Audiometric Correlates of the Unaided APHAB

Robyn M. Cox*†
Genevieve C. Alexander*†
Ginger A. Gray†

Abstract

The Abbreviated Profile of Hearing Aid Benefit (APHAB) is a self-report questionnaire that is used to quantify the impact of a hearing problem on an individual's daily life. In this investigation, the relationships were explored between typical clinical audiometric data and the four subscale scores of the APHAB administered in the unaided (without-amplification) condition. Sixty subjects provided APHAB scores, audiograms, and speech recognition data. Analyses revealed significant relationships between audiometric data and each of the three APHAB subscales that reflect speech communication (EC, RV, and BN). None of these subscales was significantly more strongly related to any specific audiological variable. However, the pattern of associations between audiometric variables and subscale scores was consistent with predictions based on item content for subscales EC and RV, but not for BN. As predicted, no relationship was found between audiometric data and scores for the Aversiveness subscale (AV). Even for the subscales with the strongest associations, differences in audiometric data could be used to explain half or less of the variance in self-report data.

Key Words: Hearing loss, speech intelligibility, subjective outcomes

Abbreviations: RSIN = revised speech in noise; RAU = rationalized arcsine unit; SBR = signal-to-babble ratio; SBR-50 = the SBR corresponding to a score of 50% correct; SIN = speech in noise; PTA = pure-tone average; APHAB = Abbreviated Profile of Hearing Aid Benefit; WRS = word recognition score; SRT = speech reception threshold; EC = Ease of Communication; RV = Reverberation; BN = Background Noise; AV = Aversiveness

Sumario:

El Perfil Abreviado de Beneficio del Auxiliar Auditivo (APHAB) es un cuestionario de auto-reporte que se utiliza para cuantificar el impacto de un problema auditivo en la vida diaria de un individuo. En esta investigación, se exploraron las relaciones entre los datos clínicos audiométricos típicos y los puntajes de la cuatro sub-escalas del APHAB, administrado en condiciones no amplificadas. Sesenta sujetos aportaron puntajes del APHAB, audiogramas y datos de reconocimiento de lenguaje. Los análisis revelaron relaciones significativas entre los datos audiométricos y cada una de las tres sub-escalas del APHAB diseñadas para reflejar la comunicación por lenguaje (EC, RV y BN). Ninguna de estas sub-escalas se relacionó fuerte y significativamente con ninguna variable audiológica específica. Sin embargo, el patrón de asociación entre las variables audiométricas y los puntajes de las sub-escalas fue consistente con las predicciones basadas en el contenido de las sub-escalas EC y RV, pero no para BN. Como se predijo, no se encontró relación entre los datos audiométricos y el puntaje de la escala de Aversión (AV). Aún para las escalas con la asociación más fuerte, las diferencias en los datos audiométricos pueden ser utilizadas para explicar la mitad o menos en la variancia de la información de auto-reporte.

*Department of Veterans Affairs Medical Center, Memphis, TN; † University of Memphis, Memphis, TN

Reprint requests: Robyn M. Cox, Ph.D., Memphis Speech and Hearing Center, 807 Jefferson Ave., Memphis, TN 38105, USA; Phone: 901-678-5831; Fax: 901-525-1282; E-mail: robyncox@memphis.edu

Palabras Clave: Pérdida auditiva, inteligibilidad del lenguaje, resultado subjetivo.

Abreviaturas: RSIN = lenguaje en ruido revisado; RAU = unidad racionalizada de arcoseno; SBR = tasa de señal/bable; SBR-50 = el SBR correspondiente a un puntaje correcto del 50%; SIN = lenguaje en ruido; PTA = promedio tonal puro; APHAB = Perfil Abreviado de Beneficio del Auxiliar Auditivo; WRS = Puntaje de reconocimiento de palabras; SRT = umbral de recepción del lenguaje; EC = Facilidad de comunicación; RV = reverberación; BN = ruido de fondo; AV = aversión

There is an ongoing debate concerning the role in hearing health care of self-report data about daily hearing problems. At one time, it was widely assumed that responses to questions about real-life hearing difficulties would be closely related to clinical data such as sensitivity thresholds and speech recognition scores. To evaluate this assumption, investigators have typically explored the relationship between clinical data and data from questionnaires. The literature contains several articles having titles similar to the title of this paper but encompassing different self-report instruments (e.g., Weinstein and Ventry, 1983; Brainerd and Frankel, 1985; Hawes and Niswander, 1985). A review of these articles reveals a fairly consistent pattern in which correlations between traditional audiometric measures and questionnaire data have been found to be low to moderate in strength. However, a close examination of these studies suggests that different questionnaires might tap different aspects of auditory performance. For example, the highest correlation observed by Brainerd and Frankel (1985) between the Social Hearing Handicap Index and audiogram pure-tone average (PTA) was only 0.35. Weinstein and Ventry (1983) studied a different questionnaire (the Hearing Handicap Inventory for the Elderly [HHIE]) and noted a higher correlation of 0.61 between HHIE score and PTA but only -0.42 between the HHIE score and word recognition score. On the other hand, Hawes and Niswander (1985), who studied the Revised Hearing Performance Inventory, found the highest correlation to be with word recognition score (-0.67), and a lower correlation (0.54) with PTA.

