is therefore easier to listen to the pieces in a relatively fast sequence and to do pairwise comparisons. To make fast comparisons of two hearing aids in the way they are normally used (i.e. with earmoulds) is, however, not possible to do. So far we have no solution to this problem. Another problem is the fact that some patients have stuck to a fixed value for certain scales in almost all cases, especially for the scales Disturbance and Total impression. Memory effects may have played a role here since there was a constant order of the programs and of the scales, and since the patient also knew which hearing aid he was listening to.

The patient's final choice of the hearing aid is not only influenced by the perceived sound quality. Other important factors may be the size of the aid, its appearance, easiness in handling etc. Therefore one should not expect a perfect correlation between the ratings of perceived sound quality and the actual choice of a hearing aid. It is desirable to get information from the patient regarding the weights he attaches to the perceived sound quality for different programs and scales as well as to various other factors as those mentioned above. Another question is whether the final choice is influenced by the order in which the different hearing aids were used in actual practice (that is, A-B-C). The present data do not permit any definite answers to this question, which is left open for further investigation.

5 REFERENCES

- Gabrielsson, A. (1979). Dimension analyses of perceived sound quality of sound-reproducing systems. Scand J Psychol 20, 159-169.
- Gabrielsson, A. & Sjögren, H. (1979a). Perceived sound quality of sound-reproducing systems. J Acoust Soc Am 65, 1019-1033.
- Gabrielsson, A. & Sjögren, H. (1979b). Perceived sound quality of hearing aids. Scand Audiol 8, 159-169.