

Comparison of Two Hearing Aid Receiver-Amplifier Combinations Using Sound Quality Judgments

Catherine V. Palmer, Mead C. Killion, Laura A. Wilber, and William J. Ballad

Sound quality judgments were obtained on two binaural pairs of laboratory hearing aids with similar battery drain. One pair had a traditional low-current-drain "starved Class A" output stage. The other had a new (at the time) "Class D" output stage. Speech and music reproduction was rated, for seven input levels between 70 and 100 dB SPL, on an overall quality scale by juries of normal-hearing and hearing-impaired subjects. The same subjects also were asked to assign a dollar value to each condition by answering the question "What would you pay for a hearing aid that sounded like that?" Both subject groups rated the hearing aids with the Class D output stage as having superior sound quality across a variety of input levels and test materials, consistent with objective distortion measurements. On the average, each one-percentage point increase in sound quality rating corresponded to a \$6.75 increase in perceived value in these experiments.

(*Ear & Hearing* 1995;16;587-598)

Gabrielsson, Hagerman, Berg, Ovegard, and Anggard (1980) reported that "clarity" was the most highly correlated parameter to the overall impression of a hearing aid for 12 hearing-impaired subjects. Hagerman and Gabrielsson (1984) further reported that "sound quality" was ranked as the most important property of hearing aids by the majority of hearing-impaired patients in their study. Hagerman and Gabrielsson (1985) later had a group of potential hearing aid users rate the importance of a variety of features of the hearing aid such as sound quality, ease of use, size, and battery drain. Sound quality was rated as being more important than any other factor. Agnew (1988) cited nonlinear distortion of the hearing aid processed signal as a cause of subjective judgments of poor sound quality. Preves and Newton (1989) report that poor sound quality judgments often lead to nonacceptance of a hearing aid.

Preves (1989) cites inadequate "headroom" as one cause for the nonlinear distortion that occurs when the input signal is sufficiently intense to drive the hearing aid into saturation. Inadequate headroom can result from low SSPL90 combined with high

gain, which causes clipping and other types of nonlinear distortion at high input levels. A more subtle problem, inadequate headroom only at high frequencies (Killion, 1980), is usually an indirect result of using what might be called a "starved Class A" output stage in a hearing aid. Until recently, this was the only practical choice, a choice dictated by the low battery current required in the smallest hearing aids to provide an acceptable battery life with the smallest batteries. The other possible choice, a Class B output stage, was not a practical choice for canal aids because of the large physical size of the capacitors required by Class B circuits at the time this study was undertaken.

Hearing aid distortion is most often measured according to the ANSI S3.22-1987 standard, which requires measurement of total harmonic distortion for an input SPL of 65 dB at 500, 800, and 1600 Hz. When inadequate headroom is limited to the high-frequency region, the harmonic distortion measured at the ANSI test frequencies and levels may appear acceptable for a hearing aid that is nonetheless described by listeners as having poor sound quality. Thus, we have seen the recent rediscovery of the high-frequency intermodulation distortion problem first reported by Peterson (1951), who described a hearing aid that looked good on paper but sounded awful. Peterson found that a high-frequency harmonic distortion measurement was of little value because the frequencies of the harmonic distortion products exceeded the high-frequency cutoff of the hearing aid, making it impossible to measure the distortion accurately. Peterson found that a high-frequency intermodulation distortion measurement (of the lower-frequency distortion products that were created below the frequency of his two high-frequency test tones) correlated well with the subjective complaints. More recently, it has become popular to measure intermodulation distortion using a coherence-function analysis, which provides a combined measure of the noise and distortion in a hearing aid. The result of such an analysis is a number from 0 to 1 that indicates what portion of the output signal is due only to the input signal (Burnett, Corliss, & Nedzelnitsky, 1982). Although the coherence-function analysis provides no way of separating out the various defects, it can be argued that both noise and nonlinear distortion products

Northwestern University (C.V.P., L.A.W.), Evanston, Illinois; Etymotic Research, Inc. (M.C.K.), Elk Grove Village, Illinois; and Knowles Electronics (W.J.B.), Itasca, Illinois

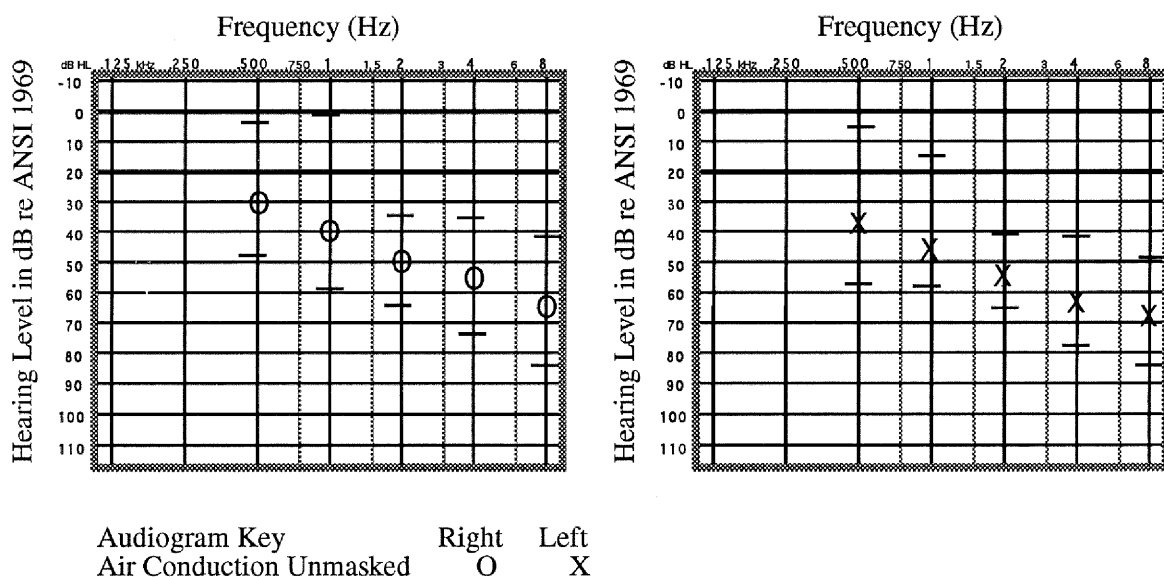


Figure 1. Mean threshold and threshold ranges for 11 hearing-impaired subjects.

act somewhat similarly in masking the desired signal and thus from a practical standpoint the lack of separation makes little difference.

The purpose of this study was to quantify the subjective sound quality of two hearing aids, one with a traditional starved Class A output stage and one with a new (at the time the study was undertaken) Class D output stage. To reflect the variability of real-life listening, a variety of input intensity levels and program materials (i.e., speech and music) were selected. Both normal-hearing and hearing-impaired subjects provided quality and perceived-financial-value judgments in our attempt to quantify the practical significance of any quality differences. A secondary goal was to examine the correlation, if any, between subjective sound-quality judgments and the coherence-function-measured noise and distortion of the hearing aids.

