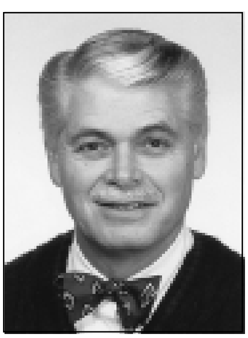
A forum in which two experts present their views on important and controversial issues of clinical or professional significance.

Measuring Hearing Aid Benefit With the APHAB: Is This as Good as It Gets?

Robert G. Paul
Feature Editor, Face-to-Face, American Journal of Audiology

Robyn M. Cox
The University of Memphis, TN



Paul

Cox

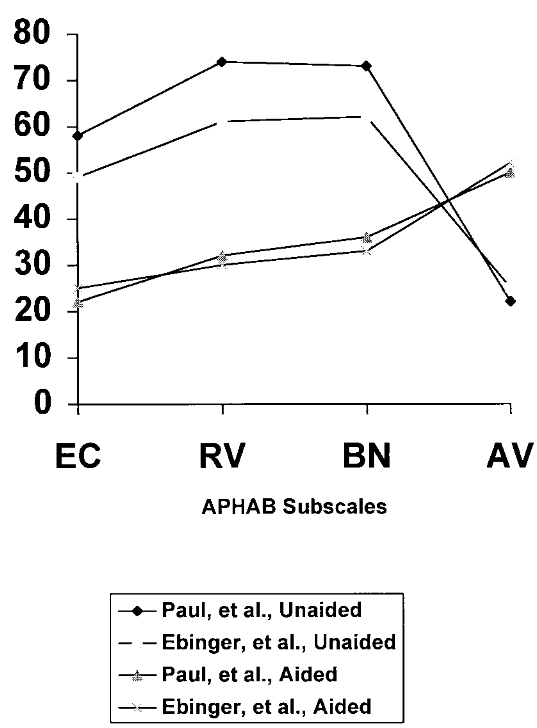
P*aul:* The results shown by Ebinger, Holland, Holland, and Mueller (1995) in a poster presentation at the 1995 American Academy of Audiology meeting caught my eye. They reported on the use of the APHAB as a means to assess whether there were differences in benefit between Class D and K-AMP completely-in-the-canal (CIC) hearing aids.

What interested me was that their subjects seemed to experience the same percentage of problems across the four APHAB listening situations in the aided condition as did subjects in our study (Paul, Talley, Pleasants, & Usrey, 1994) in which we used the PHAB to compare the benefit obtained with each of four different hearing aid gain and output selection approaches. As in the Ebinger et al. study, some of the aids we used were K-AMP, but most were Class D integrated receivers; however, all of our aids were ITEs.

I constructed Figure 1 to show a comparison of the averaged unaided and aided frequency of problems for these two studies. The results for Ebinger et al. are averaged across the two types of hearing aid circuits used, whereas for our study, the results are averaged across the four hearing aid gain and output selection approaches.

Figure 1 shows that the subject groups from these two studies appear to differ, on average, in the frequency of problems they experience across listening conditions without a hearing

FIGURE 1. Frequency of unaided and aided listening problems across APHAB subscales. (Data from Paul, Talley, Pleasants, & Usrey, 1994 and Ebinger, Holland, Holland, & Mueller, 1995).



 1059-0889/95/0403-0010

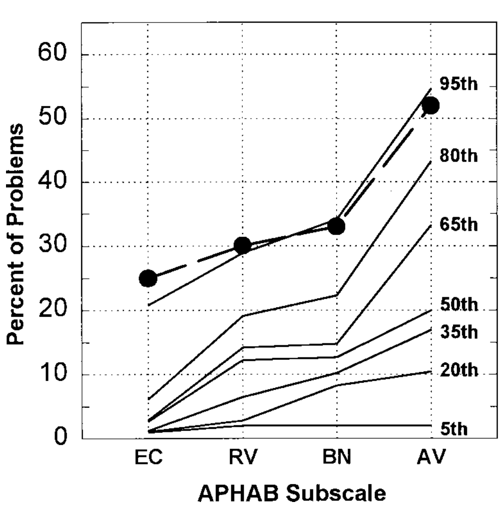
aid; however, the frequency of problems these two subject groups experience with a hearing aid across the four APHAB listening conditions is strikingly similar. I find it disturbing, from the standpoint of using the APHAB as a measure of differences among experimental conditions, that two separate groups of hearing aid users seem to have an almost identical percentage of residual aided listening problems.

I have several questions regarding the similarity of the aided results seen in these two studies. First of all, do the results suggest that these are the limits to the amount of reduction in frequency of listening problems that can be obtained with a hearing aid?

Cox: It seems obvious that there are going to be limits to the amount of improvement in listening problems that we can hope to achieve with a hearing aid. We cannot expect to reach a point where all hearing aid wearers report zero listening problems. The most ambitious goal I can imagine would be to reach the point where the frequency of problems reported by hearing aid wearers has a distribution similar to that reported by persons with normal hearing. We are not there yet, but sometimes we are at least in the ballpark. Figure 2 illustrates what I mean.

