The Performance-Perceptual Test (PPT) and Its Relationship to Aided Reported Handicap and Hearing Aid Satisfaction

Gabrielle H. Saunders and Anna Forsline

Objective: **Results of objective clinical tests (e.g., measures of speech understanding in noise) often conflict with subjective reports of hearing aid benefit and satisfaction. The Performance-Perceptual Test (PPT) is an outcome measure in which objective and subjective evaluations are made by using the same test materials, testing format, and unit of measurement (signal-to-noise ratio, S/N), permitting a direct comparison between measured and perceived ability to hear. Two variables are measured: a Performance Speech Reception Threshold in Noise (SRTN) for 50% correct performance and a Perceptual SRTN, which is the S/N at which listeners perceive that they can understand the speech material. A third variable is computed: the Performance-Perceptual Discrepancy (PPDIS); it is the difference between the Performance and Perceptual SRTNs and measures the extent to which listeners "misjudge" their hearing ability. Saunders et al. in 2004 examined the relation between PPT scores and unaided hearing handicap. In this publication, the relations between the PPT, residual aided handicap, and hearing aid satisfaction are described.**

Design: **Ninety-four individuals between the ages of 47 and 86 yr participated. All had symmetrical sensorineural hearing loss and had worn binaural hearing aids for at least 6 wk before participating. All subjects underwent routine audiological examination and completed the PPT, the Hearing Handicap Inventory for the Elderly/Adults (HHIE/A), and the Satisfaction for Amplification in Daily Life questionnaire. Sixty-five subjects attended one research visit for participation in this study, and 29 attended a second visit to complete the PPT a second time.**

Results: **Performance and Perceptual SRTN and PPDIS scores were normally distributed and showed excellent test-retest reliability. Aided SRTNs were significantly better than unaided SRTNs; aided and unaided PPDIS values did not differ. Stepwise multiple linear regression showed that the PPDIS, the Performance SRTN, and age were significant predictors of scores on the HHIE/A such that greater reported handicap is associated with underestimating hearing ability, poorer aided ability to understand speech in noise, and being younger. Scores on the Satisfaction with Amplification in Daily Life were not well explained by the PPT, age, or audiometric thresholds. When individuals were grouped by their HHIE/A scores, it was seen that individuals who report more handicap than expected based on their audiometric thresholds, have a more negative PPDIS, i.e., underestimate their hearing ability, relative to individuals who report expected handicap, who in turn have a more negative PPDIS than individuals who report less handicap than expected. No such patterns were apparent for the Performance SRTN.**

Conclusions: **The study showed the PPT to be a reliable outcome measure that can provide more information than a performance measure and/or a questionnaire measure alone, in that the PPDIS can provide the clinician with an explanation for discrepant objective and subjective reports of hearing difficulties. The finding that self-reported handicap is affected independently by both actual ability to hear and the (mis)perception of ability to hear underscores the difficulty clinicians encounter when trying to interpret outcomes questionnaires. We suggest that this variable should be measured and taken into account when interpreting questionnaires and counseling patients.**

(Ear & Hearing 2006;27;229–242)

Hearing aids are routinely used for amelioration of hearing impairment, and they play a key role in most aural rehabilitation programs. Satisfaction with hearing aids, however, is disappointingly low, although it is improving. For instance, Kochkin (2005) reports that in 2004, as many as 18.4% of individuals were "dissatisfied" with their hearing aids, and 73.1% of individuals were "satisfied" with their hearing aids. User dissatisfaction with hearing aids is a problem because it often results in abandonment of the aids, which in turn has negative psychosocial consequences, such as stressed interpersonal relations (Hetu, Jones, & Getty, 1993) and depression (Cacciatore et al., 1999). It is necessary, therefore, for researchers and clinicians alike to find ways to improve hearing aid use and satisfaction.

There is often a disconnect between reported satisfaction and measured performance with hearing aids such that the results of objective clinical

National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, Oregon.

0196/0202/06/2703-0229/0 • Ear & Hearing • Copyright © 2006 by Lippincott Williams & Wilkins • Printed in the U.S.A.

tests (such as measures of speech understanding in noise) conflict with subjective reports of hearing aid benefit and satisfaction. For instance, there are investigations in which participants' speech intelligibility scores do not differ between hearing aids/hearing aid settings, although subjective evaluations show a strong listener preference for one model or another (e.g., Horwitz, Turner, & Fabry, 1991; Preminger, Neuman, & Bakke, 2000; Ricketts & Bentler, 1992; Valente, Fabry, Potts, & Sandlin, 1998). There are studies in which little or no relation between reported benefit and measured benefit is found (e.g., Cox & Alexander, 1992; Haggard, Foster, & Iredale, 1981). Yet other investigations find participants reporting strong preferences for one of two pairs of hearing aids worn during a study, even though both pairs of hearing aids were identical (Bentler, Niebuhr, Johnson, & Flamme, 2000; McClymont, Browning, & Gatehouse, 1991). Finally, there are investigations in which hearing aid setting preferences in the laboratory do not transfer to real-world use (Preminger & Cunningham, 2003). It is difficult to reconcile such data, in part because the materials and procedures used for measuring speech performance do not directly relate to the questionnaires used for subjective evaluations.