It is important to note that this was not intended to be a comparison of Class A and Class D amplifiers per se. Starved Class A amplifiers have become traditional because providing a Class A amplifier enough current for proper operation would result in a battery life of at most 1 or 2 days. If a dramatically higher battery drain was practical in Class A amplifiers, then the two types of experimental hearing aids in the present study could have been made practically indistinguishable, at most requiring a slightly different Class A amplifier design. Indeed, Killion (1979) used such Class A amplifiers in his demonstration that experimental hearing aids could be constructed with a fidelity comparable to expensive high-fidelity systems. It was, however, precisely the real-life comparison that was of interest here. Class A amplifiers with 1 or 2 days' battery life, such as Killion used in his study, were considered to be of only academic interest.

METHOD

Subjects

The first group of subjects consisted of 15 normal-hearing subjects with a mean age of 26 yr and with overall thresholds better than 10 dB HL at octave interval frequencies between 250 and 8000 Hz. Eight of the subjects reported previously participating in a listening experiment. Normal-hearing subjects were recruited from the student and staff populations at Northwestern University.

A second group of subjects consisted of 11 hearing-impaired subjects with a mean age of 48 yr. All hearing-impaired subjects displayed mild-to-moderate hearing impairments and would be considered appropriate canal or in-the-ear hearing aid candidates. All of the hearing-impaired subjects had at least 2 yr of experience wearing binaural hearing aids. Figure 1 reveals the mean thresholds and threshold ranges for the hearing-impaired group. Seven of the 11 subjects reported that they were satisfied with the sound quality of their present hearing aids. One of the hearing-impaired subjects reported previously participating in a listening experiment. The hearing-impaired subjects were recruited from the active patient population at the Northwestern University Hearing Clinic.

Materials

A speech passage ("Joe took father's shoe bench out. She was sitting at my lawn."), a piano passage (Oscar Peterson piano trio recording of Peterson's blues "The Smudge"), and an orchestra passage (Beethoven Violin Concerto in D, N. Y. Philharmonic) were used as the input signals. Each passage

lasted approximately 20 sec. An Ampex 456 master tape created by Killion (1979) contained the music passages as reproduced through AR-3a loudspeakers in a recording-studio-like room and recorded with microphones placed on the KEMAR manikin to simulate the position of an in-the-ear hearing aid. The speech consisted of live-voice recordings made under the same conditions. When reproduced into the input circuits of the laboratory hearing aids, these master tapes provided the equivalent of a live recording of the aids on KEMAR.

Hearing Aids

One binaural pair of hearing aids, the "starved Class A hearing aids," incorporated Knowles ED1913 receivers driven by Gennum LC505 amplifiers that had been wired for the low battery drain typical of canal aids. The second binaural pair, the "Class D hearing aids" incorporated Gennum LC504 preamplifiers and EP3074 receivers (each containing its built-in Class D amplifier). These hearing aids were built on a breadboard by Industrial Research Products, Inc., then a subsidiary of Knowles Electronics, for use in this experiment.

For the purpose of this experiment, the battery drains of both types were made as nearly equal as practical. Our best efforts at equalization in battery drain produced starved Class A hearing aids that had a battery life of approximately 112 hr and Class D hearing aids with a battery life of 185 hr using a "10A" size zinc air cell (Carlson, 1989). One-week battery life seems the minimum acceptable value in the marketplace, so the battery drain (and undistorted output) of the starved Class A amplifiers was about as high as is practical with a 10A battery.

A technical problem arises in comparing Class A circuits and Class D circuits because typical Class D circuits have a low output impedance. A low output impedance gives less treble emphasis in the hearing aid frequency response than is obtained with high output impedance, such as obtained with typical Class A amplifiers designed for hearing aids. The Class D hearing aids used in this study included an R-C equalization network to make their frequency response approximate that of the starved Class A hearing aids, as shown in Figure 2. To simulate common practice in the hearing aid industry, no damping was used in any of the hearing aids. Input/output curves for the binaural sets of hearing aids with a low (500 Hz) and high (5000 Hz) frequency signal are presented in Figure 3. The difference in saturation (i.e., headroom) is most evident in the higher frequencies.

An electrical-analog-circuit simulation of the Knowles EA-1842 microphone was also included in

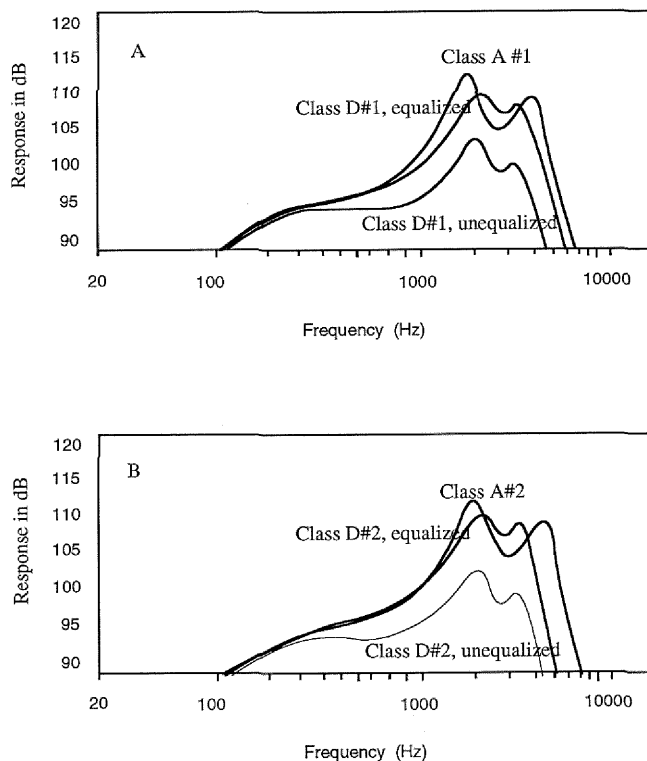


Figure 2. Frequency response curves for binaural sets (panels A and B) of starved Class A, Class D (equalized to match the starved Class A response) and Class D (unequalized) hearing aids with an a.c. input of 4 mVrms.

each circuit, so that the frequency response of those commonly used microphones (+5 dB response peak at 5 kHz and -7 dB low-frequency rolloff at 100 Hz) was included in our hearing aid response. (The master recordings had originally been made using BT-1759 microphones, which have a frequency response that is flat from below 100 Hz to beyond 10 kHz.) As in the case of the receiver response, the responses of interest represented the most common industry practice rather than high-fidelity design.