These and similar studies support two assertions. First, it appears that individual differences in audiometric data generally can account for half or less of individual differences in real-life hearing problems. The corollary is that an individual's responses to a self-report questionnaire of daily-life problems are usually a relatively poor predictor of the audiometric data for that same individual. Second, it appears that the relationship between audiometric data and self-report data is somewhat different for different questionnaires. To refute the second assertion, it could be argued that the differences across studies such as those reviewed above were possibly the result of sampling effects and might not reflect underlying differences in the questionnaires. On the other hand, it has been demonstrated that the extent to which any particular set of questionnaire items elicit responses that are predictive of, say, the clinical audiogram, depends strongly on the actual wording of the items. This was shown, for example, by Coren and Hakstian (1992) and Koike et al. (1994) in studies that attempted to develop questionnaires for the express purpose of estimating audiogram thresholds. The Hearing Screening Inventory developed by Coren and Hakstian produced a relatively high correlation of 0.81 with PTA.

Although it is clear that clinical audiometric data and self-report data about daily hearing problems are not equivalent, there are occasions when audiometric data are available but self-report data are not, and vice versa. For these situations, it is sometimes important to know whether, or to what extent, audiometric findings might be serviceable predictors of problems in daily life (e.g., Dobie and Sakai, 2001), or whether self-report data could provide any useful insights about likely audiometric findings (e.g., Kochkin, 1998). The answer to these questions depends on the statistical relationship between self-report data and audiometric data. The considerations outlined above suggest that the precise relationship between any self-report measure of hearing problems and traditional audiometric variables is determined in part by the specific content of the questionnaire items and must, therefore, be empirically established. Accordingly, this article explores the relationship between audiometric data and scores for the Abbreviated Profile of Hearing Aid Benefit (APHAB [Cox and Alexander, 1995]). Despite the widespread use of the APHAB, this relationship has not been reported in the literature.

The ABHAB questionnaire is used to quantify everyday life problems associated with hearing impairment. It can be completed by the hearing impaired person for two conditions: first, to describe the frequency of problems when amplification is not used (the unaided condition) and, second, to describe the frequency of the same problems when amplification is used (the aided condition). In addition, benefit from the hearing aid can be calculated by comparing responses for the aided and unaided conditions to determine the extent to which problems are reduced (or increased) when the hearing aid is used. For this investigation, only data for the unaided condition were explored.

The questionnaire comprises 24 items that are scored in four subscales. Each item contributes to only one subscale, and there are six items for each subscale, distributed randomly within the inventory. The subscales are:

- Ease of Communication (EC): The strain of communicating under relatively favorable conditions.
- Background Noise (BN): Communication in settings with high background noise levels.
- Reverberation (RV): Communication in reverberant rooms such as classrooms.
- Aversiveness (AV): The unpleasantness of environmental sounds.

The study examined the relationship between unaided APHAB data, audiogram data, and speech recognition scores for a group of hearing impaired subjects. Because the four APHAB subscales comprise items that are intended to characterize different circumstances of everyday listening, hypotheses were generated for each subscale, as follows:

1. EC score will be most closely related to measures describing midfrequency sensitivity and/or objective clinical tests of speech understanding in quiet conditions. This hypothesis reflects the content of the EC items, which concern speech communication in low-noise situations. Success in these situations should be determined mostly by midfrequency audibility (Siegenthaler and Strand, 1964).

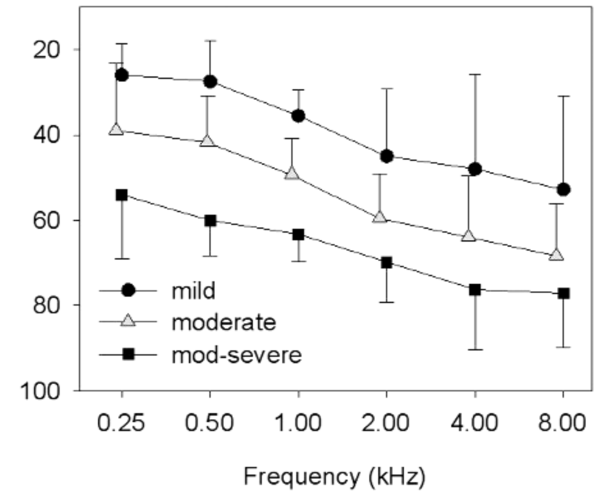

Figure 1 Test ear composite audiograms for the three groups of subjects. Bars show one standard deviation.

2. Scores for BN will be most closely related to measures of high-frequency sensitivity and to objective clinical tests of speech understanding in noise. This hypothesis reflects the fact that BN items describe speech communication in listening situations with substantial ambient noise.
3. Scores for RV will be most closely related to objective measures of speech understanding in noise and to high-frequency sensitivity. Although the RV items concern communication when speech is masked by reverberation (reflections and temporal smearing) rather than ambient noise, several studies have suggested that the effects of reverberation masking are generally similar to those of noise masking (e.g., Nabelek and Mason, 1981). Based on these findings, the hypothesis for the RV subscale was parallel to that for the BN subscale.
4. AV scores will not be related to traditional audiometric data. This hypothesis was formed because the AV items concern reactions to sounds that can be at or near the level of discomfort, whereas traditional audiometric data describe perceptibility of sounds that are well below discomfort levels.

