BEHAVIORAL CORRELATES OF HEARING-AID PERFORMANCE[a]

James Jerger, Ph. D.

Research Institute of the
Houston Speech and Hearing Center
1343 Moursund Avenue
Houston, Texas 77025

Commercially available hearing aids differ substantially in the fidelity with which they amplify the sounds of speech. These differences can be described by a variety of physical measurements. Examples are frequency response, harmonic distortion, transient distortion, and intermodulation distortion. On the basis of such measurements it is possible to identify good, average, and poor fidelity aids.

The question remaining, however, is to what extent these differences in the physical performance of hearing aids are important in terms of the user's ability to understand speech through the aid.

Our hearing-aid-research program is directed toward this single goal. We are attempting to determine how much difference in the physical characteristics of aids makes a difference to users of aids.

We are attacking the problem through an orderly series of experiments designed to answer the following five questions.

First, is it possible to find a behavioral technique that will differentiate among hearing aids?

Second, are differences among aids, as measured by such a technique, larger, as large, or smaller for hard-of-hearing listeners than for normal listeners?

Third, is there any interaction between hearing aids and type of hearing loss? In other words, are certain hearing aids best for patients with one type of loss, and other aids better for other types of loss, or is the best aid for anyone the best aid for everyone?

Fourth, what is the optimal behavioral technique for differentiating among hearing aids? What combination of signal input and experimental task best rank orders aids in a valid and reliable manner with minimal inter-subject variance, in terms of the hearing aid user's ability to understand speech?

Fifth, how much of a physical difference makes a behavioral differ-

[a] Based on work performed under VA Contract V1005M–1239.

ence? In other words, at what point on the continuum of any physical characteristic of hearing aids does the behavioral index reflect a change from relatively good to relatively poor "usability"?

Can Aids Be Differentiated Behaviorally?

Our first problem was to determine whether or not it was at all possible to show differences among hearing aids by behavioral techniques. This was a necessary first step. It was essential to identify techniques that would successfully order aids along a behavioral continuum. Unless such techniques could be found, subsequent phases of the research program would be quite meaningless.

Previous research findings with monosyllabic (PB) word lists were not encouraging. It appeared that performance on such word lists was relatively independent of hearing-aid quality.

Our basic approach to this problem was to abandon single-word test materials and to turn, instead, to sentences as units of measurement. We chose, for this task, a multiple-choice sentence intelligibility test originally developed during World War II at the Harvard Psycho-acoustic Laboratory and known as PAL–8. From an original pool of 100 sentences three equivalent forms of 30 sentences each were constructed. Each test item consists of a single sentence, either a question or command, and four alternative answers. Table 1 shows two illustrative test items.

TABLE 1.—*Sample Questions from PAL–8 Test*

The cold weather ends in: Furnaces Freezing Tickling *Spring*
When it gets very cold, water becomes: North-Pole Winter *Ice* Mines

Our next step was to choose three hearing aids that best represented the physical continuum of hearing-aid performance: a very good aid, an average aid, and a very poor aid. Our limited objective at this point was simply to find a behavioral test that was capable of differentiating these extremes of physical performance.

Accordingly, the sentence test materials were played through each aid and recorded on magnetic tape along with a competing speech message at a primary-to-secondary ratio of —6 dB. Six normal listeners were then tested over six trials on each aid.

Figure 1 shows that we were successful in reaching our first objective. Average performance curves for the three aids show substantial differences. Aid "A," characterized by flat frequency response and minimal harmonic distortion, gives the best result. Performance is much poorer for aid "C," which was characterized by considerable harmonic distortion, and intermediate for aid "B," characterized by a peaked frequency response and moderate distortion.