The high-frequency-average gain (HFA gain, the average of the gain at 1000, 1600, and 2500 Hz) of each aid was adjusted to 15 dB, simulating a typical use gain as reported by Berger (1988). Figure 2 shows the frequency responses of both sets of hearing aids with a simulated 80 dB SPL input. The "unequalized" Class D curve also is displayed to illustrate the frequency response difference in the absence of equalization, a difference that might have produced sound quality judgments based on preference for one or the other frequency response rather than differences in overload characteristics. (Indeed field reports of the superiority of the Class D receiver when it was simply substituted for the traditional receiver/Class A combination appeared partly due to the improved low-frequency response.) The item of interest in the present study was whether or

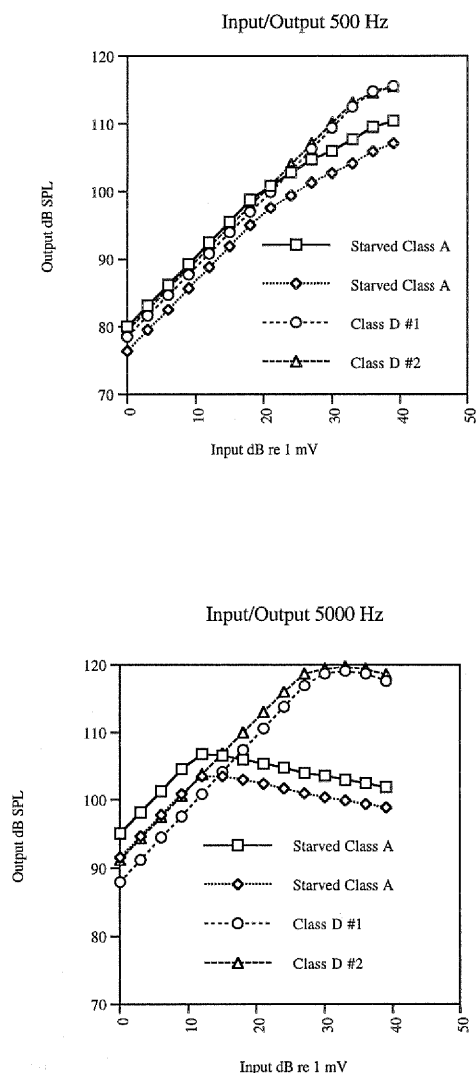


Figure 3. Input/output curves for binaural sets of starved Class A and Class D hearing aids with a 500 and 5000 Hz signal.

not the objectively measured improvement in undistorted output would be reflected in significantly improved subjective quality ratings.

Recording Procedures

Figure 4 shows a block diagram of the recording arrangement. The letters A–H are used in the following text to refer to the recording stages in that figure. All recordings were made through two channels to create binaural recordings and binaural playback. The electrical output of the Ampex playback unit was routed through a pair of Hewlett-Packard attenuators (A) to select the equivalent hearing aid input level in 5 dB steps from 70 to 100 dB SPL. The output of the attenuators was attenuated with a pair of resistors (B) chosen to provide the proper signal level and source impedance (3.5 kohm) to simulate

live hearing aid microphones. The signal was then fed to either the starved Class A or Class D hearing aids (C). The output of each hearing aid went through a Zwislocki coupler (D) to an ER-11 pre-amp (E) whose built-in diffuse-field-inverse filter was selected. (As described by Killion [1979], use of a diffuse-field-inverse filter permits the subsequent playback through high-fidelity earphones to best simulate live listening tests with actual hearing aids.) A 350 ohm series resistance (F) was used to bring the 250 ohm output impedance of the ER-11 up to the 600 ohm input impedance of the Hewlett-Packard attenuator (G) that was used to adjust the equivalent output sound level on half of the recordings. The final listening-test tape recordings were made on a SONY Digital Audio Tape (DAT) recorder (H). A VU meter with variable attenuator was used to set levels, with the rated input SPL corresponding to the frequent-peak readings on the VU meter.

Each of the seven simulated input SPL conditions (70 to 100 dB SPL in 5-dB steps) was repeated for each of the three source materials (speech, piano, and orchestra) and for two presentation-level conditions (output level determined by 15 dB hearing aid gain, and output level fixed at 70 dB SPL), resulting in binaural recordings of 42 individual listening-test passages for each hearing aid type (starved Class A and Class D). The two presentation levels were used to examine the possibility that subjects might either dislike or prefer a given presentation simply because it was louder or softer than the others, regardless of its distortion level. In the conditions where the output was determined by the gain of the hearing aids, the signal delivered to the subjects' ear canals had an across-the-board reduction of 5 dB below actual Zwislocki-coupler levels to promote listening comfort for all of the subjects. For the other half of the recordings, the recorded output level was held constant at 70 dB SPL to eliminate any bias in quality judgments due to loudness preferences.

Procedures

The previously recorded passages were reproduced on the DAT recorder, whose two output channels were fed to the two channels of a Crown Amplifier and then to a pair of Etymotic Research ER-1 insert earphones placed in the ears of the listening-test subject.

Subjects listened to blocks of 15 test passages: 14 conditions plus one anchor condition. The anchor condition, a heavily distorted passage, was presented first in each test block. The subject was told that the first passage in each block would be one of the poorest-quality sounds. Table 1 gives an example of one test block.

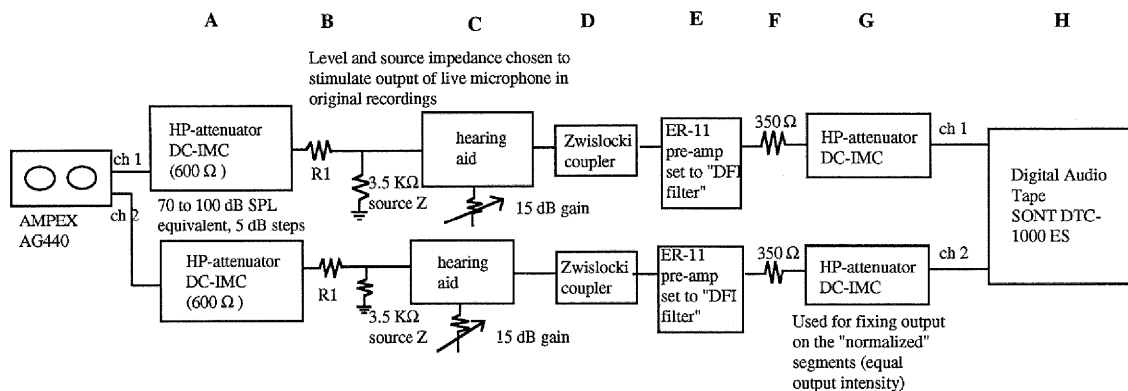


Figure 4. Schematic of the recording set up for the experiment. The letters indicate recording stages that are referred to in the text.

TABLE 1. One example test block that a subject would rate (15 ratings). There were six test blocks in one test session.

Material = Speech (Output not Fixed)		
Input Intensity	Class of Receiver	
1. Anchor (100)	A	
2. 90	D	
3. 85	A	
4. 75	D	
5. 70	D	
6. 100	D	
7. 70	A	
8. 85	D	
9. 75	A	
10. 80	A	
11. 80	D	
12. 100	A	
13. 95	D	
14. 95	A	
15. 90	A	

TABLE 2. Example order of test blocks for one test session. Half of the outputs were fixed at 70 dB SPL to eliminate any quality judgments based on loudness differences.

Material	Output Condition
1. Piano	Fixed at 70 dB SPL output
2. Orchestra	Fixed at 70 dB SPL output
3. Speech	Not fixed
4. Speech	Fixed at 70 dB SPL output
5. Piano	Not fixed
6. Orchestra	Not fixed

Six test blocks were presented in each testing session. Table 2 gives an example of the order of test blocks for one testing session. The order of the 14 conditions was randomized between test blocks, and the order of the test blocks was randomized between test sessions.

Two experiments were conducted. In Experiment I, the subjects rated sound quality on a scale of 0% to 100% (Gabrielsson, Shenkman, & Hagerman, 1988; Hagerman & Gabrielsson, 1985). In Experiment II the same subjects rated dollar value on a scale of \$0 to \$700.