This figure shows equal-percentile profiles derived from the responses of 30 young normal hearers to the unaided part of the APHAB. These data tell us the extent to which college-aged persons with normal hearing report listening problems in the situations sampled by the inventory. The percentile level of each line indicates the percentage of persons who reported a smaller proportion of problems in each subscale. As you can see, most young normal hearers reported at least some listening problems. Half of them (above the 50th percentile line) reported that environmental sounds (the AV scale) are unpleasant more than 20% of the time and that they have problems understanding speech in noisy and reverberant situations (the BN and RV scales) more than about 12% of the time. If we gathered similar data on older individuals with normal hearing, the reported proportion of problems would most likely be greater than that seen in Figure 2.

FIGURE 2. Equal-percentile APHAB profiles for normal hearers responding to the unaided portion of the inventory. The percentile label for each line gives the percentage of individuals who reported a smaller frequency of problems than shown by the line. (Filled circles show mean data from Ebinger, Holland, Holland, & Mueller, 1995).



On Figure 2, I have also plotted (filled circles) the mean data you supplied for Ebinger et al. (1995) reflecting the percentage of problems reported by their subjects wearing two types of CIC hearing aids. These data approach the normal range, and, interestingly, for the background noise (BN) and aversiveness (AV) subscales, are slightly better than the data for a small proportion of individuals with normal hearing. The ultimate goal would be for the mean or median data of the hearing aid wearers to approximate the 50th percentile profile of the normal hearers. I don't think we can hope for more. Your question, however, seems to reflect a doubt that we can ever achieve that goal.

Because Paul et al. and Ebinger et al. both reported almost the same mean frequency of aided problems, though using different styles of hearing aids, you are asking whether this is as good as it can get. Actually, I don't find it too surprising that separate studies using the same circuits (Class D and K-AMP) produced similar results, even if those circuits were packaged differently (CIC versus ITE). Although recent data (Kochkin, 1995) suggest that there may be some user-perceived differences between CIC and ITE instruments, most of these cannot be quantified by the APHAB. For example, Kochkin (1995) found that CIC users rated their instruments higher than users of other hearing aid styles on the parameters of visibility, ongoing expense, natural sound, localization, feedback, and usefulness in noisy situations. Of these parameters, only "usefulness in noisy situations" can be measured by the APHAB; and, indeed, Kochkin's APHAB results for some users do show a difference on the AV subscale between CICs and other hearing aid styles.

I think it is important to remember that there were differences in the *unaided* frequency of problems between the Ebinger et al. and Paul et al. subject groups. As Kochkin (1995) points out, our inferences about the meaning of differences in aided results may be compromised unless we compare groups with similar percentages of unaided problems or, alternatively, unless we statistically adjust for any differences in the groups' percentages of unaided problems. This is another issue that might be important in comparing the data from these two studies.

Paul: Do the results suggest that these are the limits to the amount of reduction in frequency of aided listening problems that can be measured with the APHAB?

Cox: I don't see why this would be true. After all, as is shown in Figure 2, most normal hearers respond to the inventory with a smaller percentage of listening problems than the amount reported by these two groups of hearing aid wearers. I cannot think of any inherent reason why hearing aid wearers would respond any differently on this inventory than normal hearing individuals.

Paul: Do the results imply that the measure, aided frequency of listening problems, is insensitive to the clinical decisions we make in selecting gain and output, hearing aid style, or type of circuit?

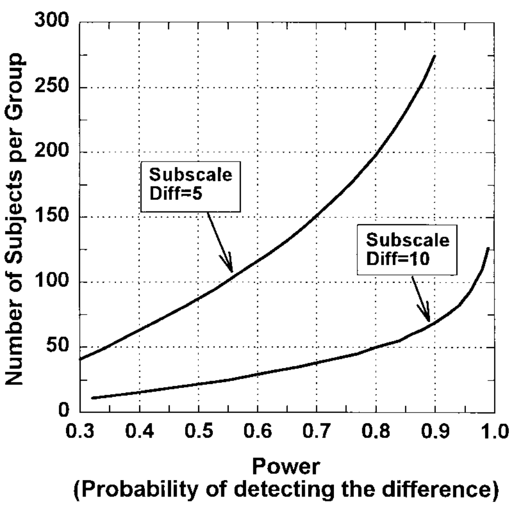
Cox: The sensitivity of the APHAB (and other similar inventories) to differences among conditions deserves careful consideration by both clinicians and researchers. It is essential to recognize that human beings are not as precise as laboratory instruments. Humans show a considerable amount of day-to-day variability in responding to any questionnaire. In addition, the time constraints of clinical practice dictate that our questionnaires cannot be very long, and short questionnaires have more variable results than long ones. As questionnaires go, the APHAB is pretty short.