Saunders, Forsline & Fausti (2004) described a test, known as the Performance-Perceptual Test (PPT), which is an outcome measure in which objective and subjective evaluations are made by using the same test materials, the same testing format, and the same unit of measurement (signal-to-noise ratio, S/N). In this test, the sentence lists, masking noise and the adaptive algorithm from the Hearing In Noise Test (HINT; Nilsson, Gelnett, Sullivan, Soli, & Goldberg, 1992) are used to measure a Performance Speech Reception Threshold in Noise (SRTN) and a Perceptual SRTN. For the Performance SRTN, participants repeat back to the experimenter what they heard (as per HINT guidelines). For the Perceptual SRTN, the experimenter alters the S/N, based on whether participants think that they can "just understand everything that was said." The Perceptual SRTN is thus the S/N at which listeners perceive that they can just understand all of the speech material. A third result is available from this test: the difference between the Performance SRTN and the Perceptual SRTN. It is known as the Performance-Perceptual Discrepancy (PPDIS) and is a measure of the extent to which the listener "misjudges" his or her hearing ability. If the Perceptual S/N is more adverse (a lower S/N) than the Performance S/N, it suggests that listeners overestimate their hearing ability. If the Perceptual S/N is less adverse (a higher S/N) than the Performance S/N, it suggests that listeners underestimate their hearing ability. The test thus permits a direct comparison between perceived ability to hear speech in noise and actual ability to hear speech in noise.

Nabelek, Tucker, & Letowski (1991) developed a measure known as the Acceptable Noise Level (ANL). It establishes the S/N of the maximum level of noise acceptable for ongoing speech played at most comfortable level. The ANL is similar in nature to the Perceptual SRTN in that it combines both a subjective component and an objective component. For the ANL, the objective component is the limitation imposed by the hearing loss, whereas the subjective component is the tolerance the individual has for listening to speech in background noise. For the Perceptual SRTN, the objective component is the limitation on understanding speech in noise imposed by the hearing loss, whereas the subjective component is the individual's perception of that ability. Nabelek and colleagues have shown a relation between ANLs and hearing aid use such that full-time users accepted more background noise than part-time users and that ANLs are unaffected by amplification (Nabelek, Tampas, & Burchfield, 2004).

The PPT has an advantage over the ANL measure in that the measurement of both the Performance and Perceptual SRTNs permits computation of the PPDIS, a variable in which the objective and subjective components are dissociated. This is not possible for the ANL because an S/N for actual performance is not measured.

Saunders et al. (2004) examined the relation between unaided hearing handicap, as measured by the Hearing Handicap Inventory for the Elderly (HHIE; Ventry & Weinstein, 1982) or the Hearing Handicap Inventory for Adults (HHIA; Newman, Weinstein, Jacobson, & Hug, 1990) and the PPT. They found that individuals who reported more handicap than would be expected from their hearing impairment underestimated their hearing ability to a greater extent than did those individuals who reported less handicap than would be expected from their impairment. They further determined that the combination of the Performance SRTN, the PPDIS, and age explained 40% of variance in HHIE/A scores, with the Performance SRTN and the PPDIS each explaining approximately 14% of the variance. The purpose of the current study was to investigate the relations between the PPT, residual aided handicap, and hearing aid satisfaction.

METHODS

Participants

Ninety-four participants between 47 and 86 yr of age (mean, 69.1 yr; SD, 9.2 yr) took part in the experiment. Eighty-nine of the participants were male; five were female. All had symmetrical sensorineural hearing loss and wore binaural hearing aids and had done so for at least 6 wk before participating. Fifty-three participants wore in-the-ear hearing aids, 11 each wore behind-the-ear (BTE) and in-the-canal (ITE) hearing aids, 10 wore completely-in-the-canal aids, 1 wore an ITE/BTE combination, and the style of hearing aids for 8 individuals was not noted. Participants were recruited for the study through fliers posted around the Portland VA Medical Center, from the Portland VA Medical Center Audiology Clinic, and from National Center for Rehabilitative Auditory Research databases. All participants came to the laboratory for the sole purpose of participating in this research study, and all signed an institutional review board–approved consent form. Participants received a $20 reimbursement after each visit.

Test Measures

Pure-Tone Audiometry, Otoscopy, and Tympanometry

Air-conduction thresholds were measured at octave frequencies between 0.25 kHz and 8 kHz along with interoctave frequencies of 1.5, 3.0, and 6.0 kHz, through ER-3A insert earphones, using the American Speech–Language–Hearing Association (1978) recommended procedure. A four-frequency pure-tone average (4F-PTA) was computed by averaging thresholds at 0.5, 1.0, 2.0, and 4.0 kHz from both ears. Otoscopy and tympanometry were conducted to check for cerumen and conductive pathology, respectively. Any participants with conductive pathology or asymmetrical hearing (difference between right and left ear 4F-PTA of more than 15 dB) were excluded from the study.