METHODS

Subjects

Sixty individuals with bilaterally symmetric sensorineural hearing impairment were recruited for the study. Inclusion criteria were adult-onset hearing loss, normal language ability, English as first language, no history of recurrent otological problems, and normal tympanogram. In addition, subjects were separated into three groups of 20 based on better ear speech reception threshold (SRT). The groups were labeled as mild (26–40 dB HL), moderate (41–55 dB HL), and moderate-severe (56–70 dB HL). The ear with the better SRT was chosen as the test ear for speech recognition testing. There were 20 left and 40 right test ears. Figure 1 depicts the test ear composite audiograms for the three groups.

Hearing aid use was not a consideration in subject selection. Forty-five subjects reported using amplification at least some of the time. All of the hearing aid wearers indicated that they were aware of the problems they experienced without amplification, and were able to report them (necessary for the unaided APHAB). Twenty-six subjects were women; 34 were men. Ages ranged from 26 to 84. For the mild, moderate, and moderate-severe impairment groups, the mean ages were 68, 74, and 74 years, respectively.

Procedure

Each subject completed the test protocol in the same sequence. The unaided APHAB was completed first using a paper and pencil procedure. The subject was instructed to describe problems experienced when no amplification was used. Second, the standard audiogram was measured, and these data were used to verify group assignment and to select the test ear for speech recognition tests. Third, speech recognition tests were administered.

The APHAB Questionnaire

Each item of the APHAB is a statement, such as, "I can understand my family at the dinner table." The subject must decide how often the statement is true in his/her daily life and respond by choosing from a list of seven alternatives ranging from "Never (1%)" to "Always (99%)." Each response alternative's descriptive word is associated with a percentage of occasions, to help the subject choose the best response. The administration and interpretation of the APHAB is described in detail elsewhere (Cox, 1997).

Speech Recognition Tests

Three speech recognition tests were administered. They included the standard speech reception threshold for spondee words (SRT), a 50-item monosyllabic word list (Auditec recording of the NU-6 Test), and one modified dual block (see below) of the Revised Speech in Noise (RSIN) test (Cox et al., 2001). The monosyllabic word list was delivered without noise competition at a level of 40 dB above the SRT, or lower if required for comfort.

The RSIN test is a revised version of the Speech in Noise (SIN) test (Killion and Villchur, 1993). The SIN test sentences are spoken by a female talker in the presence of a four-talker speech babble. The original SIN test has nine test lists (blocks), with each block composed of 40 sentences. The Revised SIN test was developed to enhance the utility of the test in research settings by increasing the equivalence and reliability of the test blocks. The revision was accomplished by reallocating the prerecorded SIN test material on compact disc into different blocks. The RSIN has four test blocks, with each block composed of 80 sentences.

The SIN test was designed to explore lifelike listening conditions. Thus, it incorporates loud, soft, and noisy speech. These appealing features were retained in the revised version and used in this study. Half of the sentences were presented at 70 dB HL (the loud speech condition). The other half of the sentences were presented at 40 dB HL (the soft speech condition). There were 40 loud sentences and 40 soft sentences. The 40 sentences designated for each presentation level were further divided into four signal-to-babble ratios (SBR): 0 dB, +5 dB, +10 dB, and +15 dB. Thus, for both loud and soft conditions, 10 sentences were administered at each SBR. Each time the SBR was changed, five practice sentences were presented before the test sentences.

Instrumentation and Materials

The audiometric tests were performed in a double-walled sound-treated booth. Signals were routed via a Madsen Orbiter 922 audiometer, with standard calibration (ANSI, 1996), and delivered using an ER3A insert earphone with compressible foam eartip. All speech recognition tests used commercially available stimuli prerecorded on compact discs (Auditec of St. Louis) and replayed using a Denon CD player (model DCD-1290).

RESULTS

For each subject, the responses to the APHAB questionnaire were scored using standard procedures (Cox, 1997) to determine the frequency of problems in daily life in each of the four subscales described above. Figure 2 depicts the mean subscale scores for each of the three impairment groups.

Mean NU-6 word list percent correct word recognition scores (WRS) for the three impairment groups were mild = 93%, moderate = 83%, and moderate-severe = 71%. For statistical treatment, the NU-6 percent correct scores were transformed into rationalized arcsine units (RAU) to homogenize the variance (Studebaker, 1985).