Appendix A contains an example of the rating scales. Gabrielsson et al. (1988) originally published 10 scales covering different dimensions of sound quality. Since we were interested only in the listener's overall impression of sound quality, as defined by his/her personal view of what constitutes sound quality, we used only the general rating scale. In all probability, judgments in the present experiment

were based primarily on the subject's reaction to the amount of distortion in the program and to the overall frequency responses (which were not high fidelity, but typical of traditional hearing aid design), but in no case were subjects asked to define their rating methods. In the interest of time, we did not directly investigate the question of whether a given level of distortion would be rated differently depending on the frequency response, although the ratings for the flatter-response unequalized Class D condition provided indirect evidence of significant interaction (see Fig. 6).

The order of quality or value judgments on day 1 and day 2 was randomized among subjects: If a subject made quality judgments first on day 1, then the same subject automatically made value judgments first on day 2 and vice versa.

RESULTS AND DISCUSSION

Test versus Retest Data

Figure 5 shows the quality-judgment results obtained on day 1 compared with those obtained on the same passages on day 2. The 4.3% standard error of the mean calculated using a standard ANOVA procedure on the complete set of ratings is displayed as a bar ($\pm 4.3\%$) around each point. If there were no important learning or order effects, one would ex-

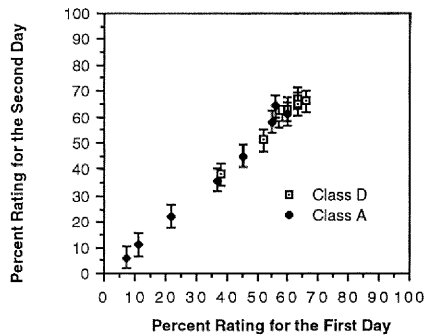


Figure 5. Quality ratings on first and second days, collapsed across material type and presentation method. Bars indicate standard error of the mean.

pect 68% (9.5 of the 14) of the first-day-second-day comparisons to fall within one standard deviation; pleasantly enough, 10 of 14 did so. A small tendency toward higher ratings on the second day was evident, but learning and order effects appeared relatively unimportant. In particular, the test-retest reliability of the data appeared adequate to permit combining the first and second day's results. Allowing for seven comparisons gives a critical difference of approximately 18% for statistical significance at the 95% confidence level for differences in the two-day-average across-subject, across-materials sound-quality ratings. (The Bonferoni method [Miller, 1966] was used to allow for multiple comparisons, so a significance level of $0.05/7 = 0.007$ was used, giving a t -test critical difference of $2.99 \cdot 1.41 \cdot 4.3\% = 18.1\%$.)

Experiment I: Sound Quality Judgments

Figure 6 illustrates the overall sound quality ratings for the combined listeners (5A), as well as the ratings for the normal hearing listeners (5B) and the hearing-impaired listeners (5C). The sound quality judgments represent an average of the test and retest ratings. The standard error of the mean derived from the ANOVA is represented on each graph by bars.

Figure 6A also includes ratings for the Class D (EP3074) receiver before equalizing its frequency response to that of the starved Class A (ED1913) receiver. The highest quality ratings at high input levels were awarded to the unequalized Class D circuit (labeled Class D, unequal), which had the best low-frequency response. This is consistent with industry reports on the enthusiastic reception for "Class D substitution," suggesting that the traditional amount of low-frequency rolloff employed in hearing aid design may be more than necessary.

The highest ratings did not exceed 70% under any conditions. This is consistent with the results ob-

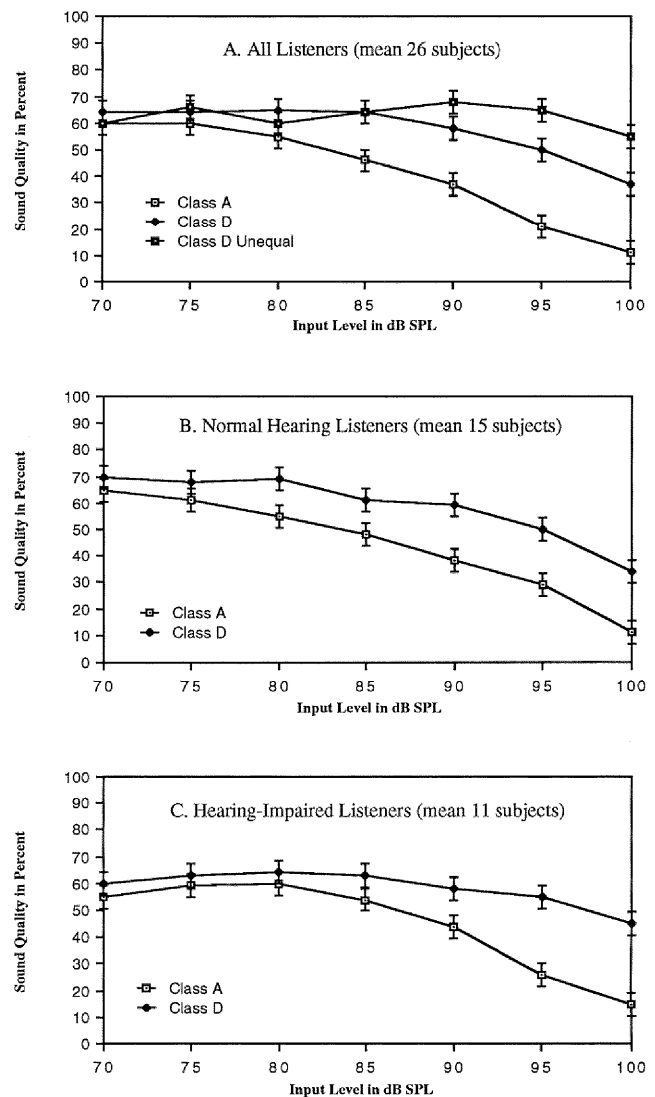


Figure 6. Overall sound quality ratings and sound quality ratings as a function of subject group. Ratings for the 14 conditions are collapsed across material type, presentation method, and test/retest administration. Ratings for the Class D unequalized frequency response appear in panel A only. Bars indicate the standard error of the mean.

tained by Killion (1989). Experimental high-fidelity-response hearing aids produced ratings of 80% or better, while a pair of undamped-response (peaky, as in Fig. 2) hearing aids produced ratings of 58%. In other words, the presence of response peaks reduced the maximum ratings to 72% of the high-fidelity hearing aids in Killion's study. It is tempting to speculate that the reason the maximum sound-quality ratings were as low as they were here was because the frequency responses of the laboratory hearing aids had a restricted bandwidth, reduced low-frequency response, and undamped response peaks, being intentionally designed to simulate then-common commercial practice rather than a

now-increasingly-popular high-fidelity frequency response.

Sound quality differences between the two hearing aid types were clearly perceived at the four higher-intensity input levels, where all of the Class D quality ratings were significantly higher than the starved Class A ratings (i.e., the grand-average differences exceeded the critical difference of 18%). The levels at which the differences became apparent, 80 dB SPL inputs for the normal-hearing listeners and 90 dB SPL for the hearing-impaired listeners, correspond to the approximate input level at which the starved Class A hearing aid started to overload.