Performance-Perceptual Test

The PPT is run by using the HINT adaptive protocol, test materials, and speaker configuration (Nilsson et al., 1992). The test materials consist of twelve 20-item sentence lists and accompanying masking noise shaped to the average long-term spectrum of the sentences. The test can be run either in the sound field or using head-related transfer functions to simulate the sound field under headphones. All data presented here are for sound field testing conducted in two conditions: (1) two loudspeakers, 1 m from the listener's head with speech presented from 0° degrees azimuth, noise from 90° azimuth, and (2) two loudspeakers, 1 m from the listeners head with speech presented from 0° degrees azimuth, noise from 270° azimuth. Results from the two conditions were averaged because all participants had symmetrical hearing and thus differences in performance in the two conditions were not expected.

For the PPT, two SRTNs are obtained: (1) a Performance SRTN: the S/N for 50% correct identification of the sentences; (2) a Perceptual SRTN: the S/N at which the listener perceives that he or she can understand everything that was said.

From these, a third variable is computed: the PPDIS: The PPDIS is the difference between the Performance SRTN and the Perceptual SRTN. It measures the degree to which a listener accurately judges his or her hearing ability. It is computed by subtracting the Perceptual SRTN from the Performance SRTN. For example, if the Performance SRTN was −10 dB and the Perceptual SRTN was −5 dB, the PPDIS would be −5. A negative PPDIS indicates that the participant selected a less adverse S/N than they required to perform and can be interpreted as the participant underestimating his/her hearing ability. Conversely, had the Performance SRTN been −5 dB and the Perceptual SRTN been −10 dB, the PPDIS would have been + 5. A positive PPDIS indicates that the participant selected a more adverse S/N than they required to perform and can be interpreted as the participant overestimating his or her hearing ability.

Hearing Handicap Inventory for the Elderly or the Hearing Handicap Inventory for Adults, as Appropriate

The HHIE and HHIA are 25-item questionnaires that assess the social and emotional consequences of hearing loss. The HHIE is for individuals over age 65 yr, whereas the HHIA is for individuals 65 yr and younger. The questionnaires are companion questionnaires that differ in the wording of three questions only. The higher the score on the HHIE/A, the more difficulties the participant reports.

Satisfaction with Amplification in Daily Life

The Satisfaction with Amplification in Daily Life (SADL; Cox & Alexander, 1999) is a 15-item questionnaire. Hearing aid users rate hearing aid satisfaction on four subscales: (1) Positive Effects, (2) Service and Cost, (3) Personal Image, and (4) Negative Features. The higher the score on the SADL, the greater the reported hearing aid satisfaction. Many of the participants were veterans who received their hearing aids free of charge from the VA, thus the Cost item of the Service and Cost scale is omitted from the analyses.

Hearing Aid Use

This entailed participants responding to two questions. (1) For how many years have you worn hearing aids? Less than 8 wk, between 8 wk and 1 yr, between 1 yr and 5 yr, more than 5 yr. (2) On average, for how many hours each day do you wear your hearing aids? Less than 1 hour, between 1 and 4 hours, between 4 and 8 hours, more than 8 hours.

Procedures

Design Overview • After audiometric evaluation, participants completed the HHIE/A, SADL, and hearing aid usage items. The output of each participant's hearing aids at the user gain setting was then measured with a Fonix 6500-CX hearing aid analyzer to document hearing aid output and to ensure that all hearing aids were functioning appropriately. Participants then carried out the PPT for both unaided and aided listening. Sixty-five participants attended just one research visit for participation in this study, whereas 29 returned between 4 days and 42 days later to complete the PPT a second time to establish test-retest reliability. At this second visit, testing was conducted with the hearing aid output matched to that used for testing at the first visit. The Fonix hearing aid analyzer was used to confirm this.

PPT Testing

For PPT testing, the level of the noise masker was fixed at 65 dB SPL, and the level of the sentences was adjusted adaptively. The order of loudspeaker configuration conditions was counterbalanced across participants. For consistency, and in keeping with the procedure recommended for clinical testing in Saunders, Field, & Haggard (1992), the Perceptual SRTN was always measured before the Performance SRTN. Unaided testing was always completed before aided testing. For aided testing where applicable, participants were told to select the program they used most often and were permitted to adjust the volume control of their hearing aids to a comfortable level while listening to a practice list of HINT sentences played at 65 dB SPL in quiet. If subjects adjusted the hearing aid program or the volume, the output of the hearing aid was measured a second time, and these settings were used at retest.

For the Perceptual SRTN (i.e., the S/N at which the listener perceives that he or she can understand everything that was said), participants were seated in a sound-attenuating booth. They were instructed as follows: "In this test you will be hearing some sentences in background noise. After each sentence, we want to know whether you could understand everything that was said. Say 'yes' if you could understand everything or 'no' if you could not. The volume of the sentences will change; sometimes they will be quite loud, sometimes they will be quiet. This is intentional. So, as I said, we want to know whether or not you could just understand everything that was said." The sentences were then presented by using the computerized HINT system. The level of the noise was held constant and the level of the sentences was varied. The S/N was made more adverse if the participants said they could understand everything and less adverse if they said they could not. If participants said they were uncertain, the S/N was kept the same. The software engineers from House Ear Institute provided us with an additional response button for "uncertain" responses that retained the previous S/N. The default automated HINT adaptive algorithm was used, in which the level of the noise is fixed while the level of the speech is adjusted in 4-dB steps for sentences 1 though 4 and in 2-dB steps thereafter. The final Perceptual SRTN was computed by averaging the S/N for sentences 5 though 20, along with the S/N at which a 21st sentence would have been presented.