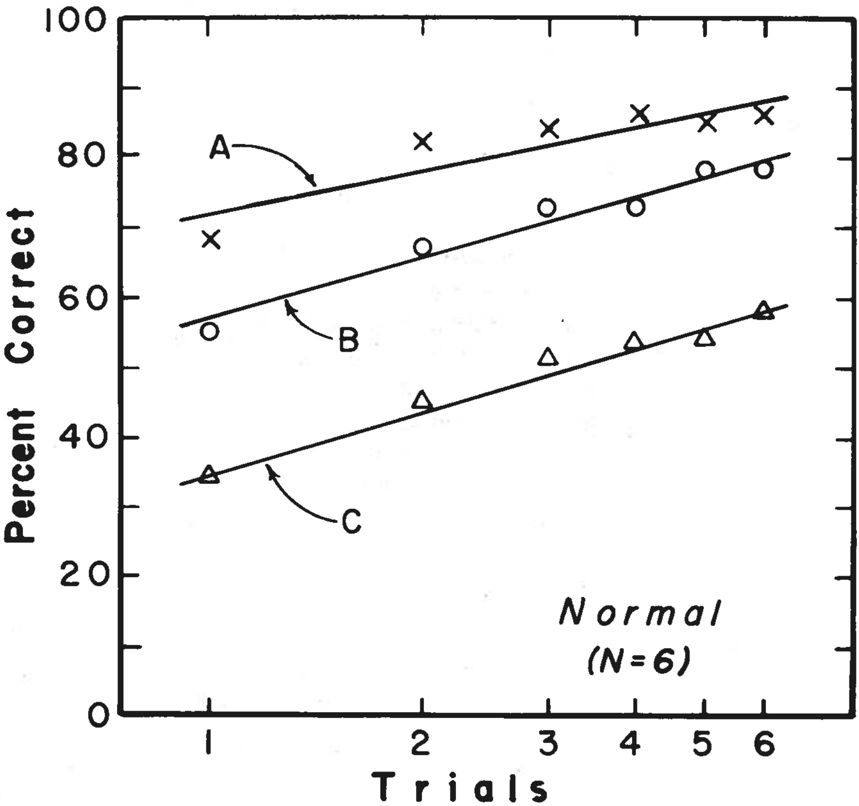


FIGURE 1.—Average performance curves for normal-hearing individuals listening to sentences recorded through three different hearing aids. The differences among the three curves show that the best hearing aid (A) always yielded the best results and the worst hearing aid (C) the poorest results, with the intermediate aid (B) ranked in between.

This was a most encouraging result. It suggested that appropriate sentence test materials would reflect differences in the physical characteristics of various aids, at least in normal ears.

Are Aid Differences Less Important to Normals?

Our second problem was to determine whether such behavioral differences were smaller, as large, or larger in patients with hearing loss.

Accordingly, we carried out the same procedure on six patients with moderate degrees of sensorineural hearing loss. In Figure 2 the performance of these patients can be compared with the previous results obtained on normals (Fig. 1). We see that there is remarkably little difference between the two groups in terms of the variation

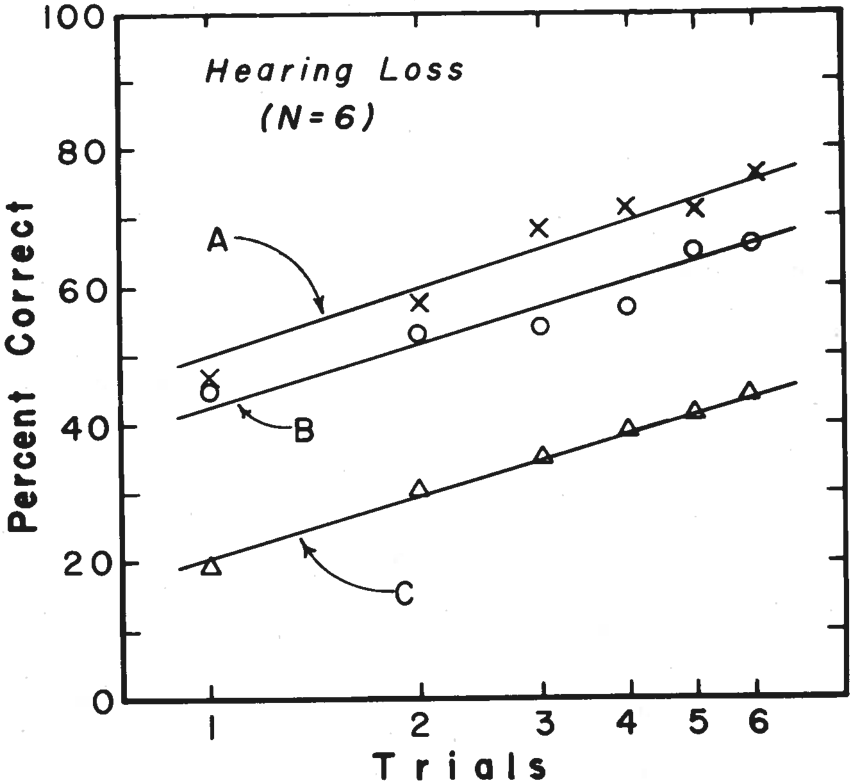


FIGURE 2.—Average performance curves for hearing-impaired individuals listening to sentences recorded through three different hearing aids. The scores for all aids are in the same rank order as the normal listeners' scores. The best hearing aid (A) always gave the best results and the poorest aid (C) the worst scores.

across aids. If anything, the differences are slightly larger in the normal group.

