Development of the Profile of Hearing Aid Performance (PHAP)

Article *in* Journal of speech and hearing research · July 1990

DOI: 10.1044/jshr.3302.343 · Source: PubMed

Robyn M Cox
The University of Memphis

Christine Gilmore Eubanks
Longwood University

Journal of Speech and Hearing Research, Volume 33, 343–357, June 1990

DEVELOPMENT OF THE PROFILE OF HEARING AID PERFORMANCE (PHAP)

ROBYN M. COX CHRISTINE GILMORE
Department of Audiology and Speech Pathology, Memphis State University, and Veteran's Administration Medical Center, Memphis, TN.

This paper reports the development of the Profile of Hearing Aid Performance (PHAP), a 66-item self-administered inventory that quantifies performance with a hearing aid in everyday life using both seven-score and four-score profiles. The profiles assess experience with amplification in terms of speech communication in three types of listening situations and in terms of reactions to amplified environmental sounds. The PHAP has been evaluated using groups consisting mainly of elderly hearing aid wearers having mild to moderate hearing impairments. Internal consistency reliability of the scales and subscales ranges from .70 to .91. Test-retest correlations range from .66 to .88. Ninety percent and 95% critical differences are presented to facilitate evaluation of differences between scores for the same individual under different conditions. It is envisioned that the PHAP will have its principal application in research evaluating and comparing different approaches to hearing aid fittings. In addition, the inventory can be used clinically to assess existing hearing aid fittings.

KEY WORDS: hearing aids, questionnaire, hearing aid performance, hearing aid benefit, profile

Many laboratory studies of performance with hearing-aid-processed stimuli have been reported and the results have been used to support theoretical approaches to hearing aid fitting. Following the lead set by this research, clinical measurements of insertion gain or ability to recognize speech in noise, for example, often form the basis of amplification recommendations. Although this approach to hearing aid provision is valuable, necessary, and practical, the ultimate appraisal of the real-world efficacy of a hearing aid fitting demands valid quantification of the effects of hearing aid use in everyday life.

Investigations with hearing-impaired individuals have verified the frequently heard assertion that the most important problem they experience is difficulty in understanding conversational speech (Barcham & Stephens, 1980; Hagerman & Gabrielsson, 1984). Thus, ability to communicate in everyday life situations while wearing a hearing aid must be a primary factor in determining the caliber of the hearing aid fitting. Another issue that frequently has been implicated in the success of hearing aid rehabilitation relates to the amplification of environmental sounds. Numerous studies have shown that dissatisfaction with either the level of or the interfering nature of background sounds is a primary reason for rejection of hearing aids (e.g., Berger et al., 1983; Brooks, 1985; Franks & Beckman, 1985). Thus, the adverse effects of amplifying environmental sounds could be expected to offset, to some extent, improvements in speech understanding that might be realized with hearing aid use. These considerations suggest that quantification of the overall effects of hearing aid use in everyday life should include attention to the positive effects of improved speech understanding and the negative effects of objectionable environmental sounds.

A measure of users' experiences wearing hearing aids in everyday life can quantify either performance with the hearing aid or hearing aid benefit. A measure of performance with the hearing aid determines functioning in everyday circumstances on an absolute scale. For example, it could determine what proportion of times a particular situation presents problems for the individual when the hearing aid is worn. A measure of hearing aid benefit determines functioning in everyday situations when the hearing aid is worn relative to functioning in the same situations when no hearing aid is worn. In other words, a measure of hearing aid benefit quantifies the change in performance that is attributable to hearing aid use.

Sometimes an appraisal of hearing aid benefit is of limited value without a concomitant consideration of absolute performance. For example, a measure of hearing aid benefit may determine that very little benefit is obtained with a particular hearing aid when the wearer is conversing with family members. This would be a matter for serious concern if absolute performance in conversing with family members was poor. On the other hand, if the hearing aid wearer experiences few problems conversing with family members (i.e., performance is good on an absolute scale), lack of significant hearing aid benefit in this type of situation may not present a problem.

Throughout this article, the term *hearing aid performance* is used to denote a listener's performance, on an absolute scale, while wearing a hearing aid. The term *hearing aid benefit* signifies the change in listener's performance that is attributable to hearing aid use.

A number of self-assessment tools for the hearing impaired have appeared in the literature. Noteworthy contributions in this area have been made by High, Fairbanks, and Glorig (1964), Giolas, Owens, Lamb, and Schubert (1979), Ventry and Weinstein (1982), and Demorest and Erdman (1987), among others. Most of this work has been directed towards the measurement of hearing handicap or communication problems in everyday life situations. Although parts of some of these inventories might be adapted to quantify performance with

  0022-4685/90/3302-0343$01.00/0

hearing aids, this was not the original purpose for which they were developed. Efforts have been made to quantify hearing aid benefit in terms of the reduction of hearing handicap scores (e.g., Newman & Weinstein, 1988; Tannahill, 1979). This approach, which produces a single number to characterize benefit, can result in a useful global measure of change related to hearing aid use. However, a single index depicting reduction in hearing handicap does not yield analytic information about hearing aid performance that might indicate appropriate directions for further rehabilitation efforts.

A more analytic inventory for self-assessment of hearing aid benefit was developed by Walden, Demorest, and Hepler (1984). These investigators devised a 64-item questionnaire, the Hearing Aid Performance Inventory (HAPI), to explore 12 bipolar features that were hypothesized to be relevant to success with amplification. Each item described a specific situation. Subjects responded by choosing one of five categories to describe the help provided by the hearing aid in that situation. After analysis of results from 128 hearing aid wearers, four major groupings of everyday situations emerged. These were: noisy situations, quiet situations with the talker nearby, situations with reduced speech cues, and nonspeech stimuli. Four subscales were defined to quantify each of these factors separately. Internal consistency reliability was high, especially for the first three subscales.

Overall, the outcome of the study by Walden et al. (1984) suggested that self-assessment of hearing aid benefit offered a promising avenue for validation and comparison of hearing aid fittings. However, the HAPI does not offer an evaluation of the negative effects of amplified environmental sounds. Also, because the scale assesses hearing aid benefit directly, it is not possible to measure hearing aid performance, as defined here, using the HAPI. In addition, the sensitivity of this instrument may be limited by the small number of response alternatives. Finally, interpretation of differences between scores obtained in successive administrations of the HAPI is complicated by the fact that critical differences for the different scales have not been reported.

The work reported in this article was undertaken to develop a new self-assessment inventory to quantify users' experience with hearing aids in daily life. This inventory is called the Profile of Hearing Aid Performance (PHAP). It was designed to measure two aspects of performance with hearing aids: (a) speech communication in a variety of typical workaday situations, and (b) reactions to the loudness or quality of environmental sounds. It was determined in advance that the PHAP would quantify hearing aid performance rather than hearing aid benefit. However, the items and response mode were designed so that, with minor modifications, the inventory could be used to quantify listeners' performance without, as well as with, hearing aids. Using this approach, hearing aid benefit could be derived, if desired, by comparing aided and unaided responses.

It is anticipated that the inventory will have its major application in research efforts. It could be used, for example, to quantify, for later comparison, performance with hearing aids that have been selected according to different fitting rationales. Likewise, the effectiveness in daily life of different signal processing strategies could be assessed using the PHAP. Furthermore, it is expected that data obtained using this inventory will allow the development of mean profiles for different groups of hearing aid users and that comparison of these profiles may shed light on long-standing questions about the efficacy of different approaches to amplification. Potentially interesting comparisons would include monaural versus binaural fittings, older versus younger individuals, and successful versus unsuccessful hearing aid users.

In the clinical realm, it is not expected that results of the PHAP will be used for initial selection of hearing aids. However, the inventory could be used to evaluate an existing hearing aid fitting or a recently fitted instrument. Further discussion and examples of possible clinical applications are presented in a later section of this article. The complete inventory is presented in the Appendix.