At the two lowest input levels, 70 and 75 dB SPL, both hearing aid types should have been free from overload distortion. Any difference in ratings should thus reflect differences in frequency response of the two types of experimental hearing aids (which were matched as closely as possible using R-C equalization) or some subtle experimental error. The rated quality differences at those two levels was only a few percent, and not statistically significant, indicating an adequate match in frequency responses for the present experiment.

As seen in Figure 6B and 6C, the normal-hearing and hearing-impaired subjects gave generally similar ratings, although normal-hearing subjects appeared to give lower quality ratings to the more distorted passages than did the hearing-impaired subjects. This could be because the distortion was more audible to the normal-hearing subjects, or because the hearing-impaired subjects had become accustomed to high levels of distortion in their own hearing aids. In this regard, it is worth noting that other studies which have investigated the detectability of distortion have generally found little differences between normal-hearing and hearing-impaired subjects as long as the presentation levels were high enough (Gabrielsson, 1980). Interestingly, at 95 and 100 dB SPL inputs the hearing-impaired listeners rated the starved Class A hearing aids some 30% inferior in overall quality to the Class D aids.

Speech/fixed-gain ratings • One purpose of this investigation was to examine the quality ratings collapsed across a variety of materials (speech, orchestra, and piano) because most hearing aid wearers listen to all of these sounds through the same hearing aids. In other words, the global rating of quality across a realistic variety of listening materials was the target. Nonetheless, examination of the impact of material type on quality ratings in this experiment revealed that fixed-gain-presentation speech-material ratings were significantly lower than ratings obtained with some of the other signal

TABLE 3. The results of a repeated analysis of variance examining the impact of material type and input level. See Figures 7 and 8 for post-hoc analysis results.

Source	SS	df	MS	F	p
Material type					
Hypothesis	107,323	5	21,465	16	0.0000
Error	166,825	125	1335		
Material type and input level					
Hypothesis	78,997	65	1215	5	0.0000
Error	389,831	1625	240		

materials. This was true for both the starved Class A and Class D hearing aids. Table 3 provides the results of a repeated measures analysis of variance that indicated a significant difference between material types as well as significant interaction between material type and input level. Scheffe's post-hoc analysis (Marascuilo & Levin, 1983) was used to examine which material types were judged to have significantly different qualities while hearing aid type and input level were the same. This resulted in 210 comparisons, which are not listed in Table 3 in the interest of space conservation. The significant findings are supplied in Figures 7 and 8.

Figure 7 shows the across-subject-groups quality ratings of the Class D hearing aids across input levels. Material type (speech, orchestra, and piano) and presentation type (fixed-gain, i.e., determined from the input by the 15 dB gain of the hearing, or fixed-output, "normalized" to 70 dB SPL) are parameters in the various graphs. The fixed-gain-presentation speech material was the only material judged to be significantly different, and so those ratings are displayed in each graph panel against one of the other five combinations of material and presentation type.

Figure 8 shows the same data for the starved Class A hearing aids. Again, the speech material with fixed-gain presentation was the combination resulting in significantly different quality judgments from the other conditions.

Although there was an important interaction between material type and quality ratings, it was primarily due to the fixed-gain-presentation speech material, which received reduced ratings at the higher levels even with the Class D hearing aids. The likely explanation was that a substantial amount of tape hiss was audible between sentences for the speech passages at the higher input intensities. Tape hiss was not audible on the music passages in the present experiment because the latter had been recorded at a higher level and there were no pauses in the music. To explain: All materials came from the original Killion (1979) master-tape recordings, which used the same tape-recorder gain

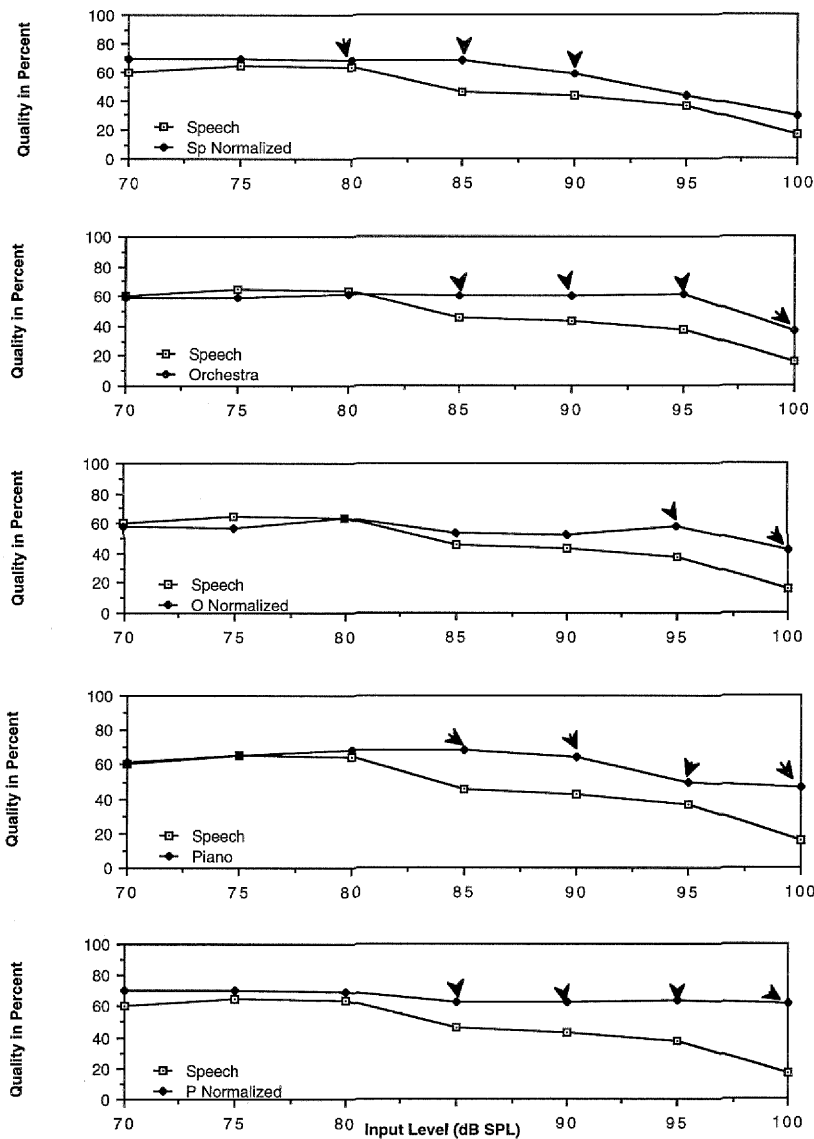


Figure 7. The difference between material types (s = speech, o = orchestra, p = piano) with output level fixed (normalized) and output level not fixed as a function of input level for the Class D receiver. Arrowheads indicate levels of materials that were significantly different ($p < 0.01$) from the speech material at the same input level.

settings for all materials (to permit error-free reproduction at original, live performance, levels). In these recordings, 0 VU corresponded to the 100 dB SPL orchestral peaks. The 65 to 70 dB SPL speech levels were thus recorded at -30 dB VU, much closer to the tape noise floor. Tape noise had been no problem in the original Killion listening tests, where each passage was reproduced at original levels (tape noise was approximately 30 dB(A) SPL, 10 dB below typical room noise, during reproduction), nor was it a problem in the present study under fixed-output (70 dB SPL) presentation conditions. It became a problem with fixed-gain presentation in these experiments, however, because the same speech passages were reproduced with up to 45 dB of gain: 15 dB of hearing aid gain plus the 10 to 30 dB of reproducer gain required to reach the equivalent of 80 to 100 dB SPL at the hearing aid microphone input. This provided an inadvertent test of the addition of poor

signal-to-noise ratio speech material at high presentation levels. While overload was still objectionable, as indicated by the fact that the Class D hearing aids received higher ratings at all input levels, a noisy signal presented at high levels was given a low quality rating even when it was not distorted.