These, too, are significant results. They imply that differences among hearing aids are at least as important to normal listeners as to hard-of-hearing listeners. Thus one may confidently evaluate aids on normal listeners alone without fear that differences important to the hearing-impaired will fail to appear.

Do Aids and People Interact?

Our third problem was to determine whether differences among aids were the same for all hearing-impaired patients regardless of type or extent of hearing loss. It has long been presumed that, because of the uniquely different natures of their losses, some hard-of-hearing patients may do best with one hearing aid, while others will do better with a different aid. The alternative to this argument is that the best hearing aid for any patient is the best hearing aid for all patients; in other words that no interaction exists between aids and people.

To answer this question we recorded a wide variety of speech materials through the three experimental aids, then played the tapes to 36 hearing-impaired listeners who represented every conceivable type and degree of hearing loss for which a wearable hearing aid would be considered appropriate. Results showed that, on the whole, subjects ranked the three aids in the same order on the sentence intelligibility test. In other words, aid A was best, aid B second best, and aid C worst, irrespective of type or degree of hearing loss.

Interestingly enough, results obtained with conventional monosyllabic word lists through the same three aids (Fig. 3) were quite ambiguous. There was no clearcut pattern for either aids or aid-by-listener interaction. Some word lists suggested superiority for one aid, while other lists ranked the same aid as poorest.

The results tend to confirm what many investigators have long surmised: that the error of measurement associated with single lists of monosyllabic words is much larger than the inherent performance differerences they are capable of reflecting. Such material, therefore, cannot be used for meaningful differentiation among hearing aids.

Further analysis of these data casts considerable doubt on still another time-honored maxim of hearing-aid fitting. It has long been held that differences in the physical characteristics of aids are less important for patients with mild conductive loss and good discrimination scores than for severe sensorineural losses with relatively poorer discrimination scores.

Figures 4, 5, and 6 suggest that, if anything, the opposite is true. When subjects are regrouped according to particular dimensions of hearing loss it becomes clear that, in terms of a behavioral correlate