Correct recognition scores for the RSIN scoring words were computed in percent (each test sentence contains five scoring words). Thus, a percent correct score was obtained for each SBR condition for both loud and soft sentences. It was observed that several subjects yielded scores of 0% even for loud speech presented with a 15 dB SBR. Very low scores were especially prevalent for the soft speech condition. Figure 3 illustrates the mean RSIN percentage scores in each listening condition for each hearing impairment group. The left panel of Figure 3 gives data for loud speech, and the right panel gives data for soft speech. All three impairment groups provided reasonable data for loud speech, but only the mild impairment group was consistently able to provide non-zero data for soft speech. This result was undoubtedly due to the limited audibility of the 40 dB HL test sentences for subjects with moderate and worse hearing loss. Because so many subjects were not able to respond to many of the soft RSIN sentences, the soft speech data were not used in subsequent statistical analyses. For statistical analyses, the RSIN percentage scores were transformed into rationalized arcsine units. As described by Cox et al. (2001), it is possible to apply weights to RSIN scores to maximize equivalence across blocks. However, the weights were not used in this study.

Relationship between APHAB Scores and Each Audiological Variable

The first step toward exploring the statistical relationship between audiometric data and APHAB scores was to compute several three-frequency averages from the test ear audiogram. These are described in Table 1. Linear correlations were then computed between each APHAB score and the eight audiometric variables. All 60 subjects

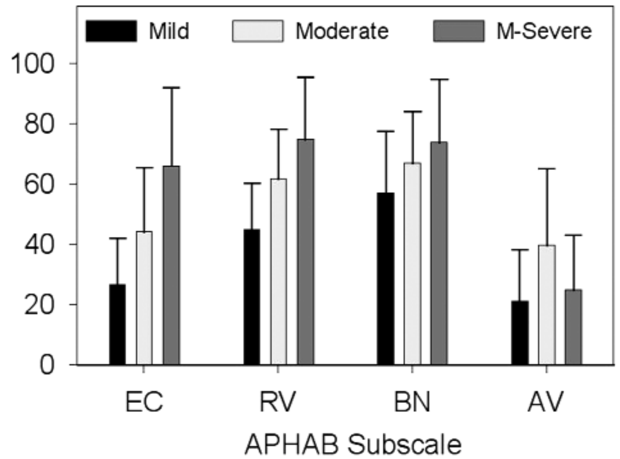

Figure 2 Mean unaided APHAB subscale scores for each of the three subject groups. Bars show one standard deviation.

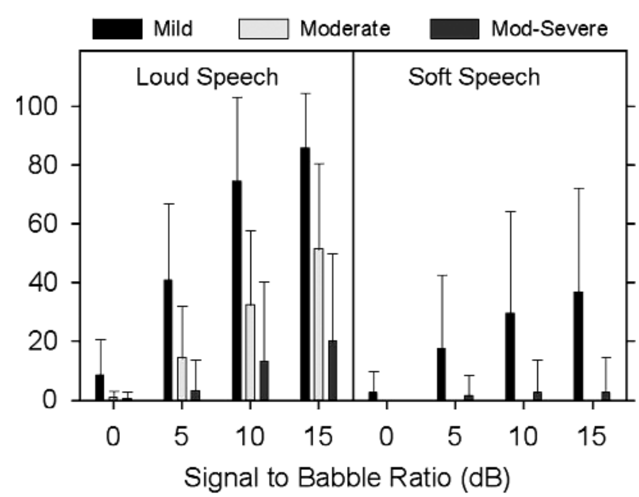

Figure 3 Mean RSIN percentage scores in each listening condition for each hearing impairment group. The left panel gives data for loud speech, and the right panel gives data for soft speech. Bars show one standard deviation.

were used for these analyses. The resulting correlation coefficients are given in Table 2. For the three APHAB subscales that reflect difficulty in everyday speech communication situations (EC, RV, BN), there was a consistent pattern of moderate relationships, with correlation coefficients ranging from .45 to .68. These are all statistically significant ($p < .001$). In contrast, there were no significant relationships between audiometric variables and the scores for the AV subscale.

Table 1 Three-Frequency Averages of Pure-Tone Thresholds from the Better (Test) Ear That Were Used in Statistical Analyses

Variable Name	Average Frequencies (Hz)
PTA1	250, 500, 1000
PTA2	500, 1000, 2000
PTA3	1000, 2000, 4000

Based on item content, it was hypothesized that EC scores would be most closely related to measures describing midfrequency sensitivity and/or speech understanding in quiet conditions or favorable signal-to-noise ratios. The correlations of audiometric variables with EC scores ranged from 0.49 to 0.63. These correlation coefficients are not significantly different from each other ($p > .05$), so we cannot conclude that any of the variables is more strongly related to EC scores. Nevertheless, the trend in the results is consistent with the hypothesis in that the strongest relationships are seen between EC scores and PTA2 (midfrequency sensitivity), WRS (word recognition in quiet), and RSIN15dB (sentence recognition in the most favorable signal-to-noise ratio).

It was hypothesized that scores for the BN subscale would be most closely related to measures describing speech understanding in noise, and to high-frequency sensitivity. Thus, we would predict the highest correlations to occur between BN and PTA3, RSIN10dB, RSIN5dB, and RSIN0dB. This hypothesis was not supported. The correlations between BN scores and audiological variables ranged from 0.45 to 0.51 and were not significantly different from each other. The trend was for the strongest relationships to be with the speech-in-low-noise variables (WRS and RSIN15dB).