For the Performance SRTN (i.e., the S/N for 50% correct identification of the sentences), participants remained seated in the sound-attenuating booth. They were instructed in accordance with the recommended HINT protocol to repeat back as much as they could of each sentence, even if it was not complete. As for the Perceptual SRTN, the default automated HINT adaptive algorithm was used. The final Performance SRTN was computed by averaging the S/N for sentences 5 though 20, along with the S/N at which a 21st sentence would have been presented. Any result for which the standard deviation of test presentation levels met or exceeded the 95th percentile for the distribution of standard deviations, as defined in the HINT manual, was rerun.

Note that the only difference in procedure between the Perceptual SRTN and the Performance SRTN is the response given by the participant. The adaptive procedure, the masking noise, and the speaker configurations are identical.

Hearing Aid Settings

For testing, 94 of the participants used omnidirectional microphones; of these, 93 selected an all-purpose program (program 1), whereas one selected an omnidirectional program for listening in noise. Just one individual used directional microphones.

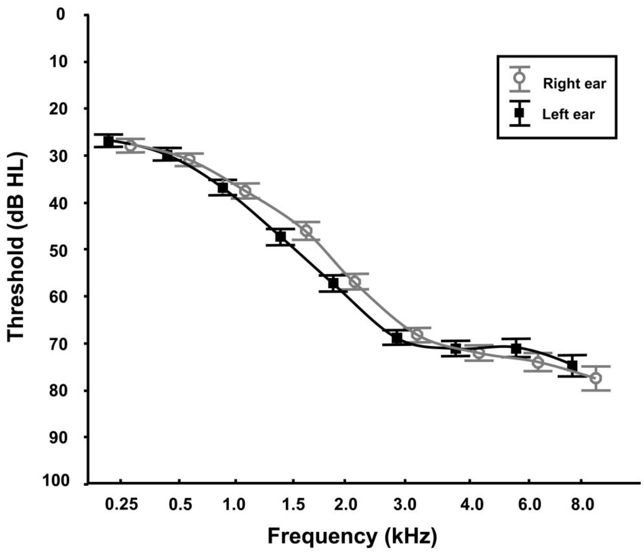


Fig. 1. Mean pure-tone thresholds (in dB HL) with ±1 standard error bars. Left ear data are shown by filled squares and right ear data by open circles.

RESULTS

Audiometric Data

Figure 1 shows the mean pure-tone thresholds for the left and right ears averaged across participants with error bars showing ±1 standard error. These data confirm that participants had symmetrical hearing loss and show a typical audiometric configuration of participants with mild to moderate sensorineural impairment.

Performance-Perceptual Test Data

Performance and Perceptual SRTNs • Figure 2 shows the distributions of unaided and aided SRTN values. It is seen that scores are normally distributed around the mean and that aided performance is better than unaided performance, i.e., participants obtain a lower (more adverse) S/N in the aided condition than in the unaided condition. Analysis of variance (ANOVA) shows the main effects of aiding and of type of SRTN to be statistically significant