With the exception of the fixed-gain-presentation speech material, the general shapes of the quality rating curves, and the differences between the starved Class A and Class D hearing aids, are fairly consistent across materials, subjects, and output conditions. Even with the fixed-gain-presentation speech material, the general trend is the same as for the other ratings.

The consistency in ratings seen especially in Figure 8 between fixed-output and fixed-gain presentations indicates that loudness (either loudness discomfort or a desire for increased loudness) was not a significant factor in rating quality in this experiment.

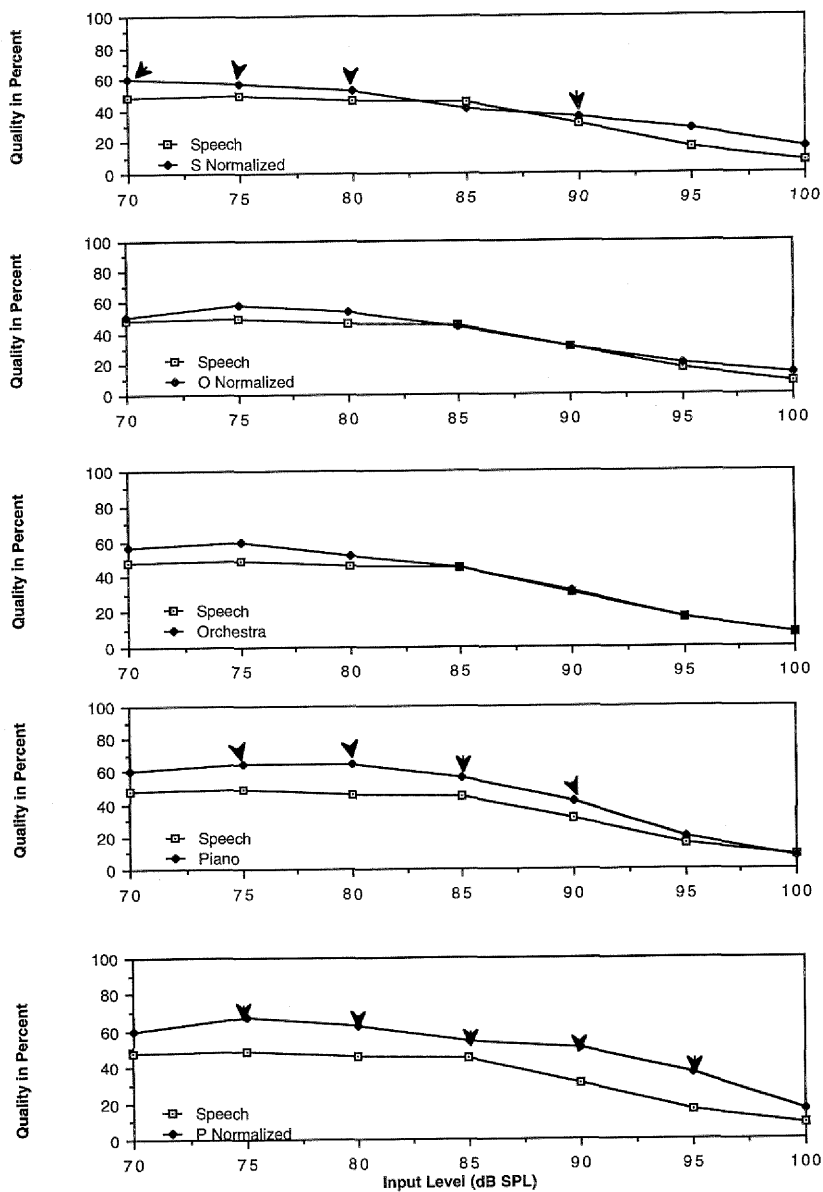


Figure 8. The difference between material types (s = speech, o = orchestra, p = piano) with output level fixed (normalized) and output level not fixed as a function of input level for the starved Class A receiver. Arrowheads indicate levels of materials that were significantly different ($p < 0.01$) from the speech material at the same input level.

Experiment II: Dollar Value Judgments

Experiment I was replicated as Experiment II with all conditions the same except the subjects were asked in Experiment II to make value ratings instead of quality ratings. The subjects indicated how much they would be willing to pay for a hearing aid that produced the sound quality in each passage. The maximum a subject was allowed to pay was \$700. A rating of \$0.00 implied that the subject would rather not be able to hear quiet sounds than wear this hearing aid. The hearing-impaired subjects most likely had a stronger sense of what it is like to miss quiet sounds than the normal-hearing subjects with respect to the \$0.00 rating. Although the value results must be interpreted with caution considering differing hearing loss experience, income, etc., the ratings are of some interest because,

to the authors' knowledge, they represent the first systematic attempt to assign a dollar value to the sound quality of hearing aids.

Figure 9 illustrates the grand-average value ratings, plotted on a scale from \$0.00 to \$500. No one rated the value higher than \$500. The likely explanation (limited bandwidth, response peaks, reduced low-frequency response) was discussed above. It is tempting to speculate that the subjects might have value-rated each condition substantially higher if the upper limit had been set at \$2000 per aid instead of \$700.

Interestingly enough, multiplying each of the original sound quality ratings from Figure 6A by a factor of \$6.75 produced the second, overlapping, set of curves in Figure 9. The close correlation indicates that dollar value and overall quality are highly correlated, which in turn permits the observation

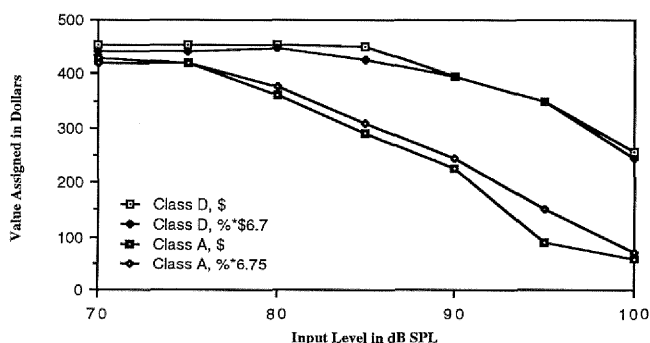


Figure 9. Value (\$) versus quality rating comparisons using the average data of all subjects, all test materials, fixed and not fixed outputs, and test and retest administration. The original sound quality ratings have been multiplied by a factor of \$6.75.

that these individuals indirectly rated each sound-quality percentage point as worth about \$6.75. More important, both the quality and the dollar-value ratings had maximums of approximately 70% of full scale, suggesting that had a much higher dollar limit been used in the valuation ratings a much higher apparent valuation than \$6.75 per percentage point sound quality would have resulted.