The hypothesized relationships between audiological variables and RV scores were the same as for subscale BN, but the results appear to be somewhat different. First, every audiological variable is more strongly correlated with RV scores than with BN scores. This observation is compelling even though the differences are not statistically significant for any individual audiological variable. Second, as predicted, the strongest relationship is seen between RV scores and sentence

Table 2 Linear Correlations between APHAB Scores and Audiometric Variables

APHAB Subscale	PTA1	PTA2	PTA3	WRS	RSIN 15 dB	RSIN 10 dB	RSIN 5 dB	RSIN 0 dB
EC	.61	.63	.61	-.62	-.62	-.58	-.58	-.49
RV	.58	.60	.56	-.61	-.66	-.66	-.68	-.58
BN	.45	.49	.45	-.51	-.50	-.47	-.49	-.45
AV	-.02	-.00	-.05	-.02	.06	.06	.05	.11

Note: RSIN variables are scores for loud sentences (70 dB HL) only. N=60.

Table 3 Summary of Stepwise Multiple Regression Analyses for Each APHAB Scores

APHAB Score	Variables Include	Variance Accounted For	Multiple Regression Coefficient
EC	PTA2 +RSIN15	40% +4.5%	.67
RV	RSIN5 +PTA1	46% +5%	.71
BN	WRS	26%	.51
AV	–	0%	–

Note: N=60.

recognition in moderate noise (RSIN5dB). Note, however, that the correlations between RV scores and audiological variables (ranging from 0.56 to 0.68) were not significantly different from each other.

As predicted, the scores for the AV subscale were not significantly related to any audiometric variable.

Relationship between APHAB Scores and Combinations of Audiological Variables

It was postulated that real-life communication problems reflected by the APHAB scores might be predicted more accurately by combined sensitivity and speech recognition variables than by any one variable alone. To explore this matter, a stepwise multiple regression analysis was performed for each of the four APHAB subscales. To maximize the subjects-to-variables ratio, some of the audiometric variables in Table 2 were not included in the multiple regression analyses. For each analysis, the APHAB score was the dependent variable. The independent variables were (1) the single PTA or RSIN variable with the strongest relationship to the APHAB score, (2) all the variables from either the PTA category or the RSIN category, whichever was not represented in the first variable, and (3) the WRS. The criterion for inclusion in the model was $p < .05$. For removal from the model, the criterion was $p > .10$.

The results of the four multiple regression analyses are summarized in Table 3 and support the following conclusions:

- For EC and RV subscale scores, 4–5% of additional variance was accounted for by combining a second audiological variable with the one most strongly related to the subscale. For these subscales, the combined audiological variables accounted for 44.5% and 51%, respectively, of the variance in APHAB scores.
- In contrast, for BN subscale scores, combining audiometric variables did not improve their ability to predict APHAB scores. Monosyllabic word recognition in quiet was the audiometric variable most closely related to BN score, and differences in WRS accounted for 26% of the differences in BN scores.
- Scores for the Aversiveness subscale (AV) were not significantly associated with any combination of audiometric variables.

Relationship between APHAB Scores and SBR-50 Scores

One further type of analysis was performed using the RSIN data. In addition to yielding percent correct scores for each SBR condition, the set of RSIN scores at a given level for each subject can, in principle, be used to estimate the SBR that would be needed to produce a score of 50% correct (SBR-50) for that subject. It has been suggested that the SBR-50 score might provide a more accurate objective prediction of the severity of communication problems in daily life (Killion and Niquette, 2000). To compute SBR-50, the scores for the four SBR conditions are plotted; the data points that bracket the 50% score are connected by a straight line; and this line is used to determine the SBR that would correspond to a score of 50%. In the current study, we computed SBR-50 scores for the loud speech data. Using the above method, scores were obtainable for only 36 subjects because the rest did not have any scores above 50%. For a further five subjects, a reasonable estimation of SBR-50 could be made by minimal extrapolation of the RSIN data. For these 41 (36 + 5) subjects, SBR-50 scores ranged from 0.5 dB to 17.5 dB.

Of the remaining 19 subjects, nine had

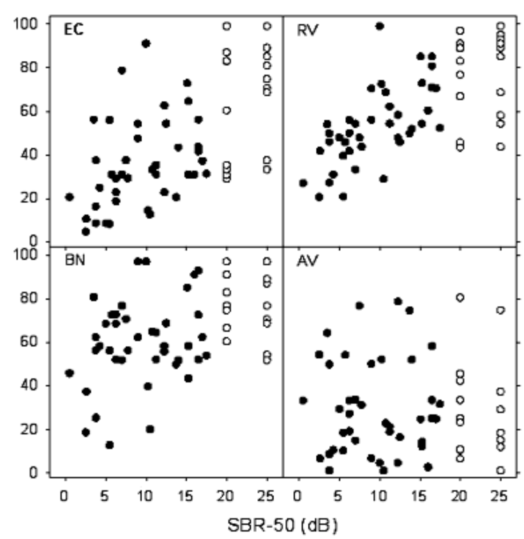

Figure 4 Relationship between SBR-50 scores and APHAB scores for each APHAB subscale. The filled circles show data for subjects whose RSIN data yielded valid SBR-50 scores. The open circles (at SBR-50 values of 20 and 25) show the APHAB scores for subjects whose RSIN data were too poor to yield an SBR-50 score.

low but non-zero scores on at least the easier conditions of the RSIN test, whereas ten had scores of 0% correct on every SBR condition. Even though these subjects did not give valid SBR-50 scores, it would be misleading to exclude them entirely from consideration of the relationship between SBR-50 scores and APHAB scores, because all of these subjects are known to have SBR-50 scores poorer than 15 dB. Thus, for illustrative purposes in Figure 4, these individuals were assigned SBR-50 scores of 20 and 25 dB, respectively.