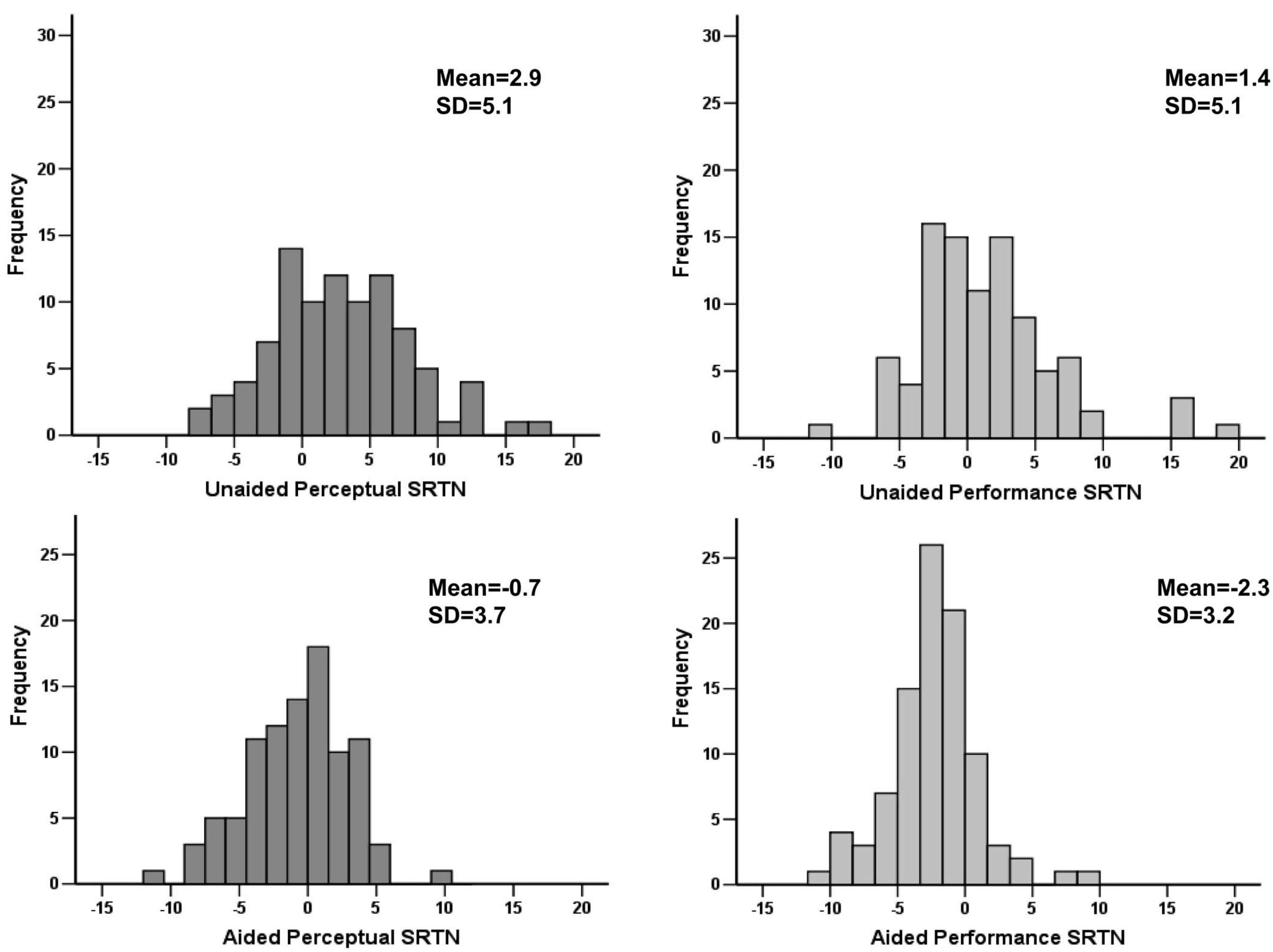


Fig. 2. Histograms showing the distribution of unaided and aided Performance and Perceptual SRTNs. Group means and standard deviations are shown on each graph.

(Aiding: $F = 28.4$, $p < 0.001$; SRTN: $F = 132.2$, $p < 0.001$) but the interaction between aiding and SRTN to be nonsignificant ($F = 0.1$, $p = 0.752$). In other words, participants' scores improved significantly when tested wearing hearing aids, but the change was independent of whether they were completing the Performance SRTN or the Perceptual SRTN. Although hearing aid output was not specifically compared with a target, the presence of aided performance benefit confirms that the hearing aid settings were acceptable.

Figure 3 shows scatterplots of the relation between the Perceptual and Performance SRTNs for unaided and aided listening. Perceptual SRTN scores are plotted on the *x*-axis, Performance SRTN scores on the *y*-axis. The regression line for the points is plotted (solid line), along with dashed lines delineating the 68% confidence interval (CI). Participants with scores above the regression line have a Performance SRTN that is higher than their Perceptual SRTN, indicating that these individuals err in the direction of overestimating their hearing ability. Individuals with scores falling below the regression line have a Performance SRTN that is lower than their Perceptual SRTN, indicating that they err in the direction of underestimating their hearing ability. From a clinical perspective, our interest is in those participants whose Performance and Perceptual SRTNs differ "substantially." For purposes of this study, a "substantial" difference is defined as a correlation between the Performance and Perceptual SRTN that is more than ±0.5 SD from the mean correlation of the study population, i.e., those participants whose scores fall outside of the 68% CI shown on Figure 3. The number of individuals falling above and below the 68% CI are shown on each graph in bold print.

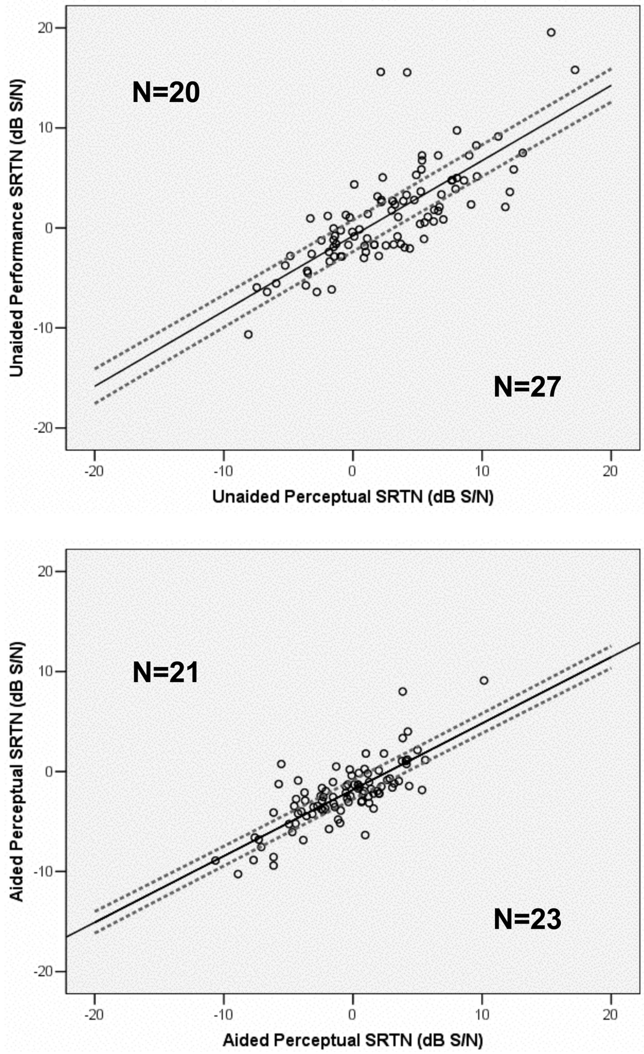


Fig. 3. Scatterplots of Perceptual versus Performance SRTNs for unaided and aided listening. Regression line for the points is plotted (solid line), along with dashed lines delineating the 68% CI.