The PHAP was developed in the following stages: (a) Content domains were defined and initial items were developed to sample these domains; (b) a large number of hearing aid users responded to the initial pool of items; (c) based on analyses of these initial responses, items were deleted and scales and subscales were constructed; (d) internal consistency of these scales and subscales was investigated with a second group of subjects; and (e) test-retest reliability of the final scales and subscales was investigated and critical differences were determined to facilitate the evaluation of score differences.

Experiment 1: Selection of Items

METHOD

Design objectives for the PHAP included the following: (a) Hearing aid performance would be quantified in four scales, encompassing communication in three basic types of listening situations and the effects of amplified environmental sounds; (b) these scales would be broken into subscales if this appeared appropriate, based on the data obtained; (c) a relatively large number of response categories would be provided; (d) the items and response mode would be designed so that the inventory measured hearing aid performance as defined earlier.

Both theoretical considerations and the data of Walden et al. (1984) suggest that three basic listening environments can be defined. These place distinctly different demands on the listener and together represent a large proportion of everyday listening situations experienced by the typical hearing aid wearer. These three environments have also been used by us in other studies (e.g., Cox, Alexander, & Gilmore, 1987). Environment A represents communication in a situation in which speech is at normal conversational level, visual cues are fully available, and background noise and reverberation are low.

Examples of Environment A include face-to-face conversation in a typical living room or quiet office. Environment B represents communication in a situation in which external environmental sound is low but speech cues are reduced because of reverberation, low speech intensity, or limited or absent visual cues. Examples of Environment B include listening as an audience member to a lecture delivered in an unamplified classroom, communicating over a distance, and listening to someone whose face is not visible. Environment C represents communication in a situation where external environmental sound is relatively high, speech levels are somewhat raised, and visual cues are available. Examples of Environment C include face-to-face communication at a social event with numerous people present and communication with a clerk in a busy store.

Most of the PHAP items exploring communication in the three basic environments were adapted from those used by Walden et al. (1984), Demorest and Erdman (1986) and Chung and Stevens (1986). Environments A, B, and C were represented by 20, 20, and 28 items, respectively.

Studies of noise annoyance reported by Angevine (1975), Graeven (1975), and Weinstein (1978) suggested that several features in addition to overall level are important in the annoyance or acceptability of environmental sounds. Features identified for inclusion in the initial PHAP items were: quality or naturalness; predictability (anticipated vs. unexpected); environment (inside vs. outside); repetition (frequent vs. occasional); and controllability (self-generated vs. other-generated). A total of 32 items were written for this content area with 9–12 items for each feature. Most of the items sampled more than one feature; for example, a slamming door was categorized as both unpredictable and infrequent.

Each of the 100 original items was a statement: for example, "I miss a lot of information when I'm listening to a lecture." Preceding the first item and repeated at the top of each subsequent page was the stem "when I wear my hearing aid:" (to evaluate performance without a hearing aid, this stem is omitted). The respondent's task was to choose from a 7-point response scale to indicate the proportion of occasions on which the statement was true. The response scale was taken from Hutton (1987). Each point consisted of both a descriptor and an associated percentage. They were: Never (1%), Seldom (12%), Occasionally (25%), Half-the-time (50%), Generally (75%), Practically Always (87%), and Always (99%). If subjects had not experienced the exact situation described, they were encouraged to imagine how they would respond in a similar situation. In addition, a response of N/A for *not applicable* was provided. Responses of N/A were treated as missing data. Most of the items were worded so that a response of *always* signified frequent problems in the described situation. However, to discourage response bias, 41 items were worded so that a response of always was indicative of few problems in the described situation. In the scoring process, the responses for these items were reversed so that for all items a high percentage score indicated a high proportion of problems.

Subjects

The PHAP was developed to evaluate hearing aid performance in adults who had experience wearing amplification in daily life. There were no limitations placed on this group of intended users. Thus, in evaluating the inventory, an effort was made to sample widely among adult hearing aid users without regard to age, length of hearing aid experience, hours of use per day, or satisfaction with amplification.

Subjects for Experiment 1 (Group 1) were 225 individuals. Distribution of age, reported hearing aid experience, and reported daily hearing aid use are shown in Figures 1, 2, and 3, respectively. These figures indicate that the subjects covered a wide age range but 73% of them were more than 60 years old. Forty-eight percent reported hearing aid use of 7–16 hrs/day, 46% reported using their aids 1–7 hr/day, and 6% used their aids less than 1 hr/day. Twenty-eight percent were relatively new hearing aid wearers, having worn hearing aids for less than 1 year.

Information about hearing loss extent and configuration was available for 142 (63%) of the subjects. There is no obvious reason why this subgroup would not be representative of the entire subject group. These data are shown in Table 1. Most of the subjects had mild or moderate hearing losses with flat or gently sloping audiometric configurations.

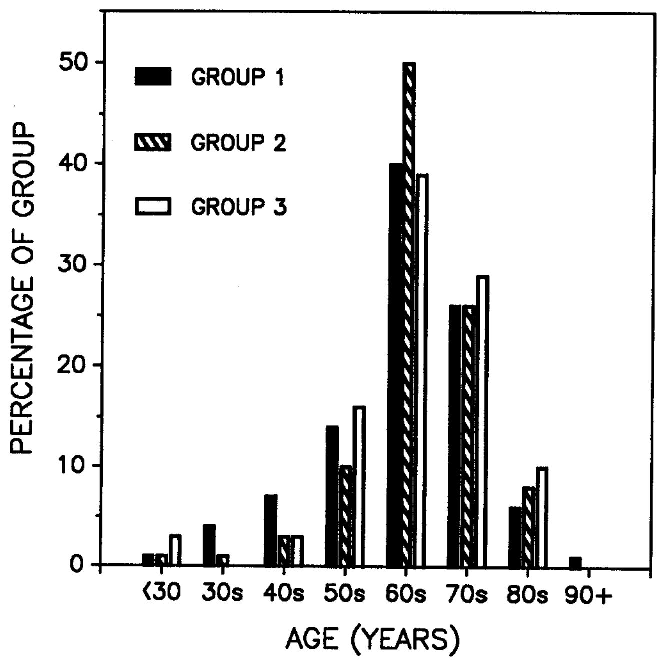

Figure 1. Distribution of ages for the three subject groups.

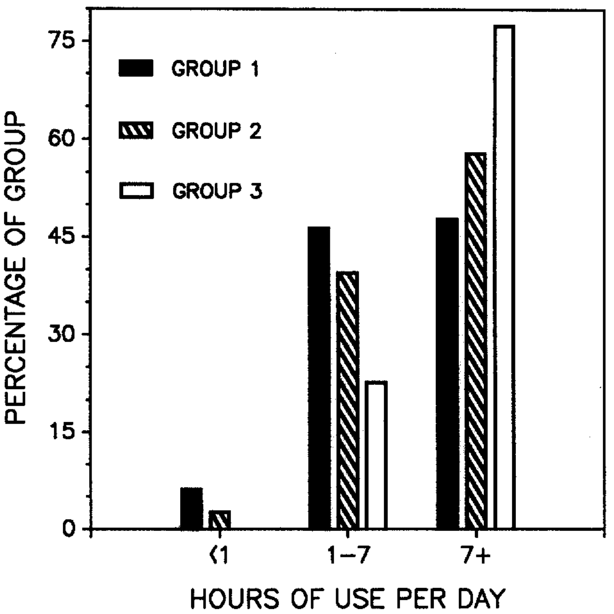

FIGURE 2. Distribution of reported hours of hearing aid use per day for the three subject groups.