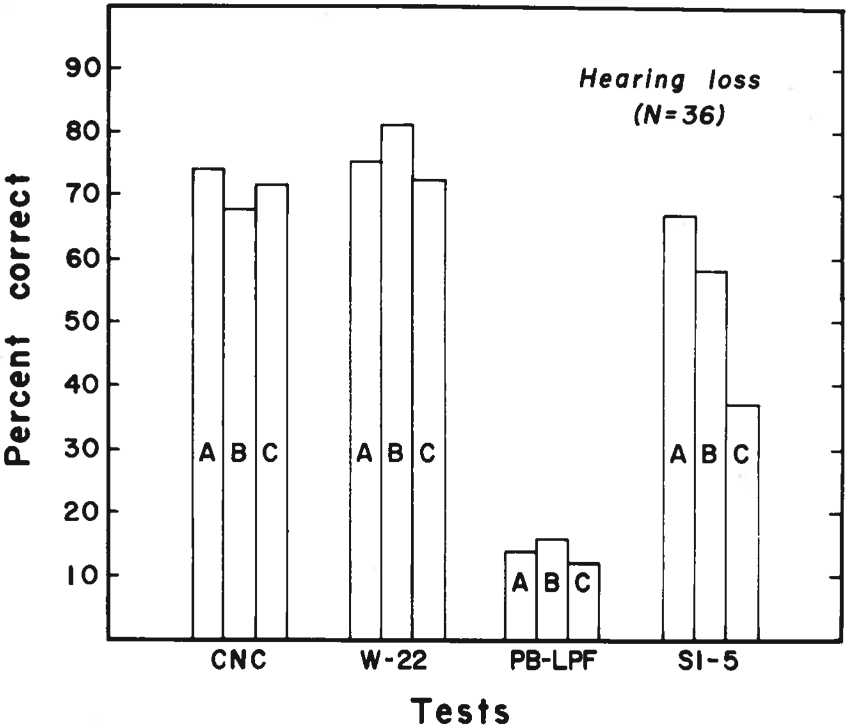


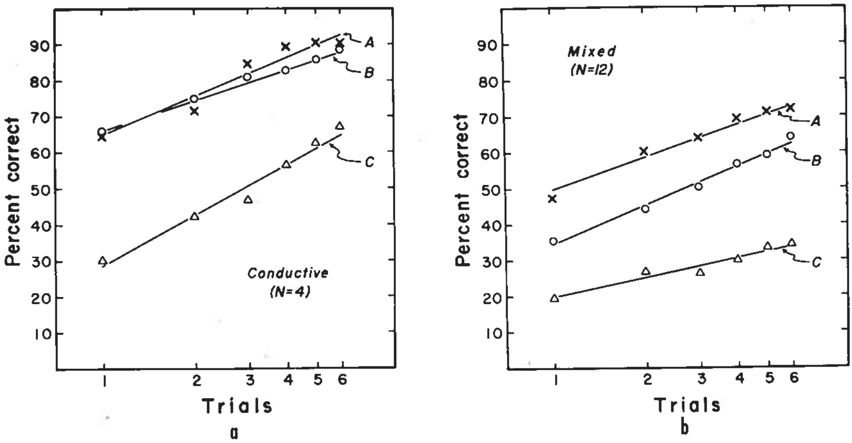
FIGURE 3.—Average scores for hearing-impaired individuals listening to one-syllable words and to sentences recorded through three hearing aids. Only the sentence materials showed reliable differences among the aids and rank ordered the best aid first and the worst aid last.

of hearing-aid performance, differences among aids are larger for young, mild, flat, conductive losses with good PB discrimination than for old, severe, sloping sensorineural losses with poor PB discrimination scores.

The Optimal Behavioral Technique

The fourth problem was to determine the optimal behavioral technique for differentiating among hearing aids. Ideally, test materials should consist of some form of on-going speech since this bears the highest face validity in terms of what we seek to predict about hearing-aid performance.

Speech materials have several inherent disadvantages, however. First, our previous research has shown that only with great difficulty can a suitable task involving speech materials be constructed. In order to render the sentence intelligibility materials satisfactory from the standpoint of differentiating among aids, it was necessary to add a competing speech signal at a relatively high level. Second, both the test



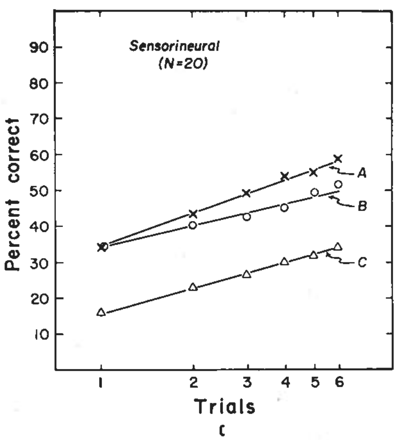


FIGURE 4.—Average performance curves for sentences recorded through three hearing aids for patients with (a) conductive hearing loss, (b) mixed hearing loss, and (c) sensorineural hearing loss. The differences between the three curves are largest for the conductive hearing loss patients and smallest for the patients with sensorineural hearing losses. In other words, the physical characteristics of the aids were more important for the patients with conductive hearing loss than for other patients.

signal and the competing signal are difficult to specify because of the complex time course of real speech. Third, equivalent forms are difficult and tedious to construct. Also, for any given type of material, words or sentences, the number of possible equivalent forms is finite and severely limited.

We have sought, therefore, a more satisfactory signal than actual speech; one that would adequately reflect the varying physical characteristics of aids, would correlate highly with performance on speech materials, and would be exactly specifiable as an input signal to the hearing aid.

We took, as a point of departure, the phenomenon of intermodulation distortion. When any complex signal is transduced by a hearing aid, energy in two or more frequency regions may interact or "inter-