Clinicians will often use the APHAB to measure the progress of a specific individual, either comparing aided scores with unaided scores or comparing aided scores with one fitting to aided scores with another fitting. How can we use these data? Many clinicians have reported using subscale scores in an intuitive way to diagnose fitting problems and counsel the patient. They have also noted that the process of completing the inventory is quite beneficial for the patient in that it helps to focus him or her on specific areas where performance might be affected by the hearing aid. These factors alone would justify the use of a subjective inventory. However, when it comes to making statements about the reality of differences that you see between two APHAB profiles, it is important to be aware of the statistical constraints. Small differences between scores for the same subscale on two occasions might, or might not, reflect real changes or real differences. You need to know how certain you can be that any difference between two scores indicates a real difference rather than simply measurement error. Guidelines on drawing conclusions from APHAB scores can be found in Cox and Alexander (1995). They present the 90- and 95-percent confidence level critical difference values for APHAB profiles obtained from the same person. These values specify the numerical differences between scores that would allow one to state, at those confidence levels, that an obtained difference is real and not a chance occurrence. For example, if the difference between two scores exceeds, say, the critical difference value for the 95-percent confidence level, then one could be assured that out of 100 such comparisons, a difference of this magnitude would occur by chance in fewer than five comparisons. In other words, one is not very likely to see a difference this big unless there is a real difference between the hearing aid fittings being compared.

In addition to the above information, the Cox and Alexander article also specifies equal-percentile profiles that compare an individual's scores to those of experienced, regular hearing aid wearers. These profiles allow one to put the patient's fitting outcome in perspective by comparing the results to normative data. Such a comparison might provide beneficial information to the audiologist, the patient, and—potentially—to third-party payers.

Researchers use the APHAB to explore the effects of differences in circuits or other variables. In this application, the most important thing to consider is probably the number of subjects needed to yield a reasonable chance of detecting a difference if it really exists. This number depends on the following factors: the size of differences in subscale scores that you think is of practical importance; the certainty with which you wish to be able to conclude that the difference is real (the alpha level of your statistical test); and the variability of subscale scores in your subject population. Figure 3 illustrates how these issues are related. Assume that you are planning a study to compare CIC circuit X with CIC circuit Y. You intend to recruit two groups of subjects, fitting members of group 1 with circuit X, and fitting members of group 2 with circuit Y. After an adjustment period, you will have the members of each group complete the APHAB. To determine

FIGURE 3. Relationship between the number of subjects in each experimental group and the probability of detecting a difference that exists between APHAB subscale scores. Data are given for differences of 5 and 10.



whether the two circuits perform differently, you will compare the mean scores of the two groups on each of the four APHAB subscales. Figure 3 depicts the relationship between the number of subjects in each experimental group and the probability that the experiment will be able to detect a difference if it exists. To generate Figure 3, I assumed that the between-subject standard deviation of subscale scores is 20. The upper solid line in the figure gives the results for a difference of 5 between two subscale scores. The lower curve gives the corresponding results for a difference of 10.

Personally, I think that a difference of 10 between the two circuits would probably be noticeable to the hearing aid wearer, but I'm not sure whether a difference of 5 would have practical significance.

If we look at the line for a difference of 10, we see that in order to have an 80-percent probability of statistically detecting a difference, if it exists, we need to use at least 50 subjects in each group. If we want to have an 80-percent chance of detecting a difference of 5 between the two circuits, we need to use a whopping 200 subjects in each group! These are big subject numbers and difficult for many of us to achieve. However, as Figure 3 illustrates, using fewer than about 50 subjects per group really diminishes the chance of finding a significant difference.

On the other hand, if large enough groups are used, APHAB scores can detect quite small differences between experimental conditions (e.g., Kochkin, 1995).

Paul: What else do you think can be done with hearing aids to reduce the percentage of aided listening problems?

Cox: Further progress can occur on at least two fronts. The first will be improvements in the hardware capabilities of amplifying devices. The field, in general, is energetically exploring the potential of different signal-processing strategies. The most promising technologies that are within our grasp right now seem to be those that include various forms of compression and multimicrophone techniques. If we can solve the problems encountered in evaluating these devices, I expect to see major advances within the next few years.

The second line of attack is available to us right now; namely, improving the services we offer to help hearing aid users to learn how to optimize the benefit they receive from their instruments. I am often surprised to learn how little is being done with postfitting management. As a profession, audiology should do much more in this area. The time invested should result in a more successful, informed, and satisfied hearing aid consumer.

References

Cox, R. M., & Alexander, S. C. (1995). The abbreviated profile of hearing aid benefit. *Ear and Hearing, 16*, 176–186.

Ebinger, K., Holland, S., Holland, J. W., & Mueller, H. G. (1995). *Using the APHAB to assess benefit from CIC hearing aids.* Poster presented at the meeting of the American Academy of Audiology, Dallas, TX.

Kochkin, S. (1995). Customer satisfaction and benefit with CIC hearing instruments. *The Hearing Review, 2,* 16, 19–26.

Paul, R. G., Talley, J. N., Pleasants, D. M., & Usrey, S. C. (1994). *Perceived benefit and use-gain compared to target at hearing aid fitting and one year later.* Poster presented at the meeting of the American Academy of Audiology, Richmond,VA.

Received July 19, 1995
Accepted August 21, 1995

Contact author: Robyn M. Cox, PhD, Professor, School of Audiology and Speech-Language Pathology, The University of Memphis, 807 Jefferson Avenue, Memphis, TN 38105

Key Words: hearing aid, APHAB, benefit, self-assessment