Procedure

The 100 original items were printed in random order. The inventory was then distributed by mail to 412 current or previous hearing aid wearers. The letter accompanying the inventory noted that it was not necessary for the respondent to be a current hearing aid wearer: the respondent was eligible to complete the inventory if he or she had worn a hearing aid long enough to have formed opinions about its performance in the various situations

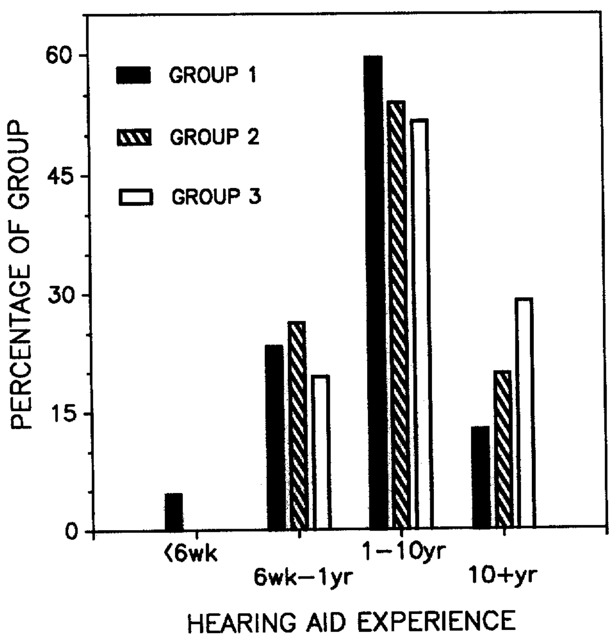

FIGURE 3. Distribution of reported duration of hearing aid experience for the three subject groups.

queried in the questionnaire. Usable responses were received from the 225 individuals that comprised Group 1, described above. A further 25 were unable to complete the inventory for a variety of reasons. The rest did not respond.

RESULTS

Responses to each item were scored in terms of the percentage corresponding to the selected answer. A few items were eliminated because inspection of the data suggested that they were confusing or anomalous in some way. The remaining items were divided into three groups: listening under favorable conditions (Environment A), listening under unfavorable conditions (Environments B and C), and environmental sounds. Each of the three groups of items was subjected to principal components analysis with varimax rotation using SPSS/PC+, version 2.0 (Norusis, 1988). The purpose of this analysis was to explore the potential for dividing the items into meaningful subscales.[1] Only factor loadings greater than .33 were considered in interpretation of the results (see Tables 2, 3, and 4).

Nineteen items representing favorable conditions (Environment A) were analyzed. The first two factors accounted for 50% of the total variance. The first factor was interpreted as representing listening to the speech of familiar talkers. The second factor was interpreted as ease of communication under favorable conditions. Most of the items in the latter factor include wording such as "I have to strain" or "I have difficulty." Two subscales were constructed for Environment A, focussing on the *familiar talkers* and *ease of communication* areas.

Forty-five items representing communication under unfavorable conditions were analyzed. The first three factors accounted for 51% of the total variance. The interpretation of these factors was not as clear as in the analysis of data for favorable listening conditions. Most of the items classified as representing listening to speech in a noisy setting (Environment C) had loadings on Factor 1 or Factor 3 with Factor 1 representing unsuccessful communication and Factor 3 comprising items depicting successful communication. Most of the items considered to represent listening to speech in a reverberant setting had loadings on Factor 2. The items representing reduced cues (low intensity or low visual cues) did not show any strong commonality.

Because items representing different unfavorable listening categories (noise, reverberation, and reduced cues) did not fall clearly into separate factors, these data suggest that the different types of signal degradations may have similar effects on performance with hearing aids, at least for the instruments worn by the Group 1 subjects. Thus, assessment of hearing aid performance separately

[1]It was necessary to divide the items into groups before analysis in order to obtain the subject-to-item ratio required for factor analysis (Nunnally, 1978, p. 334).

in the different types of unfavorable listening situations may be somewhat redundant. However, there is mounting evidence suggesting that speech interference effects may be different for reverberation and background noise (Cox et al., 1987; Helfer, 1989; Nabelek, 1988, 1989). In addition, there is considerable interest, currently, in exploring signal processing and shaping schemes to optimize speech understanding in different types of unfavorable listening conditions. These considerations suggest that it may be useful to maintain the division between noisy and reverberant environments to allow separate assessment of the effects of these two types of degradation. In addition, because the content area encompassed by the reduced cues items may produce important clinical information (e.g., a hearing aid worn with insufficient gain may give unusually low scores on these items), it was also retained. Two subscales (reverberation and reduced cues) were constructed to represent Environment B. Items representing Environment C were grouped into a single scale, without defined subscales.

Thirty-two items exploring effects of environmental sounds were analyzed. The first two factors accounted for 38% of the total variance. Factor 1 was interpreted as representing the aversiveness of environmental sounds. Factor 2 was interpreted as representing distortion of environmental sounds. Many of the environmental sound items appeared to have little in common with other items. There was no evidence in the outcome of this analysis to support the importance of predictability, environment, repetition or controllability in an individual's responses to amplified environmental sounds. Two subscales were defined to measure effects of environmental sounds, focusing on *aversiveness* and *distortion*.

To determine which items to include in the various subscales, item analyses were performed using the SPSS/PC+, version 2.0 reliability procedure. All items had satisfactory mean scores (not too extreme) and fairly broad response distributions. Item retention was based on the following general guidelines: (a) items with low corrected item-total correlations were deleted, (b) items answered N/A by a large number of subjects were deleted, (c) items with high factor loadings were give preference, (d) the number of items was further reduced with the goal that coefficient alpha for subscales should be >.85 and for scales should be >.90 (recall that four scales were determined a priori to represent effects of environmental sounds and communication in Environments A, B, and C).

Ultimately, 66 items were retained for the final inventory. Twenty-one of these are reversed for scoring. Item statistics for the subscales are shown in Tables 2, 3, and 4. Summary statistics for scales and subscales are given in Table 5. Table 6 gives the intercorrelations between the various scales and subscales.

The profile may be thought of as consisting of 7 subscale scores or 4 scale scores. Because a profile of 7 scores is potentially more informative than one consisting of four scores, both configurations were retained although further experience with the inventory may indicate that one should be abandoned. The Familiar Talkers (FT) and Ease of Communication (EC) subscales are combined to form the Speech, Environment A (SA) scale. The Reverberation (RV) and Reduced Cues (RC) subscales are combined in the Speech, Environment B (SB) scale. The Aversiveness of Sounds (AV) and Distortion of Sounds (DS) subscales contribute the items of the Environmental Sounds (ES) scale. The Background Noise (BN) subscale is the same as the Speech, Environment C (SC) scale. Because BN and SC contain the same items, it is not accurate to refer to BN as a subscale of the Environment C scale. Nevertheless, this terminology facilitates discussion of the 7-score and 4-score profiles. Thus, when used as part of the 7-score profile, these 16 items are referred to as the BN subscale whereas, when the discussion centers on the 4-score profile, the same items are termed the SC scale.

Perusal of the mean scores for the scales and subscales in Table 5 indicates that the hearing aid wearers reported fewer problems understanding speech in Environment A situations than in either Environments B or C. This is consistent with many anecdotal reports of hearing aid performance in everyday life. Also, Environment B and C situations provided very similar mean data—subjects reported difficulty about half the time in situations exemplifying both types of environments. Furthermore, it is noteworthy that the average subject reported difficulty with environmental sounds in about 40% of the situations in which the sounds were experienced.

Table 6 shows that the correlations between the subscales were generally in the range .40–.60. These correlations are significantly different from zero ($p < .01$) and, thus, indicate that the various scales are not measuring completely independent aspects of hearing aid performance. However, because a correlation between two scales of .60 indicates that only 36% of the variability in scores on one scale can be attributed to the variability in scores on the other scale, these data also suggest that the various subscales assess aspects of hearing aid performance that are related but different. An exception to this overall outcome is seen in the results for subscales RV, RC, and BN. Higher correlations (.81 to .87) among these three subscales reveal that they assess aspects of performance with hearing aids that are more similar, at least for the generation of hearing aids (circa 1987) worn by the

TABLE 1. Classification of hearing losses for Group 1 subjects. The data shown are for 63% of the subjects and are probably representative of the entire group. Data have been included for both ears of all subjects (exception: each of four subjects had one dead ear). Data are in percentages. SRT = speech reception threshold for spondee words (dB HL re ANSI, 1969). Slope = slope of audiogram from 500 to 4000 Hz in dB/octave.