Performance-Perceptual Discrepancy

Figure 4 shows the distributions of unaided and aided PPDIS values. It is seen that scores are normally distributed around the mean and that the aided PPDIS has less variance than the unaided PPDIS. A *t*-test showed that the mean unaided and aided PPDIS values do not differ statistically ($t = 0.3$, $p = 0.75$).

Test-Retest Reliability

Test-retest reliability was examined for the 29 individuals who completed the PPT twice. The Pearson *r* values and associated 95% confidence limits for both aided and unaided testing are shown in Table 1. All values are significant at $p < 0.001$, and even the lower confidence limits can be considered excellent correlations. The test-retest reliability for the Perceptual SRTNs is almost identical to that for the Performance SRTNs. These test-retest values are similar to those published for unaided listening in Saunders et al. (2004). As was discussed in that paper, the fact that participants are as reliable at carrying out the Perceptual task as they are at the Performance task is perhaps surprising because it generally assumed that individuals are not as reliable at "subjective" tasks as they are at "objective" tasks. It is likely that participants are highly reliable at the perceptual task here because it requires only a simple judgment.

The PPDIS test-retest values are slightly lower than the values for the SRTNs but are still within the range that would be considered excellent. The reason the PPDIS test-retest values are lower than for the SRTNs is because the PPDIS is derived from

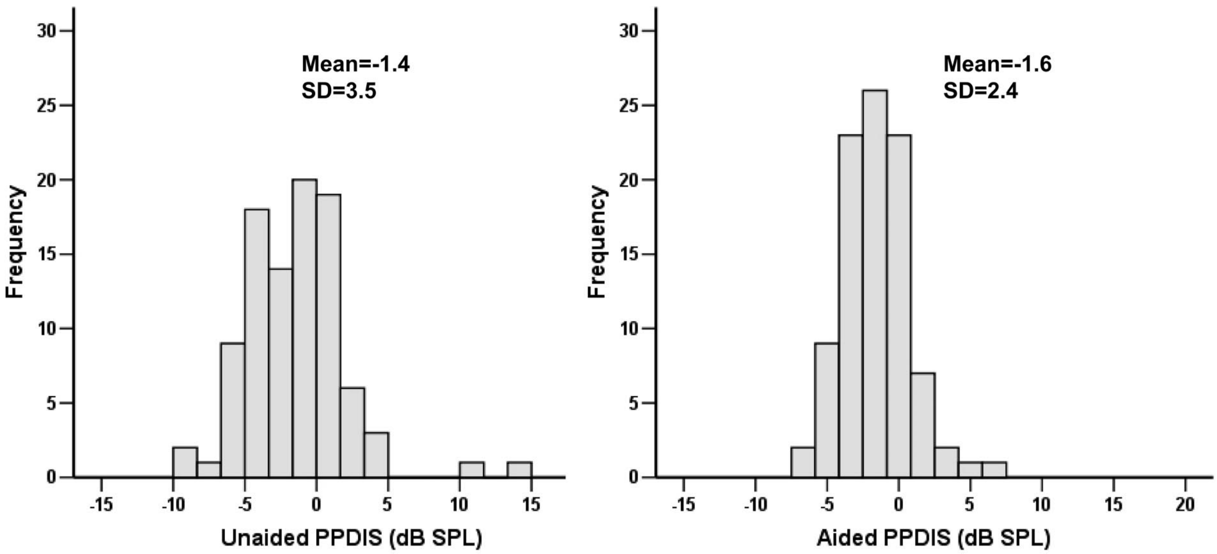


Fig. 4. Histograms showing the distribution of PPDIS for unaided and aided listening.

two other variables that are correlated, each of which has its own variability.

Having confirmed that the unaided PPT data and unaided test-retest values are similar to those reported in Saunders et al. (2004), the remainder of this paper will be devoted to aided PPT data only.

Male Versus Female Participants

As stated above, only five of the 89 participants were female. ANOVAs were used to confirm that the five female participants did not differ from the 89 male participants in terms of their pure-tone thresholds and/or PPT scores. Table 2 shows the mean scores for the male and female participants and the results of the ANOVAs comparing them. No significant group differences were found to exist; thus, for all analyses, the data from male and female participants are combined.

Hearing Aid Styles

To examine whether hearing aid style affected performance on the PPT, analyses of covariance (ANCOVAs) using hearing aid style as the fixed factor and 4F-PTA as a covariate were conducted. The 4F-PTA was used as a covariate because there is a relation between hearing aid style and the 4F-PTA ($F = 3.5$, $p = 0.01$) such that individuals wearing BTEs had significantly poorer hearing than individuals wearing completely-in-the-canal aids. The ANCOVAs showed no differences in PPT performance across styles of hearing aid (Perceptual SRTN: $F = 0.6$, $p = 0.634$; Performance SRTN: $F = 0.4$, $p = 0.778$; PPDIS: $F = 0.6$, $p = 0.660$); thus, style of hearing aid will not be taken into consideration in any later analyses.