One of the important original questions was; "Even if Class D sounds better, will anyone be willing to pay for it?" The judged value difference between the two types of aids is as great as \$200 in the higher intensities, although one might expect the hearing aid user to turn down the volume control in these situations where prolonged high-level inputs prevail. Although the relationship between laboratory pencil judgments and willingness to pay out real dollars is unknown, the differences shown in Figure 9 indicated that the answer is "yes." Since the time of this study, when Class

D hearing aids comprised less than 1% of sales, there has been substantial market confirmation of this answer; Class D hearing aids now approach 25% of U.S. sales.

Subjective Results versus Physical Measurements

Coherence versus frequency measurements were made on one sample of each type of hearing aid at a variety of input levels. Discussed briefly above, the coherence function has a value that ranges from 0 to 1.0, with a value of 1.0 indicating perfect reproduction at that frequency in the sense that all of the output at that frequency is due to the input signal. A value of 0.9 results when only 90% of the output power at a given frequency is directly correlated with the input at that frequency, with the other 10% of the power (31.6% of the signal) due to noise and/or distortion products from inputs at other frequencies. (If the output consists of noise-free distortion, a coherence-function value of 0.99 corresponds to a more traditional THD distortion of 10%, 0.9 to 31.6%, and 0.75 to 50% distortion.) A coherence value of 0 results when the only measurable output at a given frequency consists of noise and distortion products unrelated to the input signal at that frequency. Figures 10 and 11 present overall coherence functions for the starved Class A and Class D hearing aids, respectively. The graphs show the coherence function versus frequency for each type of hearing aid at the various input levels used in the present investigation. The coherence factor decreases with increasing level in an orderly fashion for both sets of hearing aids. The two lowest input intensities result in similar coherence functions for

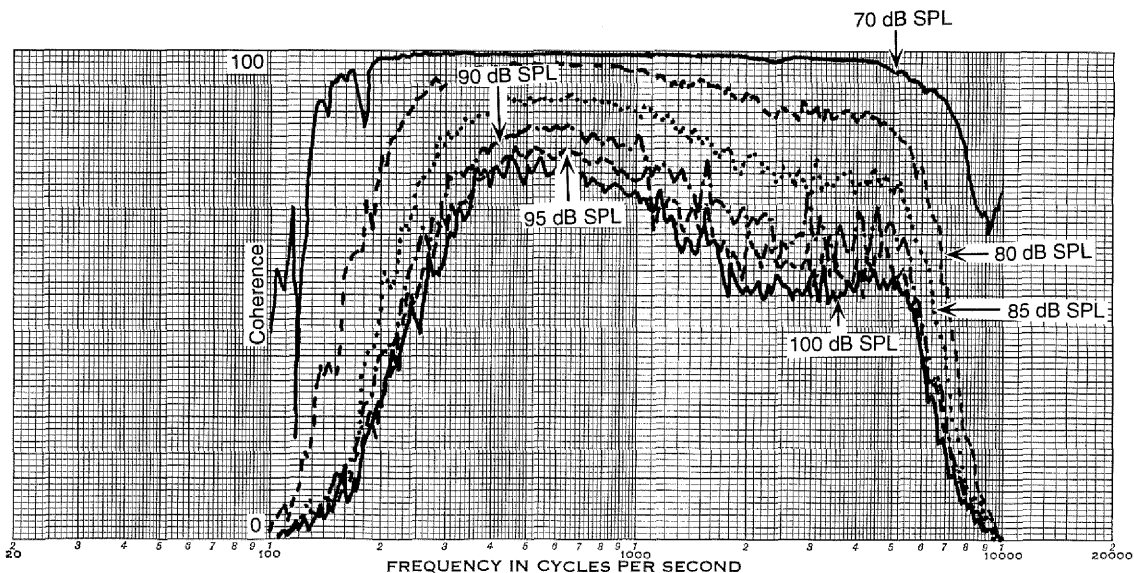


Figure 10. Coherence function response (0% to 100%) with input level as parameter for the Class A hearing aid used in this study.

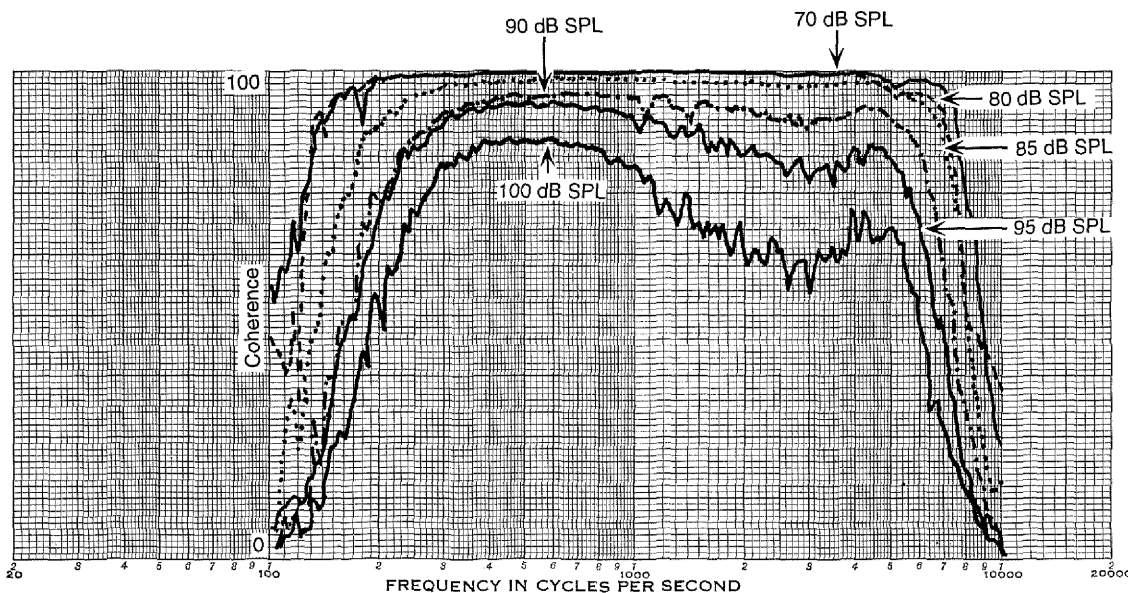


Figure 11. Coherence function response (0% to 100%) with input level as parameter for the frequency response equalized Class D hearing aid used in this study.

both sets of hearing aids, which was the expected result since distortion-free operation of the starved Class A amplifier is expected at those levels. The higher intensities reveal the largest differences between the two receiver-amplifier combinations, with the Class D hearing aid exhibiting the higher coherence function at each level.