		Slope		
SRT	*<6*	*6–14*	*>14*	*Total*
<40	10	36	19	65
40–60	15	13	3	31
>60	3	1	0	4
Total	28	50	22	100

TABLE 2. Item statistics and factor loadings for items retained from the favorable listening conditions (Environment A) category. Items are numbered according to their appearance in the complete inventory, given in the Appendix. $Corr_{i-t}$ = correlation between item and other items of same subscale. F1 = Factor 1, etc.

Condition	*Mean*	*SD*	$Corr_{i-t}$	*F1*	*F2*
Ease of Communication					
12. Conversation, people nearby	59.0	31.3	.52		.49
17. Conversation with a companion	32.0	27.4	.71	.37	.57
25. Speaker, small group	35.8	31.5	.66		.68
29. Conversation with my doctor	28.6	29.8	.58	.34	.45
33. One-on-one conversation	34.3	29.9	.68		.79
34. One other person at home	31.9	28.5	.67		.77
38. Small office, interviewing	37.8	30.8	.58		.47
Familiar Talkers					
1. Small group, no noise	23.3	20.8	.61	.68	
20. Walking with a friend	20.0	17.9	.66	.75	
32. Family, normal voice	28.0	25.4	.54	.67	
48. Quiet dinner with family	22.4	21.7	.63	.67	
52. Newscaster, TV news at home	15.2	17.2	.63	.76	
61. Conversation, family member	14.3	18.2	.68	.76	
66. Talking to a bank teller	22.6	21.7	.51	.55	

Group 1 subjects. The justification for retaining the separation of these subscale areas, as noted above, is the prospect of using them in an attempt to evaluate a new generation of signal processing hearing aids that might perform differentially for the different types of degradations encompassed by these subscales.

EXPERIMENT 2: INTERNAL CONSISTENCY

The internal consistency reliability of the scales and subscales, as indicated by coefficient alpha in Table 5, were all quite high. The .82 value of coefficient alpha for the RC subscale was the lowest and reflects the lower commonality of items in that subscale noted in the results of the principal components analysis. However, it is important to keep in mind that when statistical criteria are used to select items from a pool, sampling errors will influence the choice of items, to some extent. When these same items are used with a different group of subjects, it can be expected that the internal consistency reliability will be lower than measured in the original group. To assess the extent to which measurement errors may have led to the selection of inappropriate items, it is necessary to re-evaluate internal consistency with a new group of subjects. To this end, the inventory was administered to a second group of subjects. Determination of coefficient alpha on this new group was expected to indicate the extent to which it would be reasonable to generalize scores for the various PHAP subscales beyond the actual items to other similar items within that content domain.

METHOD

Subjects

Seventy-six hearing aid wearers served in Group 2. Their ages, reported hours of hearing aid use per day, and reported years of hearing aid experience are shown in Figures 1, 2, and 3, respectively. Overall, this group of subjects was quite similar to Group 1.

Procedure

Forty-five subjects completed the inventory by mail. Thirty-one subjects completed the inventory during a visit to either an audiology office or the audiology department of a Veterans Administration hospital. The inventory was always administered as a paper-and-pencil task and the instructions were printed on the form. The N/A response option was deleted with the rationale that item selection had eliminated items that were likely to be difficult to answer.

RESULTS

Scoring and statistical analyses were identical to those used on the original data set. Items were occasionally omitted and these were treated as missing data. Table 7 gives internal consistency and inter-item correlation statistics for the 66 items. Comparison of Tables 5 and 7 reveals that, for 3 of the 7 subscales (FT, RC, and DS) coefficient alpha was substantially reduced relative to values for the original group of subjects. This outcome indicates that the items in these subscales are not as closely related to each other as data from Group 1 suggested. Consequently, the scores for these subscales may be only moderately accurate predictors of scores for other items from the same content areas. Internal consistency was quite high for subscales EC, RV, BN, and AV and was similar to the values determined for the original group of subjects. This suggests that scores for these subscales are good predictors of scores for different items sampling the same content areas.

TABLE 3. Item statistics and factor loadings for items retained from the unfavorable listening conditions (Environments B and C) category. Items are numbered according to their appearance in the complete inventory, given in the Appendix. $Corr_{i\text{-}t}$ = correlation between item and other items of same subscale. F1 = Factor 1, etc.

Condition	*Mean*	*SD*	$Corr_{i\text{-}t}$	*F1*	*F2*	*F3*
Background Noise						
5. Busy department store	50.8	30.1	.74	.42	.40	
7. Speaker in a small group	48.6	29.6	.56			.69
8. Conversation, several people	62.9	29.5	.70			.66
10. Ask people to repeat	60.2	30.5	.77	.59	.31	
11. Crowded grocery store	38.0	27.7	.62			.37
16. Crowded reception room	44.5	31.4	.67	.33	.38	
18. News, family members talking	71.8	27.7	.76	.75		
23. Busy street, asking directions	52.8	31.2	.68	.45		
27. In a crowd, trouble hearing	72.9	27.9	.66	.78		
28. Dinner with several people	50.2	31.6	.73	.61		
30. Air conditioner on	50.5	31.1	.62	.47		
42. Restaurant, waitress questions	41.2	29.1	.60			
44. Large noisy party	76.5	28.5	.72	.70		.36
56. Outdoors on a windy day	55.0	30.2	.66	.41	.37	
62. Meeting, several other people	40.2	27.8	.58			.39
65. Communicate in a crowd	48.0	29.5	.70			.65
Reverberation						
2. Listening from rear of room	78.1	24.4	.65	.54	.35	.40
14. Question from back of room	56.9	36.3	.67	.51	.38	
22. Converse across empty room	56.7	32.5	.51			.63
26. Dialogue, movie or theater	48.4	31.6	.60	.36	.63	
31. Listening to a lecture	49.3	32.1	.79	.35	.60	.33
39. Follow lecturer's instructions	39.9	27.8	.71		.50	.43
45. Theater, people whispering	63.4	29.8	.57	.32		.63
53. Words of a sermon	36.9	29.9	.64		.65	
58. Lectures or church services	48.1	32.8	.60		.73	
Reduced Cues						
4. Most people speak too softly	60.9	30.9	.49	.36	.35	
6. Family speaking softly	43.6	31.2	.59			.40
35. Back seat, listen to driver	69.8	28.2	.61	.55		
37. Teller, drive-in window	30.6	26.1	.50			
46. Quiet restaurant, soft voice	32.2	25.8	.50			
47. TV, volume set by normal hearer	51.5	35.7	.50		.45	
49. News on car radio	19.0	18.5	.37			
54. Talking, person in other room	66.5	29.9	.62	.60		
59. Overhear conversation outside	84.1	22.2	.46	.66		

Internal consistency of the 4 scales declined somewhat in this second group of subjects but, overall, remained fairly high. Even though the scales contain a greater number of items than the subscales (except for Environment C which is the same as subscale BN), their greater diversity of content is reflected in the fact that this increase in item numbers did not result in a substantial improvement in internal consistency.

EXPERIMENT 3: TEST-RETEST RELIABILITY

Because the PHAP will be used for the comparison of hearing aid conditions in individual subjects (e.g., aided vs. unaided or hearing aid #1 vs. hearing aid #2), it is important to know the extent of measurement error associated with a single administration of the inventory. Such errors have their basis in random fluctuations in such factors as mood, health, and alertness from one test to another. An estimation of measurement error is required for the construction of critical differences that may be used to evaluate the significance of differences between scores for the same individual in two or more conditions. If an obtained difference between two scores exceeds, for example, the 95% critical difference, it may be concluded with 95% certainty that there was a real difference between the two tested conditions. There is a 5% probability that the obtained difference occurred by chance.