Figure 4 uses scatter plots to illustrate the relationship between SBR-50 scores and APHAB scores for each APHAB subscale. The filled circles show the data for subjects whose RSIN data yielded valid SBR-50 scores. The open circles (at SBR-50 values of 20 and 25) show the APHAB scores for subjects whose RSIN data were too poor to yield an SBR-50 score. These data support the following observations:

- The closest correspondence between SBR-50 and APHAB scores was seen for the RV subscale. The correlation coefficient between RV scores and SBR-50 scores was significant and moderately strong ($r = .64$, $p < .01$). Further, all the subjects whose RSIN scores were too poor to yield an SBR-50 score also had APHAB scores indicative of a relatively high frequency of problems (more than 40).
- The relationship between EC scores and SBR-50 scores was statistically significant but quite weak ($r = .37$, $p = .02$). Furthermore, six subjects who had quite low EC scores (less than 40), indicative of few real-life problems, performed so poorly on the RSIN test that an SBR-50 score could not be determined.
- For both the BN and AV subscales, there was no significant relationship between SBR-50 and APHAB scores.

DISCUSSION

By design, the subjects selected for this investigation represented a wide range of impairment levels. Because relationships between variables are most easily revealed when each variable contains data that cover a wide range of its potential values, this design enhanced our ability to delineate the audiometric correlates of APHAB scores. A potential drawback of this design was that some of the subjects were hearing aid users, and it can be argued that their ability to accurately report the extent of unaided problems could be compromised by their limited experience with unaided listening. In the current study, this issue was addressed (to the extent possible) by ascertaining, in advance of data collection, that each subject believed him-/herself able to report unaided problems.

When the data are viewed with the subjects grouped by impairment level, as in Figures 2 and 3, the patterns of results are exactly as we would predict. On average, individuals with more pure-tone threshold impairment had poorer measured speech understanding (Figure 3) and more frequent self-assessed problems in daily-life situations (Figure 2).

To illustrate, consider the results for the EC subscale in Figure 2. The mild impairment group reported a lower frequency of problems than the moderate impairment group that, in turn, reported a lower frequency of problems than the moderate-severe impairment group. The same systematic pattern is seen in the other two speech communication subscales, RV and BN. Note, however, that the pattern is not carried out for the AV subscale because both the mild and the moderate-severe impairment groups reported fewer problems with aversiveness of environmental sounds than did the moderate impairment group. This outcome supports the observation that aversiveness of environmental sounds (suprathreshold stimuli) does not have a simple linear relationship to extent of hearing threshold impairment.

Figure 3 (left panel, loud speech) also reveals predictable results. Within each SBR condition, subjects in the mild impairment group yielded higher RSIN intelligibility scores than did the subjects in the moderate impairment group. The moderate group, in turn, produced better intelligibility scores than the moderate-severe impairment group. In addition, within each impairment level, mean score improved as SBR improved in four conditions from 0 dB to 15 dB. The right panel of Figure 3 depicts the results for soft speech. Although measurable (and predictable) results were realized for the mild impairment group, the soft speech level was essentially inaudible to the moderate and moderate-severe impairment groups.

We had noted that, in previous studies,

the pattern of audiometric correlates was somewhat different for different questionnaires with different item content. For example, Kramer et al. (1996) found that audiometric measures of localization ability and interaural threshold asymmetry were significantly related to self-reports of localization problems but not to self-reports of ability to understand speech in quiet. Based on these kinds of observations, we postulated that the specific audiometric correlates of any questionnaire would be determined to some extent by item content. Accordingly, separate hypotheses were generated about the relationship between audiological variables and scores for each of the four APHAB subscales.

For subscales EC and RV, the correlations between individual audiometric variables and self-report APHAB data (Table 2) ranged from 0.49 to 0.68. These values were similar to those reported in several previous investigations using other questionnaires (e.g., McCartney et al., 1976; Weinstein and Ventry, 1983; Hawes and Niswander, 1985; Lutman et al., 1987; Kramer et al., 1996). The trends in the data supported the specific hypotheses for EC and RV scores, but the effects were not statistically significant. Thus, this study indicated that EC and RV scores are moderately related to both threshold sensitivity and speech understanding in quiet and noise.

The association between audiometric variables and scores for the BN subscale ranged from 0.45 to 0.51 and were notably weaker than the corresponding associations for EC and RV scores. Further, contrary to prediction, the scores for subscale BN, which quantifies communication in background noise, were not more strongly related to objective speech recognition in noise. In fact, BN scores were associated with threshold sensitivity as strongly as they were associated with speech recognition. Thus, the trends in the data did not support the hypothesis for BN scores.