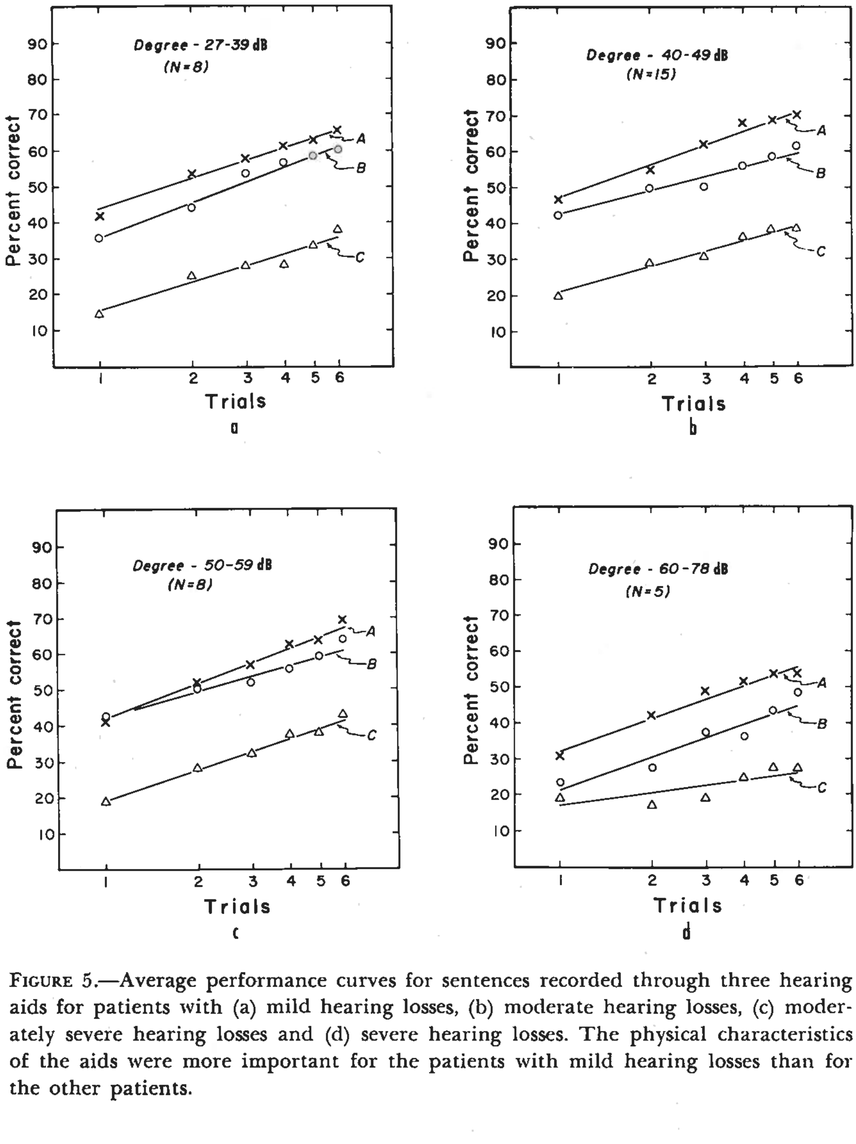
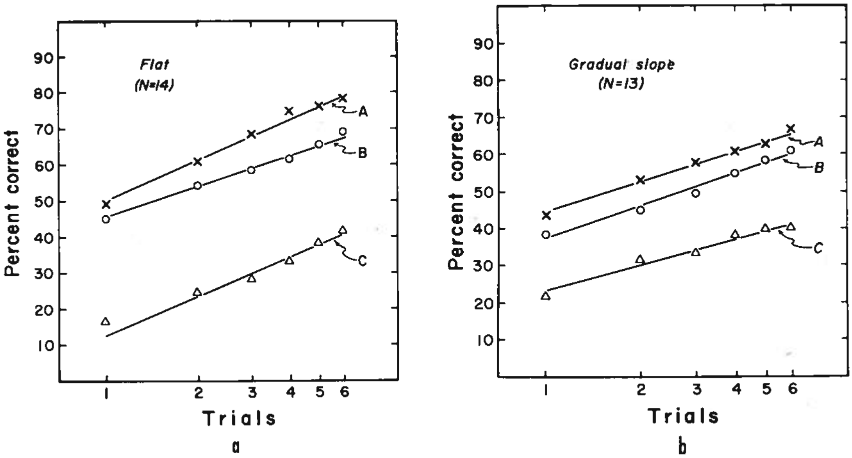


FIGURE 5.—Average performance curves for sentences recorded through three hearing aids for patients with (a) mild hearing losses, (b) moderate hearing losses, (c) moderately severe hearing losses and (d) severe hearing losses. The physical characteristics of the aids were more important for the patients with mild hearing losses than for the other patients.

modulate" to form distortion products at sum and difference frequencies or multiples thereof. Conventional wearable hearing aids are frequently characterized by excessive amounts of such distortion products, and it is reasonable to assume that unwanted energy in appropriate spectral locations could be a serious detriment to speech intelligibility.