There are numerous opportunities for error in studying test-retest reliability of a self-assessment inventory because factors other than random fluctuations in such factors as mood, health, and concentration may affect a subject's responses. For example, the process of responding to the items on the first test may change the subject's perception of the content area, and consequently, have an impact on subsequent responses to those same items. Also, a subject's opinions about the item content areas may undergo a true change between tests. Both of these circumstances would be expected to deflate the measured retest reliability. On the other hand, subjects may remem-

TABLE 4. Item statistics and factor loadings for items retained from the environmental sounds category. Items are numbered according to their appearance in the complete inventory, given in the Appendix. $Corr_{i\text{-}t}$ = correlation between item and other items of same subscale. F1 = Factor 1, etc.

Condition	*Mean*	*SD*	$Corr_{i\text{-}t}$	*F1*	*F2*
Aversiveness of Sounds					
9. Telephone ring	26.2	28.8	.57		
13. Car horn makes me jump	36.2	34.1	.65	.65	
15. Construction work	57.7	34.9	.71	.64	
19. Screeching tires	49.0	35.3	.66	.73	
21. Fire engine siren	40.9	36.9	.61	.56	.34
24. Running water	32.1	33.0	.55		
41. Household appliances	25.5	31.2	.58		
43. Crowd noise	39.8	35.3	.58	.49	
55. Smoke detector	53.6	36.8	.60	.59	
57. Everyday sounds too loud	31.8	30.6	.56	.52	
63. Traffic noises too loud	45.5	34.7	.67	.38	.41
64. Glass breaking	44.2	36.0	.55	.34	
Distortion of Sounds					
3. Women's voices "shrill"	48.9	33.3	.60		.66
36. Music sounds distorted	30.7	30.4	.60		.55
40. Everyday sounds not clear	48.3	33.0	.57		.52
50. Quality of music	44.3	32.9	.64	.39	.53
51. Telephone sounds "tinny"	40.7	34.5	.60		.66
60. Voices sound unnatural	32.3	30.2	.63		.78

ber their responses to items on the first test and intentionally replicate these responses on the second test occasion. Obviously, this would produce a spuriously high test-retest correlation. In spite of these potential pitfalls, an estimate of test-retest reliability seemed essential before the PHAP inventory could be employed for one of its intended purposes. Thus, a study was undertaken.

Subjects

Group 3 was composed of 30 hearing aid wearers. These individuals were a portion of the group of 76 who served in the investigation of internal consistency, described above. Their ages and reported hearing aid use and experience are summarized in Figures 1, 2, and 3. To minimize the likelihood of a true change in subjects' opinions about the performance of their hearing aids, they were required to have worn the present instrument for at least 3 months before the first test and to use the instrument for a minimum of 4 hours per day. Thus, as scrutiny of Figures 1–3 reveals, these subjects were similar in terms of age to the first two groups but they were, overall, slightly more regular and experienced hearing aid users.

TABLE 5. Summary statistics for subscales and scales of the PHAP. Sample sizes vary due to missing data on some items.

Scale	*No. Items*	*Mean*	*SD*	*Coeff* α	*N*
Subscales					
Familiar talkers	7	20.8	14.7	.84	211
Ease of communication	7	37.0	22.0	.86	209
Reverberation	9	53.1	22.3	.88	191
Reduced cues	9	50.9	17.8	.82	192
Background noise	16	54.0	21.3	.94	194
Aversiveness of sounds	12	40.2	23.2	.89	181
Distortion of sounds	6	40.8	23.9	.83	207
Scales					
Environment A	14	28.6	16.3	.89	198
Environment B	18	52.2	19.1	.92	170
Environment C	16	54.0	21.3	.94	194
Environmental sounds	18	40.7	21.5	.92	170

TABLE 6. Intercorrelations among subscales and scales of the PHAP (N = 125).

Subscales	*EC*	*RV*	*RC*	*BN*	*AV*	*DS*	*Scales*	*SB*	*SC*	*ES*
FT	.56	.51	.60	.51	.23	.40	SA	.76	.69	.50
EC		.70	.66	.67	.40	.61	SB		.90	.59
RV			.81	.86	.51	.60	SC			.60
RC				.87	.42	.55				
BN					.49	.64				
AV						.60				

Procedure

Subjects completed the inventory a second time 10–20 days after the first administration. Because subjects were responding on a 7-point scale to a large number of items, it seemed unlikely that they would be able to remember their initial responses. However, as an additional precaution, the items were presented in a different random order on the second occasion.

RESULTS

Test and retest scores and test-retest differences for scales and subscales were computed for each subject. Mean test and retest scores as well as mean differences and standard deviations of the distributions of differences are given in Table 8 (Table 9 gives between-subject standard deviations for each scale and subscale). In addition, Table 8 shows 90% and 95% critical differences for each scale and subscale. These critical differences were computed using the standard deviation of test-retest differences, assuming that these differences are normally distributed. If an obtained difference between scores in two test conditions has only a 10% (or 5%) probability of being observed by chance between scores obtained under identical conditions, it may be concluded with 90% (or 95%) certainty that the two test conditions were not identical. Critical differences may also be computed using the standard error of measurement as described by Demorest and Walden (1984). The two approaches result in essentially identical critical difference values.

TABLE 7. Internal consistency reliability and inter-item correlation data for subscales and scales of the PHAP determined from subject Group 2. $Corr_{i\text{-}t}$ = range of corrected item-total correlations.

Scale	$Corr_{i\text{-}t}$	*Coeff* α	*N*
Subscales			
Familiar talkers	.35–.61	.75	73
Ease of communication	.47–.82	.90	74
Reverberation	.39–.74	.84	74
Reduced cues	.26–.64	.70	70
Background noise	.50–.68	.90	71
Aversiveness of sounds	.43–.77	.90	70
Distortion of sounds	.42–.52	.75	72
Scales			
Environment A	.14–.78	.85	71
Environment B	.27–.67	.88	69
Environment C	.50–.68	.90	71
Environmental sound	.35–.74	.91	67

Table 8 illustrates that the mean test-retest difference was small for all scales and subscales, with a maximum value of 3.1 for subscale EC.[2] Comparison with Table 7 reveals that the scales and subscales with the highest internal consistency, evidenced by coefficient alpha in Table 7, were not necessarily those with the highest retest consistency, evidenced by the standard deviation of test-retest differences in Table 8. For example, subscale FT has a relatively low coefficient alpha of .75 but also is the subscale with the most repeatable scores and, therefore, the smallest critical differences. By contrast, subscale AV was shown to have high internal consistency (.90) but, nevertheless, relatively large critical differences, requiring a change in score exceeding 20% before a significant difference can be inferred.

The advantage of combining subscales into scales can be clearly seen in Table 8. The distribution of test-retest differences is generally narrower in the scales than in most of their constituent subscales. This is presumably due to cancellation of random measurement errors with the larger numbers of items. As a result, critical differences are smaller for scales than subscales. At the same time, the internal consistency of the scales is also quite acceptable, as seen in Table 7.

Although the standard error of the test-retest differences is the essential information for evaluation of scores under different conditions, correlation analyses also provide interesting insights into test-retest patterns. Although some of the subscales have similar mean scores (Table 5), results from individual subjects actually reveal a considerable diversity of profile shapes. This is illustrated in Figure 4, which shows profiles for 2 subjects. Correlations between test and retest profiles indicate the extent to which the shape of a 4- or 7-score profile obtained for the first test was replicated in the second test. In the present study, correlations were computed for each subject between the first and second profiles. Separate correlation coefficients were determined for the 7

[2]It should be kept in mind that a small mean test-retest difference does not necessarily imply that test-retest differences were small for every subject. When a mean difference is computed, positive and negative individual differences cancel.

TABLE 8. Mean test and retest scores and means and standard deviations (*SD*) of test-retest differences for the PHAP. Data are in percent, or frequency, of problems in the scale and subscale content areas. The Table also gives 90% and 95% critical differences (*CD*), in percent, for each scale and subscale. *N* = 30.