In summary, all of the audiological variables were moderately strongly related to EC, BN, and RV scores. For a given speech communication subscale, no single audiological variable stood out as significantly more associated with the subscale score; thus, the specific hypotheses for each subscale were not supported. Further, the relationships between audiological variables and APHAB score was weakest for the BN subscale.

The results of multiple regression analyses summarized in Table 3 indicated that combining threshold sensitivity and speech understanding variables provided a small advantage over either one alone in explaining the variance in EC and RV scores. The multiple regression coefficients for EC and RV were 0.67 and 0.71, respectively. In other words, the combination of threshold sensitivity and speech recognition variables was able to account for close to half of the variance in the self-report scores.

At times it is appealing to try to predict an individual's problems in daily life based on audiometric variables such as those used in this study. The results of this study illustrate that such predictions should be undertaken with great caution because the potential inaccuracy is high, even for EC and RV scores. Despite the moderately strong aggregate relationships observed between these two subscales and audiological data, the impact of a hearing impairment on a particular individual's daily life cannot accurately be predicted from pure-tone thresholds and speech recognition scores. This is illustrated in Figure 5. Each data point depicts the relationship between predicted and actual EC or RV scores for a given individual. The predicted scores were determined using the multiple regression model derived for that subscale. If the predicted scores were substantially correct, all the data points would fall near the solid diagonal line. Clearly, they do not do so. The differences between predicted and actual

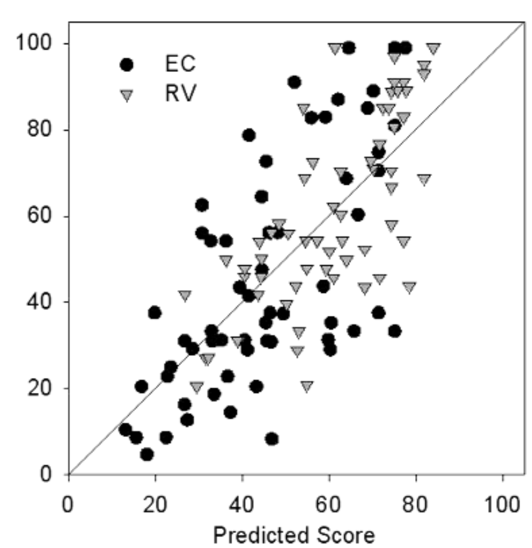

Figure 5 Relationship between predicted and actual EC or RV scores for each individual subject.

scores often exceed 20 points for an individual person, despite the significant and substantial similarities when the group is considered as a whole. We should not be surprised to observe this outcome. Experience studying and using self-reports of hearing difficulty has shown that self-report data includes influences from many variables in addition to impairment, whereas only impairment is measured by clinical audiometric data.

It was not expected that the strongest audiological correlate with BN scores (which reflect real-life communication in noisy situations) would be monosyllabic word recognition in quiet (WRS). In addition, even this correlation was relatively weak, explaining only 26% of the variance in self-report data. Further, the multiple regression approach (Table 3) did not improve the predictability of results for subscale BN. Overall, the results for BN scores suggest that this subscale is somewhat different from EC and RV in terms of its association with the audiological variables quantified in this study.

In an attempt to understand why BN scores demonstrated lower associations with audiometric variables, responses to the 24 individual APHAB items were examined, and BN items were compared with EC and RV items. Generally, the responses to BN items were qualitatively similar in pattern to those for the EC and RV items, but more variable (hence, the lower correlations). Every item in all three subscales was significantly associated with the three PTA variables and with the WRS variable; however, there were some differences in the pattern of associations with the RSIN variables. Every item in subscales EC and RV was significantly associated with all four RSIN variables, but this was not the case for the BN items. One BN item (form A, #6: listening to the car radio) was not associated with any of the sentence recognition (RSIN) variables. Two items (form A, #16 and #18: conversing when several people are talking, and "in a crowd") were significantly associated with all four RSIN variables. The three other BN items (air conditioner, grocery store, and dinner table) were variously related to the RSIN variables. These results indicate that, although all of the items in the BN subscale explore common daily-life situations where communication is limited by ambient noise, some of the items appeal to abilities or experiences that are not invoked in the RSIN test. It would be interesting to determine whether this result is observed for other objective clinical tests of speech understanding in noise.

As predicted, we observed that scores on the APHAB aversiveness subscale were not related to any of the audiometric variables examined. This should not be surprising because item content for the AV subscale relates to the unpleasantness or discomfort associated with loud environmental sounds, whereas the audiometric variables encompassed understanding speech and detection of soft sounds. This finding is consistent with previous research in which no relationship was found between AV scores and speech reception thresholds (Cox et al. 1999). On the other hand, those authors did observe a significant relationship between AV scores and a personality attribute (external locus of control). Further, ongoing research (unpublished) has also shown significant relationships between AV scores and personality attributes such as neuroticism and use of avoidance coping mechanisms. These types of results suggest that AV scores are more associated with psychological than psychoacoustic variables. In the future, it would be interesting to explore the relationship between AV scores and audiometric measures that delve into suprathreshold loudness perceptions.