We have, therefore, designed a simple test procedure based on intermodulation distortion for the purpose of differentiating among hearing



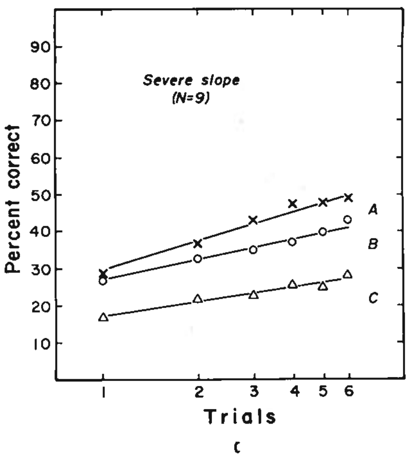


FIGURE 6.—Average performance curves for sentences recorded through three hearing aids for patients with (a) flat hearing curves, (b) gradually sloping hearing curves, and (c) severely sloping hearing curves. The physical characteristics of the aids were more important for the patients with flat hearing curves than for the other patients.

aids. After considerable preliminary exploration we chose, as input signals to the aids, two sinusoids, of equal amplitude: one at 1000 Hz (f_1); the other at 1600 Hz (f_2). When this dual-frequency signal is transduced by an aid, distortion products of considerable magnitude are typically created at 400 Hz ($2f_1-f_2$), 600 Hz (f_2-f_1), 2200 Hz ($2f_2-f_1$), and 2600 Hz (f_1+f_2).

Two-channel magnetic tape recordings were prepared by recording on one channel the "clean" two-frequency signal and on the other channel the same two-frequency signal after it had been transduced by the hearing aid (Fig. 7).

The behavioral test procedure is then carried out by presenting brief bursts from each channel to a listener in a two-alternative, forced-choice paradigm. The listener hears a pair of signals, each of 300 msec. dura-

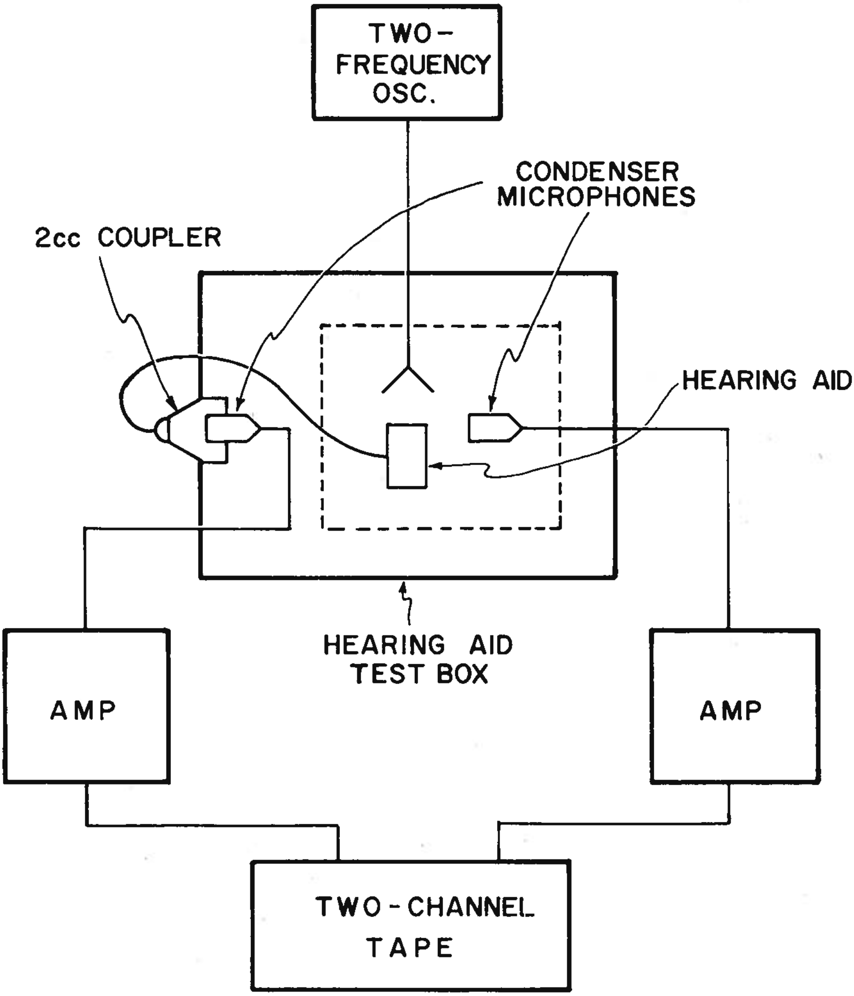


FIGURE 7.—Method for recording Intermodulation Distortion Test through each hearing aid.

tion, separated by a silent interval of 200 msec. His task is to press an appropriate response key depending on whether he heard the "clean" signal first or the "distorted" signal (hearing aid) first. The order in which the two signals actually occur is randomly varied over successive trials with an a priori probability of 0.50. The test signals are mixed with various amounts of random noise, and performance is determined as a function of signal-to-noise ratio.