Scale	*Test mean*	*Retest mean*	*Diff mean*	*Diff SD*	*90% CD*	*95% CD*
Subscales						
Familiar talkers	17.3	16.5	0.8	8.5	14.0	16.7
Ease of communication	39.9	36.8	3.1	14.1	23.3	27.6
Reverberation	50.1	52.6	−2.5	12.2	20.1	23.9
Reduced cues	52.7	51.9	0.9	10.7	17.6	21.0
Background noise	53.4	52.3	1.1	9.2	15.2	18.0
Aversiveness of sounds	46.6	45.0	1.6	14.2	23.4	27.8
Distortion of sounds	41.3	43.3	−2.0	14.4	23.7	28.2
Scales						
Environment A	28.6	26.7	1.9	8.2	13.5	16.1
Environment B	51.4	52.2	−0.8	10.1	16.7	19.8
Environment C	53.4	52.3	1.1	9.2	15.2	18.0
Environmental sound	44.8	44.4	0.4	11.6	19.1	22.7

subscales and for the 4 scales. The distributions of the obtained correlation coefficients are illustrated in Figure 5. This figure shows that for most of the subjects, the correlation was reasonably high (> .80) for both the 7- and 4-score profiles. Thus, the shape of the profile was usually rather closely replicated on test and retest.

Correlation coefficients also were computed across the 30 subjects between the test and retest scores for each of the scales and subscales. The results are shown in Table 9. In addition, Table 9 gives the between-subject standard deviation of scores for each (sub)scale as well as the standard errors of measurement computed using the correlations and standard deviations given in the table (Nunnally, 1978, p. 218). Subscale correlations ranged from a low of .66 for RC to a high of .88 for BN. These correlation coefficients should be interpreted as indicating the extent to which subjects maintained their relative standing on each (sub)scale from test to retest. It may be tempting to infer that 2 subscales with relatively low test-retest correlations (FT and RC) yield less reliable scores, but this is not necessarily so. Other things being equal, subscales with greater between-subject variance will tend to have higher retest correlations. The data given in Table 9 follow this general rule. Because correlations are influenced by between-subject differences, they do not provide a good measure of score reliability.

The standard error of measurement gives an estimate of score reliability that is essentially independent of between-subject differences. This statistic may be interpreted as the standard deviation of scores subjects would obtain for many parallel (sub)scales. A comparison of the standard errors of measurement for the different subscales suggests that the precision of measurement is rather similar for most of the subscales assessing speech communication. However, the precision is somewhat lower for the two subscales assessing response to amplified environmental sounds.

DISCUSSION

The outcome of the investigations described above was a 66-item inventory that yields a 7- or 4-score profile quantifying hearing aid performance. The paper-and-pencil profile is usually completed in 20–30 min. The

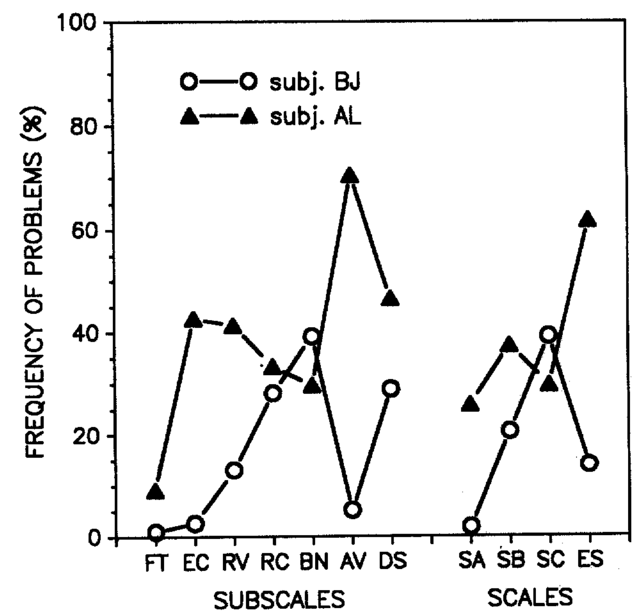

FIGURE 4. Examples of PHAP profiles obtained from 2 hearing aid wearers. For each subject, the profile composed of 7 scores is seen on the left and the 4-score profile is given on the right. FT = Familiar talkers; EC = Ease of communication; RV = Reverberation; RC = Reduced cues; BN = Background noise; AV = Aversiveness of sounds; DS = Distortion of sounds; SA = Speech, Environment A; SB = Speech, Environment B; SC = Speech, Environment C; ES = Environmental sounds.

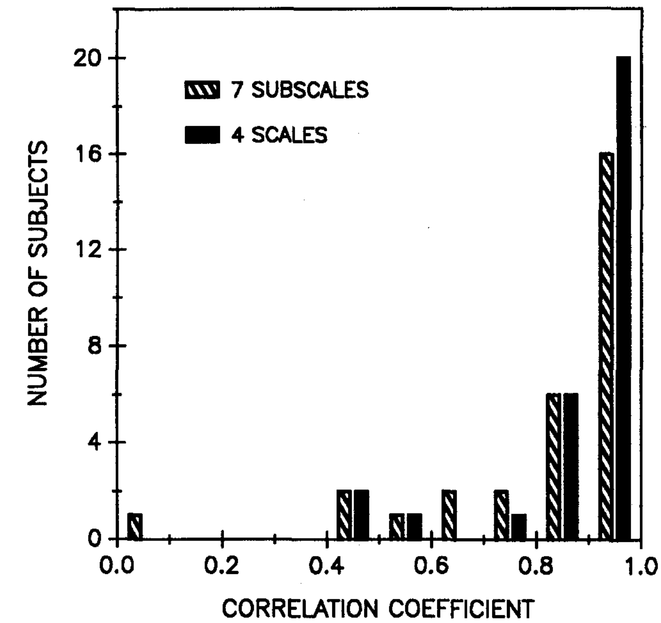

Figure 5. Distribution of correlation coefficients for 7-score and 4-score profiles determined using the PHAP. $N = 30$.

subjects for the studies were mostly elderly, adventitiously impaired individuals with mild to moderate hearing loss. Thus, the PHAP seems well suited for application with older adults having presbyacusic or noise-induced impairments.

The psychometric properties of the PHAP (internal consistency, test-retest reliability, critical differences) appear to be acceptable and are rather similar to corresponding data obtained using other self-assessment tools with hearing-impaired persons (e.g., Demorest & Erdman, 1987, 1988; Ventry & Weinstein, 1982; Weinstein, Spitzer, & Ventry, 1986). Overall, these results suggest that the PHAP can be used to elicit repeatable data about

Table 9. Test-retest correlations (r), between-subject standard deviations in percent (SD), and standard error of measurement in percent (Se) for scales and subscales of the PHAP. $N = 30$.

Scale	*r*	*SD*	*Se*
Subscales			
Familiar talkers	.69	10.7	5.9
Ease of communication	.77	20.4	9.8
Reverberation	.84	21.2	8.5
Reduced cues	.66	13.0	7.6
Background noise	.88	18.5	6.4
Aversiveness of sounds	.84	25.2	10.1
Distortion of sounds	.78	21.3	10.0
Scales			
Environment A	.81	13.2	5.8
Environment B	.81	16.2	7.1
Environment C	.88	18.5	6.4
Environmental sound	.86	21.5	8.0

the perceptions of hearing aid wearers concerning their performance with hearing aids.

Although there is some redundancy in presenting both the 7-subscale and the 4-scale profiles, each appears to have advantages not possessed by the other. The principal advantage of the 4-score profile is the greater reliability of each score, resulting in smaller critical differences. Thus, smaller interscore differences are necessary to detect a real difference between conditions. On the other hand, the 7-score profile has the potential for providing more specific information about problems with the hearing aid fitting. Some examples can be seen in the two profiles shown in Figure 4. Comparison of the SA scores indicates that Subject AL reported much more difficulty communicating in Environment A situations than did Subject BJ. However, examination of the SA subscales (FT and EC) suggests that the difference between the 2 individuals is mostly related to the amount of effort they expend to understand speech. Both report few problems understanding the speech of familiar talkers (FT) but their responses were very different when describing the ease with which they are able to communicate (the EC subscale). The additional insight provided by the subscale scores, in combination with information about the individual's hearing impairment, may point the way to appropriate intervention. For example, the score on the EC subscale might be improved if subject AL increased the hearing aid's gain in settings where background noise is relatively low.