Hearing Aid Use

Data regarding lifetime hearing aid use and daily hearing aid use were collected (see Table 3). The mean 4F-PTA and age of participants in each category are included in the table, along with the results of a one-way ANOVA comparing the age and 4F-PTA of participants across hearing aid use groups. It is seen that the majority of participants in this study had been hearing aid users for more than 5 yr and that most individuals wore their hearing aids for more than 8 hours per day. Age was not related to hearing aid use; degree of hearing loss, however, was related to lifetime use, such that individuals who were long-term hearing aid users had poorer hearing than newer users. Surprisingly, audiometric thresholds were not related to daily hearing aid use.

Correlates of Scores on the PPT

Table 4 shows the Pearson r values and associated 95% confidence limits for correlations between the PPT variables, age, and the 4F-PTA. The Perceptual SRTN correlates significantly with the 4F-PTA but not age. The Performance SRTN correlates significantly with both the 4F-PTA and age, although the latter correlation becomes nonsignificant when 4F-PTA is accounted for. The PPDIS is related neither to age or the 4F-PTA.

Predictors of HHIE/A Scores

Stepwise multiple linear regression analysis was then used to determine the relative contribution of the PPT variables, age, and audiometric thresholds toward HHIE/A scores. The results are shown in Table 5. The Performance SRTN, PPDIS, age, and

TABLE 1. Test-retest reliability with 95% confidence limits in parentheses for unaided and aided PPT variables

		Visit 1		
Visit 2	Condition	Perceptual SRTN *r* value	Performance SRTN *r* value	PPDIS *r* value
Perceptual SRTN	Unaided	0.952 (0.903–0.977)		
	Aided	0.944 (0.887–0.973)		
Performance SRTN	Unaided		0.974 (0.947–0.987)	
	Aided		0.924 (0.849–0.963)	
PPDIS	Unaided			0.880 (0.767–0.940)
	Aided			0.810 (0.643–0.904)

PPT, Performance-Perceptual Test; SRTN, Speech Reception Threshold in Noise; PPDIS, Performance-Perceptual Discrepancy.

4F-PTA were used as the independent variables, and the HHIE/A Total score was used as the dependent variable. The Emotional subscale and Social subscale of the HHIE/A were not analyzed separately because scores on these scales correlate so closely ($r = 0.867, p < 0.001$). The Perceptual SRTN was not used as an independent variable because it is comprised of a combination of the Performance SRTN and the PPDIS and it was desirable to permit both the Performance SRTN and PPDIS to enter the equation separately rather than have the Perceptual SRTN enter in place of both.

The PPDIS, the Performance SRTN, and age are all significant predictors of scores on the HHIE/A. By examining the β-value of the variable in the regression equation, it is seen that greater reported handicap is associated with underestimating hearing ability, poorer aided ability to understand speech in noise, and being younger. The Performance SRTN and the PPDIS each explain approximately 10% of the total variance in HHIE/A scores, and age explains a further 5%. Between these three variables, then, almost 25% of the total variance in HHIE/A scores is explained.

These relations are further illustrated in Figure 5. Participants were classified into three groups according to the mean expected HHIE/A score for their audiometric thresholds, as specified in Ventry et al. (1982) and Newman et al. (1990). Participants depicted by an open square (Low Handicap Group) have an HHIE/A score more than 0.5 SD below the expected HHIE/A score for their pure-tone average. Participants depicted by the filled circle (Average Handicap Group) have an HHIE/A score that is within ±0.5 SD of the mean expected score for their pure-tone average, and participants depicted by the X (High Handicap Group) have an HHIE/A score more than 0.5 SD above the mean HHIE/A score for their pure tone average. In other words, participants in the Low group report less handicap than expected, participants in the Average group report the expected degree of handicap, and participants in the High group report more handicap than expected. The mean Perceptual SRTN, mean Performance SRTN, and mean PPDIS value of participants in each of the handicap groups is plotted, along with error bars showing ±1 SEM. It is seen that individuals who report more handicap than expected select a higher S/N for the Perceptual SRTN than the individuals who report expected handicap. In turn, these individuals select a higher S/N than the individuals who report less handicap than expected. This is reflected in the PPDIS scores of individuals who report more handicap than expected having more a negative PPDIS, i.e., underestimating their hearing ability relative to individuals who report expected handicap, who in turn have a more negative PPDIS than individuals who report less handicap than expected. No such patterns were apparent for the Performance SRTN. That is, reported handicap across these three groups is independent of actual ability to understand speech in noise. ANCOVAs, using 4F-PTA as a covariate for the SRTNs, were conducted to determine whether these differences were significant. The 4F-PTA was used as a covariate for the SRTNs because of the high corre-

TABLE 2. Means (standard deviations in parentheses) and results of ANOVAs comparing scores of male (n = 89) and female (n = 5) participants

Variable	Male	Female	F = p <
Low-frequency PTA*	28.7	35.0	1.2
	(12.3)	(16.4)	0.272
Mid-frequency PTA†	54.4	52.2	0.17
	(11.9)	(11.8)	0.681
High-frequency PTA‡	74.6	59.8	3.7
	(15.4)	(33.5)	0.057
Aided Perceptual SRTN	−0.8	−0.1	0.2
	(3.5)	(6.9)	0.671
Aided Performance SRTN	−2.3	−1.6	0.2
	(2.9)	(6.4)	0.627
Aided PPDIS	−1.6	−1.6	0.0
	(2.4)	(1.7)	0.992

*Mean of thresholds at 0.25 and 0.5 kHz; †mean of thresholds at 1.0, 2.0, and 3.0 kHz; ‡mean of thresholds at 4.0, 6.0, and 8.0 kHz.