It has been proposed that a measure of the signal-to-babble ratio that results in a score of 50 percent (SBR-50) would be an improved predictor of the real-life impact of a hearing impairment (Killion and Villchur, 1993). That proposal was not supported by the results of this study. Comparison of the results displayed in Figure 4 and Table 2 reveals that the SBR-50 was not superior to other speech intelligibility variables in terms of its relationship to APHAB scores. In addition, there were many subjects for whom SBR-50 scores could not be obtained using the RSIN stimulus materials.

CONCLUSION

Like many other self-report inventories, the APHAB questionnaire generates several scores (subscales EC, RV, and BN) that are moderately correlated to pure-tone thresholds and monosyllabic word recognition scores. In addition, slightly improved prediction of scores in the EC and RV subscales can be obtained using a combination of threshold variables and speech recognition in noise variables. However, no noteworthy relationships were observed between audiological variables and scores on the APHAB Aversiveness subscale.

Results for the BN subscale revealed less overlap with audiological data than did those for the EC and RV subscales. Audiometric variables accounted for about half of the variance in EC and RV scores but only one-quarter of the variance in BN scores. This intriguing finding calls for continued scrutiny of the core attributes and underlying correlates of the BN items.

Finally, it is important to keep in mind that even the strongest relationships observed in this study do not permit accurate prediction of APHAB subscale scores for individual persons, based on their audiological data. It is clear that self-reports of daily-life hearing problems are only partly rooted in physiological impairment and in the types of psychoacoustic abilities encompassed in traditional audiometric examinations. Self-reports provide unique insights into the consequences of hearing loss that are not obtainable with conventional objective clinical assessments.

Acknowledgments. This article is based on work supported by the Office of Research and Development, Rehabilitation R&D Service, Department of Veterans Affairs (VA). Amy Koralewski and Anne Murphy helped with data collection.

REFERENCES

ANSI. (1996). *American National Standard Specification for Audiometers* (ANSI S3.6-1996). New York: ANSI.

Brainerd SH, Frankel BG. (1985). The relationship between audiometric and self-report measures of hearing handicap. *Ear Hear* 6:89–92.

Coren S, Hakstian AR. (1992). The development and cross-validation of a self-report inventory to assess pure-tone threshold hearing sensitivity. *J Speech Hear Res* 35:921–928.

Cox RM. (1997). Administration and application of the APHAB. *Hear J* 50:32–48.

Cox RM, Alexander GC. (1995). The Abbreviated Profile of Hearing Aid Benefit. *Ear Hear* 16:176–86.

Cox RM, Alexander GC, Gray GA. (1999). Personality and the subjective assessment of hearing aids. *J Am Acad Audiol* 10:1–13.

Cox RM, Gray GA, Alexander GC. (2001). Evaluation of a Revised Speech in Noise (RSIN) test. *J Am Acad Audiol* 12:423–433.

Dobie RA, Sakai CS. (2001). Estimation of hearing loss severity from the audiogram. In: Henderson D, Prasher D, Salvi RJ, Kopke R, Hamernik R, eds. *Noise Induced Hearing Loss: Basic Mechanisms, Prevention and Control*. London: NRN Publishers, 351–363.

Hawes NA, Niswander PS. (1985). Comparison of the Revised Hearing Performance Inventory with audiometric measures. *Ear Hear* 6:93–97.

Killion MC, Niquette PA. (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? *Hear J* 53:46–53.

Killion MC, Villchur E. (1993). Kessler was right—partly: but SIN test shows some aids improve hearing in noise. *Hear J* 46:31–35.

Kochkin S. (1998). MarkeTrak IV: correlates of hearing aid purchase intent. *Hear J* 51:30–41.

Koike KJ, Hurst MK, Wetmore SJ. (1994). Correlation between the American Academy of Otolaryngology—Hearing and Neck Surgery five-minute hearing test and standard audiologic data. *J Otolaryngol—Head Neck Surg* 111:625–632.

Kramer SE, Kapteyn TS, Festen JM, Tobi H. (1996). The relationships between self-reported hearing disability and measures of auditory disability. *Audiol* 35:277–287.

Lutman ME, Brown EJ, Coles RRA. (1987). Self-reported disability and handicap in the population in relation to pure-tone threshold, age, sex, and type of hearing loss. *Br J Audiol* 21:45–58.

McCartney JH, Maurer JF, Sorenson FD. (1976). A comparison of the Hearing Handicap Scale and the Hearing Measurement Scale with standard audiometric measures on a geriatric population. *J Auditory Res* 16:51–58.

Nabelek AK, Mason D. (1981). Effect of noise and reverberation on binaural and monaural word identification by subjects with various audiograms. *J Speech Hear Res* 24:375–383.

Siegenthaler BM, Strand R. (1964). Audiogram-average methods and SRT scores. *J Acoust Soc Am* 36:589–595.

Studebaker GA. (1985). A "rationalized" arcsine transform. *J Speech Hear Res* 28:255–262.

Weinstein BE, Ventry IM. (1983). Audiometric correlates of the Hearing Handicap Inventory for the Elderly. *J Speech Hear Disord* 48:379–384.