Reference to Figure 4 also indicates that these 2 subjects' scores on the ES scale were very different, with Subject AL reporting more frequent problems with amplified environmental sounds. Again, examination of the subscale scores reveals different patterns in the two profiles: AL reported a high proportion of loudness discomfort due to environmental sounds but only moderate levels of distortion. BJ reported a very low proportion of loudness discomfort but a higher frequency of sound distortion. Possibly, these results indicate that the maximum output of AL's hearing aid is too high, whereas BJ's hearing aid may produce excessive harmonic, intermodulation, or transient distortion when worn at the settings used in daily life. Further study is needed to determine whether these types of interpretations are valid.

Although it is not foreseen that the PHAP would be used routinely in hearing aid selection, the inventory could be useful in evaluating problematic hearing aid fittings. In addition, the PHAP could be used to assess the impact in everyday life of hearing aid fittings that differ, for instance, in signal processing strategies. The hearing impaired individual could use the first fitting for 1–2 weeks and then complete the inventory. The second fitting could then be tried for the same period and the inventory completed again (the time interval used in the test-retest study was chosen with this sort of application in mind). Two fittings that differ in noise reduction strategies, for example, may produce significantly different scores on the BN subscale and perhaps on other subscales also.

In another clinical application, an average profile could be determined for successful hearing aid wearers. This profile could be used as a yardstick to evaluate the success of individual hearing aid fittings. For example, it might be determined that a particular fitting results in typical scores for communication in Environment A but poorer than usual scores for Environments B and C. Such a finding might indicate a need for a change in frequency response or, perhaps, a noise-reduction type hearing aid.

A modified inventory may be used to produce a Profile of Hearing Aid Benefit (PHAB). In this application, the hearing aid wearer responds twice to each item, once to reflect unaided experiences and again to indicate performance while wearing the hearing aid. Benefit is computed by subtracting aided from unaided responses for each item and then determining scale and subscale scores. The psychometric properties of the PHAB are under investigation as well as its relationship to other subjective and objective measures of hearing aid benefit.

As these examples suggest, there are numerous potential applications for this profile of hearing aid performance. Continued study of responses from hearing aid wearers we hope will result in enhanced understanding of the impact of hearing aid features such as electroacoustic performance characteristics and signal processing schemes on the benefit received from amplification and its overall efficacy in everyday life.

Finally, it is important to keep in mind that the PHAP has been developed as an instrument for the study of one aspect of the rehabilitation effort for a hearing-impaired individual—the amplification system. The PHAP does not assess the social and emotional consequences of hearing loss or underlying personality attributes that might influence responses to a hearing handicap. Nevertheless, knowledge of these factors may be important in interpreting profiles obtained with the inventory. The effect, for example, of denial of hearing loss on self-assessed hearing aid benefit has yet to be determined. If the PHAP is employed in a clinical setting, the clinician is encouraged to evaluate social, emotional, and personality issues using one of the available instruments at the outset of the rehabilitation program. This information can establish a framework within which subsequent self-assessments such as those elicited with the PHAP can be interpreted.

ACKNOWLEDGMENTS

This work was supported by VA Rehabilitation Research and Development funds. Software to score the PHAP inventory was written by Robert M. Joyce. Rhonda Huggins and Leslie Gott collected the data for the study of test-retest reliability.

REFERENCES

AMERICAN NATIONAL STANDARD INSTITUTE (1969). *Specification for audiometers* (ANSI S3.6–1969). New York: ANSI.

ANGEVINE, O. L. (1975). Individual differences in the annoyance of noise. *Sound & Vibration, 9*, 40–42.

BARCHAM, L. J., & STEPHENS, S. D. G. (1980). The use of an open-ended problems questionnaire in auditory rehabilitation. *British Journal of Audiology, 14*, 49–51.

BERGER, K. W., ABEL, D., HAGBERG, E., PUZZ, L., VARAVVAS, D., & WELDELE, F. J. (1982). Successes and problems of hearing aid users. *Hearing Aid Journal, 35*(12), 26–30.

BROOKS, D. N. (1985). Factors relating to the under-use of postaural aids. *British Journal of Audiology, 19*, 211–217.

COX, R. M., ALEXANDER, G. C., & GILMORE, C. (1987). Intelligibility of average talkers in typical listening environments. *Journal of the Acoustical Society of America, 81*, 1598–1608.

CHUNG, S. M., & STEVENS, S. D. G. (1986). Factors influencing binaural hearing aid use. *British Journal of Audiology, 20*, 129–140.

DEMOREST, M. E., & ERDMAN, S. A. (1986). Scale composition and item analysis of the Communication Profile for the Hearing Impaired. *Journal of Speech and Hearing Research, 29*, 515–535.

DEMOREST, M. E., & ERDMAN, S. A. (1987). Development of the Communication Profile for the Hearing Impaired. *Journal of Speech and Hearing Disorders, 52*, 129–143.

DEMOREST, M. E., & ERDMAN, S. A. (1988). Retest stability of the Communication Profile for the Hearing Impaired. *Ear and Hearing, 9*, 237–242.

DEMOREST, M. E., & WALDEN, B. E. (1984). Psychometric principles in the selection, interpretation, and evaluation of communication self-assessment inventories. *Journal of Speech and Hearing Disorders, 49*, 226–240.

FRANKS, J. R., & BECKMAN, N. J. (1985). Rejection of hearing aids: Attitudes of a geriatric sample. *Ear and Hearing, 6*, 161–166.

GIOLAS, T. G., OWENS, E., LAMB, S. H., & SCHUBERT, E. D. (1979). Hearing Performance Inventory. *Journal of Speech and Hearing Disorders, 44*, 169–195.

GRAEVEN, D. B. (1975). Necessity, control and predictability of noise annoyance. *Journal of Social Psychology, 95*, 85–90.

HAGERMAN, B., & GABRIELSSON, A. (1984). *Questionnaires on desirable properties of hearing aids.* (Report TA109). Stockholm: Karolinska Institutet, Department of Technical Audiology.

HELFER, K. S. (1989, October). Aging influences on consonant confusions in reverberation and noise. *Asha, 31*, 125.

HIGH, W. S., FAIRBANKS, G., & GLORIG, A. (1964). Scale for self-assessment of hearing handicap. *Journal of Speech and Hearing Disorders, 29*, 215–230.

HUTTON, C. A. (1987). A problem-oriented questionnaire for rehabilitative audiology. *Hearing Journal, 40*(10), 47–52.

NABELEK, A. (1988). Identification of vowels in quiet, noise, and reverberation: Relationships with age and hearing loss. *Journal of the Acoustical Society of America, 84*, 476–484.

NABELEK, A. (1989). Reverberant overlap- and self-masking in consonant identification. *Journal of the Acoustical Society of America, 86*, 1259–1265.

NEWMAN, C. W., & WEINSTEIN, B. E. (1988). The Hearing Handicap Inventory for the Elderly as a measure of hearing aid benefit. *Ear and Hearing, 9*, 81–85.

NORUSIS, M. J. (1988). *SPSS/PC+ advanced statistics V2.0* (Computer program manual). Chicago, IL: SPSS.

NUNNALLY, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

TANNAHILL, J. C. (1979). The Hearing Handicap Scale as a measure of hearing aid benefit. *Journal of Speech and Hearing Disorders, XLIV*, 91–99.

VENTRY, I. M., & WEINSTEIN, B. E. (1982). The Hearing Handicap Inventory for the Elderly: A new tool. *Ear and Hearing, 3*, 128–134.