As expected, the hearing aids with starved Class A amplifiers saturate at lower intensity levels than the hearing aids with the Class D receivers. Because a total battery drain consistent with use in a canal aid was employed, the capabilities of the starved Class A amplifier were unavoidably and automatically limited. Indeed, informal listening tests confirmed that a Class A stage biased at 3 mA of battery drain would provide performance entirely comparable to that of the Class D receiver; no obvious audible difference at any input level. Thus, it should be emphasized that it was not the Class D operation per se that caused the improved overload characteristic, but rather the fact that the Class D operation permitted a greatly improved overload characteristic while maintaining a reasonable battery drain.

Figure 12 shows a comparison between sound-quality ratings and transformed 3 kHz coherence-function values. The transformation used on the coherence-function values was a multiplication of the 3 kHz value by 0.65, the highest average quality rating obtained in this study. The practical justification for this transformation is that it seems to provide a good correlation between subjective and objective measurements over a fair range of values. A second, theoretical, justification for the transformation used on the coherence-function values is that limitations in bandwidth and response

smoothness of the hearing aids used in this study appear to have limited the subjective quality ratings to a ceiling of 65%. Since the coherence-function test measures only aberrations due to noise and distortion, and not frequency-response, some type of subjectively derived adjustment is needed to account for the latter aberrations. The fact that a single-value multiplier appears adequate, and that its value is equal to the maximum value for the average quality ratings (divided by 100 to reduce the percentage rating to a decimal fraction), is of interest. Although a good correlation has been found between the subjective measure and the transformed 3 kHz coherence-function values, one cannot interpret a cause and effect relationship from these data. If the same result holds in future studies that include subjective sound quality ratings (e.g.,

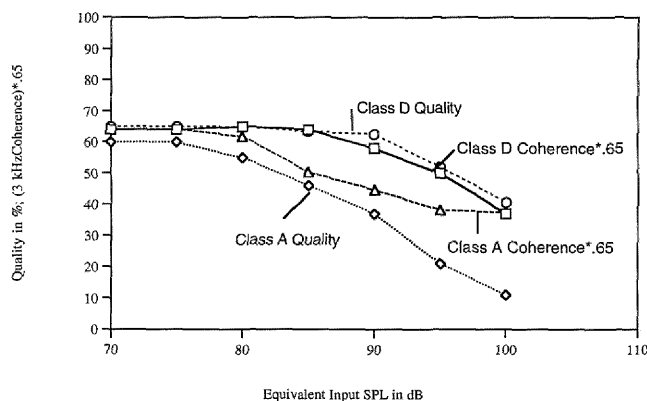


Figure 12. Comparison between across-conditions-average subjective quality ratings and measured coherence function values obtained at each input SPL with the Class A and Class D hearing aids used in this study.

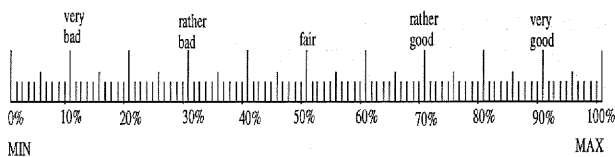
Schweitzer, Grim, Preves, Kubichek, & Woodruff, 1991), and a wider range of noise and distortion characteristics, the coherence measurement may prove useful in helping predict subjective responses to noise and distortion.

SUMMARY

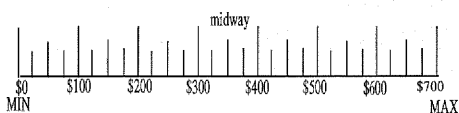
Normal-hearing and hearing-impaired subjects rated experimental hearing aids containing a Class D receiver as having superior sound quality compared with hearing aids containing starved Class A output amplifiers. This finding held across a variety of input intensities and test materials, and was consistent with the physically measured improvement in overload characteristics of the Class D receiver. Superior sound quality translated into an increase in perceived value for the hearing aid. Based on this experiment, each percentage point in sound quality appears to be worth \$6.75.

Appendix A

Rating scale for sound quality judgments.



Rating scale for value (\$) judgments.



ACKNOWLEDGMENT:

The authors would like to thank Knowles Electronics, Inc. for their support of this project.

Address for correspondence: Catherine V. Palmer, Ph.D., University of Pittsburgh, Dept. of Communication Science and Disorders, 3347 Forbes Avenue, Pittsburgh, PA 15260.

Catherine V. Palmer is now at the University of Pittsburgh, Pittsburgh, PA.

REFERENCES

- Agnew, J. (1988). Hearing instrument distortion: What does it mean for the listener? *Hearing Instruments*, 39(10), 10-20.
- Berger, K. (1988). Predicted versus actual gain usage by binaural hearing aid wearers. *The Hearing Journal*, 20-22.
- Burnett, E., Corliss, E., & Nedzelnitsky, V. (1982). Research problems in coupler and in situ measurements on hearing aids. In G. Studebaker, F. Bess (Eds.), *Vanderbilt Report* (p. 71). Upper Darby, PA: Monographs in Contemporary Audiology.
- Carlson, E. (1989). *The Knowles amplified receiver, report 10676-1*. Elk Grove Village, IL: Industrial Research Products, Inc.
- Gabrielsson, A., Hagerman, B., Berg, C., Ovegard, A., & Anggard, L. (1980). Clinical assessment of perceived sound quality in hearing aids. Report TA 98, Department of Tech. Audiology, Karolinska Institute, Stockholm.
- Hagerman, B., & Gabrielsson, A. (1984). *Questionnaires on desirable properties of hearing aids. Report TA 109*. Stockholm, Sweden: Karolinska Institute.
- Hagerman, B., & Gabrielsson, A. (1985). Questionnaires on desirable properties of hearing aids. *Scandinavian Audiology*, 14, 109-111.
- Killion, M. (1979). Design and evaluation of high-fidelity hearing aids (Unpublished doctoral dissertation, Northwestern University, 1979). *Dissertation Abstracts International* (University Microfilms No. 7917816).
- Killion, M. (1980). Problems in the application of broadband hearing aid earphones. In G.A. Studebaker, & I. Hochberg (Eds.), *Acoustic factors affecting hearing aid performance*. Baltimore: University Park Press.
- Marascuilo, L., & Levin, J. (1983). *Multivariate statistics in the social sciences*. Pacific Grove, CA: Brooks/Cole Publishing Co.
- Miller, R. G. (1966). *Simultaneous statistical inference*. New York: McGraw Hill.
- Preves, D., & Newton, J. (1989). The headroom problem and hearing aid performance. *The Hearing Journal*, 42, 19-26.
- Punch, J., & Beck, E. (1980). Low-frequency response of hearing aids and judgments of aided speech quality. *Journal of Speech and Hearing Disorders*, 45, 325-335.
- Punch, J., & Beck, L. (1986). Relative effects of low-frequency amplification on syllable recognition and speech quality. *Ear and Hearing*, 7, 57-62.
- Schweitzer, C., Grim, M., Preves, D., Kubichek, R., & Woodruff, B. (1991). *Qualitative assessment of hearing aid performance by an expert pattern recognition system*. Paper presented at the American Academy of Audiology Convention, Denver, CO.
- Skinner, M., & Miller, J. (1983). Amplification bandwidth and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. *Audiology*, 22, 253-279.
- Studebaker, G., & Sherbecoe, R. (1988). Magnitude estimations of the intelligibility and quantity of speech in noise. *Ear and Hearing*, 9, 259-267.