Figure 8 shows results for the three aids used in the previous work reported here. The best aid (A) introduces so little intermodulation

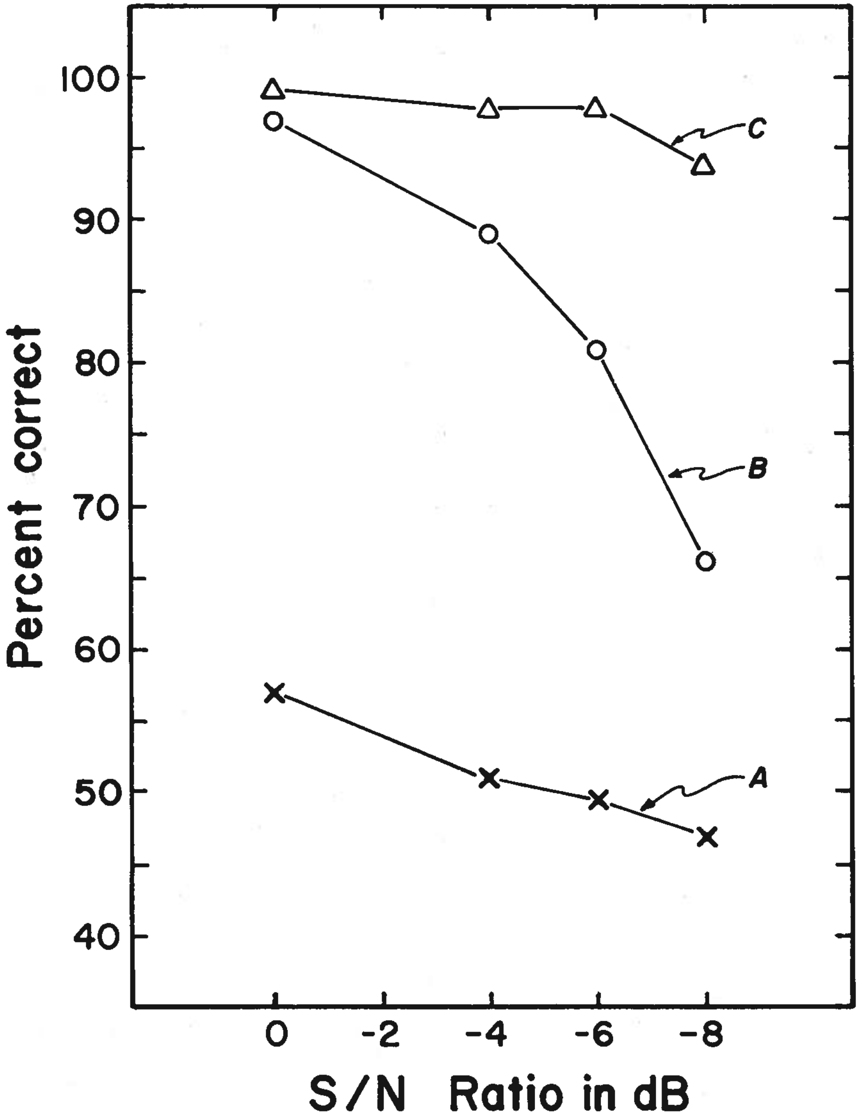


FIGURE 8.—Results of Intermodulation Distortion Test on the three hearing aids. The best aid (A) has so little distortion that it cannot be easily distinguished from the undistorted signal. The worst aid (C), on the other hand, has so much distortion that it is easily differentiated from the undistorted signal. Again, aid (B) is intermediate, being easy to distinguish at favorable signal-to-noise ratios and progressively more difficult to distinguish as increasing signal-to-noise ratio makes the task more difficult.

distortion that listeners cannot easily differentiate it from the "clean" channel until the signal-to-noise ratio becomes extremely favorable. On the other hand, the poorest aid (C) introduces so much distortion that listeners can easily differentiate it from the "clean" channel, even at very unfavorable signal-to-noise ratios.

Figure 9 compares inter-subject variablity on this new Intermodulation Distortion Test (IDT) with variability of the same six subjects on the previously described sentence intelligibility test for each of the three aids. We note a considerable advantage for the IDT procedure in terms of minimal inter-subject variance.

These various results are encouraging. They lead us to believe that the IDT procedure may have substantial value as a behavioral technique for differentiating among aids with widely varying physical characteristics. The procedure has two striking advantages over speech materials. First, intersubject variability is minimal; second, the test signal is exactly specifiable.

What is a Difference That Makes a Difference?