WALDEN, B. E., DEMOREST, M. E., & HEPLER, E. L. (1984). Self-report approach to assessing benefit derived from amplification. *Journal of Speech and Hearing Research, 27*, 49–56.

Weinstein, B. N. (1978). Individual differences in reactions to noise: A longitudinal study in a college dormitory. *Journal of Applied Psychology, 63,* 458–466.
Weinstein, B. E., Spitzer, J. B., & Ventry, I. M. (1986). Test-retest reliability of the Hearing Handicap Inventory for the Elderly. *Ear and Hearing, 7,* 295–299.

Received July 31, 1989
Accepted December 29, 1989

Requests for reprints should be sent to Robyn M. Cox, Memphis State University, Memphis Speech & Hearing Center, 807 Jefferson Avenue, Memphis, TN 38105.

APPENDIX. PROFILE OF HEARING AID PERFORMANCE

INSTRUCTIONS: Please circle the answer that comes closest to your everyday experience. Notice that each choice includes a percentage. You can use this to help you decide on your answer. For example, if the statement is true about 75% of the time, circle C for that item. If you have not experienced a particular situation, imagine how you would respond in a similar situation.

A Always (99%)
B Almost Always (87%)
C Generally (75%)
D Half-the-time (50%)
E Occasionally (25%)
F Seldom (12%)
G Never (1%)

WHEN I WEAR MY HEARING AID

1. I can understand others in a small group situation if there is no noise. A B C D E F G
2. When I am listening to a speaker who is talking to a large group, and I am seated toward the rear of the room, I must make an effort to listen. A B C D E F G
3. Women's voices sound "shrill." A B C D E F G
4. I find that most people speak too softly. A B C D E F G
5. I have trouble comprehending speech when I am in a busy department store talking with the clerk. A B C D E F G
6. I can understand my family when they speak softly to me. A B C D E F G
7. I can understand a speaker in a small group, even when those around us are speaking softly to each other. A B C D E F G
8. I can understand conversations even when several people are talking. A B C D E F G
9. When the telephone rings, the sound startles me. A B C D E F G
10. I have to ask people to repeat themselves when there is background noise. A B C D E F G
11. When I am in a crowded grocery store, talking with the cashier, I can follow the conversation. A B C D E F G
12. When I am having a conversation, and people are talking quietly nearby, I have to strain to understand the speaker. A B C D E F G
13. If a car horn sounds, it makes me jump. A B C D E F G
14. When I am talking to a group, and someone from the back of the room asks a question, I have to ask someone up front to repeat the question. A B C D E F G
15. The sounds of construction work are uncomfortably loud. A B C D E F G
16. When I am in a crowded reception room waiting to be called, I miss hearing my name. A B C D E F G
17. When I am having a quiet conversation with a companion, I have difficulty understanding. A B C D E F G
18. When I am listening to the news on the car radio, and family members are talking, I have trouble hearing the news. A B C D E F G
19. The sound of screeching tires is uncomfortably loud. A B C D E F G
20. I can understand conversation when I am walking with a friend through a quiet park. A B C D E F G
21. The sound of a fire engine siren close by is so loud that I need to cover my ears. A B C D E F G
22. When I am in conversation with someone across a large empty room (such as an auditorium), I understand the words. A B C D E F G
23. When I am on a busy street, asking someone for directions, I have to ask him to repeat them before I really understand. A B C D E F G
24. The sounds of running water, such as a toilet or shower, are uncomfortably loud. A B C D E F G
25. When a speaker is addressing a small group, and everyone is listening quietly, I have to strain to understand. A B C D E F G
26. I have trouble understanding dialogue in a movie or at the theater. A B C D E F G
27. When I am in a crowd with a friend who doesn't want others to overhear our conversation, I have trouble hearing as well. A B C D E F G
28. When I am at the dinner table with several people, and am trying to have a conversation with one person, understanding speech is difficult. A B C D E F G
29. When I'm in a quiet conversation with my doctor in an examination room, it is hard to follow the conversation. A B C D E F G
30. I have trouble understanding others when an air conditioner or fan is on. A B C D E F G
31. I miss a lot of information when I'm listening to a lecture. A B C D E F G
32. I can understand my family when they talk to me in a normal voice. A B C D E F G
33. I have to ask people to repeat themselves in one-on-one conversation in a quiet room. A B C D E F G
34. I have difficulty hearing a conversation when I'm with one other person at home. A B C D E F G
35. When I am riding in the back seat of a car, and the driver talks to me from the front, I have to strain to understand. A B C D E F G
36. Music sounds distorted to me. A B C D E F G
37. When I'm talking with the teller at the drive-in window of my bank, I understand the speech coming from the loudspeaker. A B C D E F G

38. When I am in a small office, interviewing or answering questions, I have difficulty following the conversation. A B C D E F G
39. When a lecturer is giving instructions, I can easily follow along. A B C D E F G
40. Everyday sounds are too soft for me to hear clearly. A B C D E F G
41. I avoid using certain appliances (blender, vacuum cleaner, etc.) because their loudness is uncomfortable. A B C D E F G
42. When I am in a busy restaurant and the waitress is taking my order, I can comprehend her questions. A B C D E F G
43. I avoid crowds because the noise is uncomfortably loud. A B C D E F G
44. When I'm at a large, noisy party, conversation is very confusing. A B C D E F G
45. When I am in a theater watching a movie or play, and the people around me are whispering and rustling paper wrappers, I can still make out the dialogue. A B C D E F G
46. When I am in a quiet restaurant, I can understand soft conversation. A B C D E F G
47. I can't understand the TV news when the volume is set by a normal-hearing person. A B C D E F G
48. I can understand conversation during a quiet dinner with my family. A B C D E F G
49. When I am listening to the news on my car radio, and the car windows are closed, I understand the words. A B C D E F G
50. The sound quality of music isn't very good. A B C D E F G
51. The ring of a telephone sounds "tinny." A B C D E F G
52. I understand the newscaster when I am watching TV news at home alone. A B C D E F G
53. I can follow the words of a sermon when listening to a religious service. A B C D E F G
54. When I am at home, talking with someone who is in another room, following the conversation is difficult. A B C D E F G
55. Unexpected sounds like a smoke detector or alarm bell are uncomfortable. A B C D E F G
56. When I'm talking with a friend outdoors on a windy day, I miss a lot of the conversation. A B C D E F G
57. Everyday sounds that don't bother others are too loud for me. A B C D E F G
58. It's hard for me to understand what is being said at lectures or church services. A B C D E F G
59. When I am in a room with the door closed and I want to overhear a conversation going on outside the door, I have to strain to listen. A B C D E F G
60. People's voices sound unnatural. A B C D E F G
61. When I am in face-to-face conversation with one member of my family, I can easily follow along. A B C D E F G
62. When I am in a meeting with several other people, I can comprehend speech. A B C D E F G
63. Traffic noises are too loud. A B C D E F G
64. The sound of glass breaking is uncomfortably loud. A B C D E F G
65. I can communicate with others when we are in a crowd. A B C D E F G
66. I understand speech when I am talking to a bank teller, and I am one of a few customers at the bank. A B C D E F G

TABLE A1. Listing of item numbers included in each PHAP subscale. Responses to items whose numbers are marked with an asterisk are reversed before scoring.

Subscale	*Item numbers*
Familiar talkers (FT)	1*, 20*, 32*, 48*, 52*, 61*, 66*.
Ease of communication (EC)	12, 17, 25, 29, 33, 34, 38.
Reverberation (RV)	2, 14, 22*, 26, 31, 39*, 45*, 53*, 58.
Reduced cues (RC)	4, 6*, 35, 37*, 46*, 47, 49*, 54, 59.
Background noise (BN)	5, 7*, 8*, 10, 11*, 16, 18, 23, 27, 28, 30, 42*, 44, 56, 62*, 65*.
Aversiveness of sounds (AV)	9, 13, 15, 19, 21, 24, 41, 43, 55, 57, 63.
Distortion of sounds (DS)	3, 36, 40, 50, 51, 60.