Currently, we are attempting to answer the fifth question: how much of a physical difference in aids makes a behavioral difference to listeners? To attack this problem we chose 21 representative aids, differing appreciably in physical characteristics, from among the various aids submitted to the VA in 1966. Using the data submitted by the National Bureau of Standards (NBS) on each aid, we selected 21 aids that seemed to represent the continuum of physical performance from very best to very worst.

We have made tape recordings through all aids using both the IDT approach and the synthetic sentence identification approach recently developed in our laboratory. Our plan is to play both sets of tapes to a number of normal and hearing-impaired listeners. We will then be in a position to study the inter-correlations of both behavioral and physical indices of hearing-aid performance, and to define the physical limits of behavioral tolerability if such limits do, in fact, exist.

SUMMARY

In pursuit of our primary research goal, to determine the tolerance limits for physical characteristics of hearing aids from the standpoint of behavioral performance, we have learned the following:

1. It is possible to differentiate among hearing aids by means of suitably constructed behavioral techniques.
2. Such behavioral indices suggest that physical differences among aids are at least as important to normal listeners as to hearing-impaired listeners.
3. Such behavioral indices fail to indicate a significant interaction

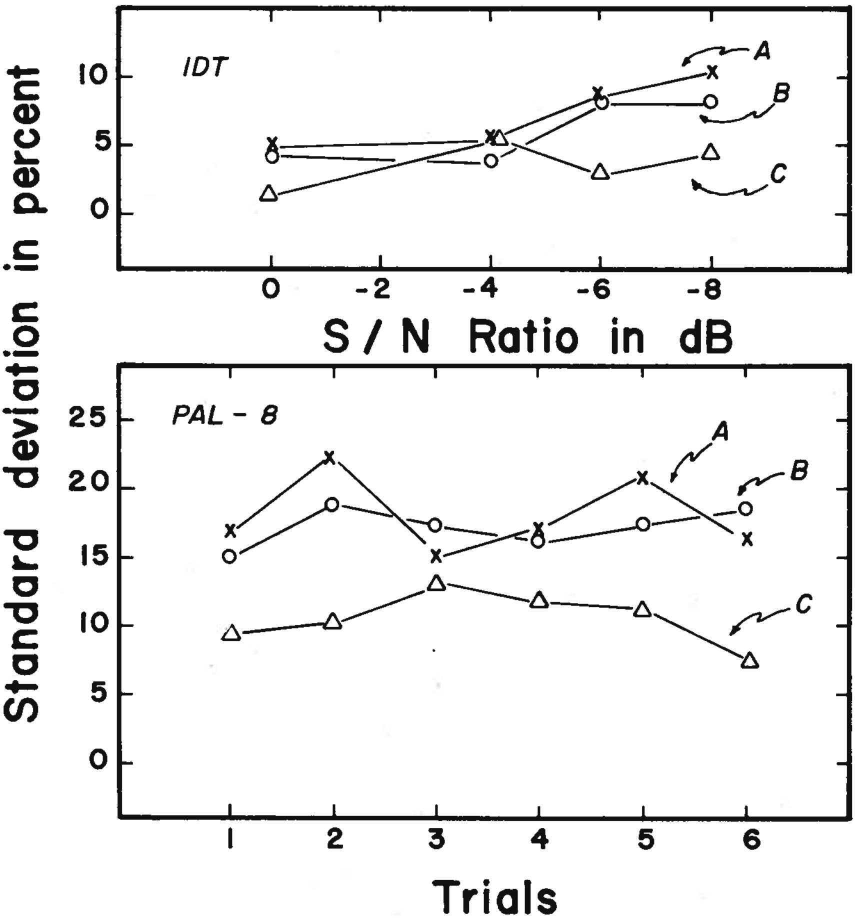


FIGURE 9.—Comparison of inter-signal-variability on Intermodulation Distortion Test and Sentence Intelligibility Test through the same three aids. Performance of different subjects is more alike (standard deviation is smaller) on the Intermodulation Distortion Test than on the sentence materials.

between aids and type or degree of hearing loss. The best aid for anyone appears to be the best aid for everyone.

4. Differences among aids are at least as important to the individual with a mild, flat, conductive loss and good PB discrimination as to the individual with a severe, sloping, sensorineural loss and poor PB discrimination.

5. A behavioral test procedure based on the phenomenon of intermodulation distortion shows great promise as a possible substitute for speech materials in the behavioral evaluation of hearing-aid performance.
6. Suitable techniques now exist for defining, in relatively precise fashion, the relationship between the physical and behavioral dimensions of hearing-aid performance.