PTA, Pure-Tone Average; SRTN, Speech Reception Threshold in Noise; PPDIS, Performance-Perceptual Discrepancy.

TABLE 3.
Hearing aid use data for all subjects

Hearing aid use	No. of subjects	4F-PTA	Age	Hours use per day	No. of subjects	4F-PTA	Age
Less than 2 mos	6	40.0	65.5	Less than 1 hr	13	46.1	66.7
2 mos to 1 yr	9	42.6	65.3	1 to 4 hrs	9	46.6	68.6
1 to 5 yrs	21	46.4	68.0	4 to 8 hrs	16	49.6	69.3
More than 5 yrs	58	52.1	70.5	More than 8 hrs	56	50.1	69.8
ANOVA:		$F = 5.4$ $p = 0.002$	$F = 1.4$ $p = 0.25$	ANOVA:		$F = 0.7$ $p = 0.54$	$F = 0.4$ $p = 0.76$

4F-PTA, Four-Frequency Pure-Tone Average.

TABLE 4. Results of Pearson correlations between PPT variables, age, and 4F-PTA

Variable	Perceptual SRTN	Performance SRTN	PPDIS
4F-PTA	0.525* (0.365 to 0.655)	0.640* (0.506 to 0.744)	0.039 (−0.161 to 0.236)
Age	0.163 (−0.037 to 0.350)	0.233† (0.036 to 4.412)	0.057 (−1.143 to 0.235)
Age partially 4F-PTA	0.048 (−0.152 to 0.244)	0.110 (−0.090 to 0.302)	0.049 (−0.151 to 0.245)

$*p < 0.001$, †$p < 0.05$.
PPT, Performance-Perceptual Test; 4F-PTA, Four-Frequency Pure-Tone Average; SRTN, Speech Reception Threshold in Noise; PPDIS, Performance-Perceptual Discrepancy.

TABLE 5. Results of stepwise multiple regression analyses for HHIE/A scores

Independent variable entering regression equation	% Variance explained by variable (R^2 change)	Value in final equation	Significance $p <$
PPDIS	8.9	−0.335	0.001
Performance SRTN	10.4	0.376	0.001
Age	4.8	−0.226	0.025
Total variance explained	24.1		

Only variables entering the regression equation with a significance of $p < 0.05$ or less are shown.
HHIE/A, Hearing Handicap Inventory for the Elderly/Adults; PPDIS, Performance-Perceptual Discrepancy; SRTN, Speech Reception Threshold in Noise.

lation between audiometric thresholds and the SRTNs. The ANCOVA showed significant differences across Handicap groups in the Perceptual SRTN ($F = 6.4$, $p = 0.002$) and the PPDIS ($F = 4.9$, $p = 0.010$) but not in the Performance SRTN ($F = 1.4$, $p = 0.244$). Tukey HSD post hoc tests were conducted to determine which Handicap groups differed significantly from one another. These are shown in Figure 5 by a horizontal line between the groups that differed significantly. For the Perceptual SRTN it is seen that the Low Handicap group differs significantly from both the Average and High Handicap groups, such that participants in the Low Handicap group have better (more adverse S/Ns) Perceptual SRTNs than participants in the other two groups. It should be noted that the data plotted in the figure are the raw data, whereas the ANCOVAs correct for the effect of audiometric thresholds. For the PPDIS, participants in the High Handicap group have significantly lower PPDIS values than participants in the Low Handicap group. In other words, individuals who reported more handicap than expected underestimated their hearing ability relative to individuals reporting less handicap than expected.

Predictors of SADL Scores

The stepwise multiple linear regression analysis conducted for the HHIE/A scores above was repeated for the SADL scores to determine the relative contribution of the PPT variables, age, and audiometric thresholds toward hearing aid satisfaction. Because it is logical to assume that hearing aid satisfaction might be a function of hearing aid benefit, measured benefit for speech in noise was also included as an independent variable. Measured hearing aid benefit was computed by subtracting the aided Performance SRTN from the unaided Performance SRTN. The results are shown in Table 6.

Variance in SADL scores is not satisfactorily explained by the independent variables. At most, 16% of the variance was explained for the SC scale, whereas 6% or less of the variance on the other three scales was explained. Only 9% of the Global score was explained. Hearing-based measures (Performance SRTN and Measured benefit) are the most important variables, explaining variance in the PE, SC, PI scales, and the Global score. In each instance, better SADL scores are associated with better performance on the Performance SRTN or more measured hearing aid benefit. The PPDIS explained 4%

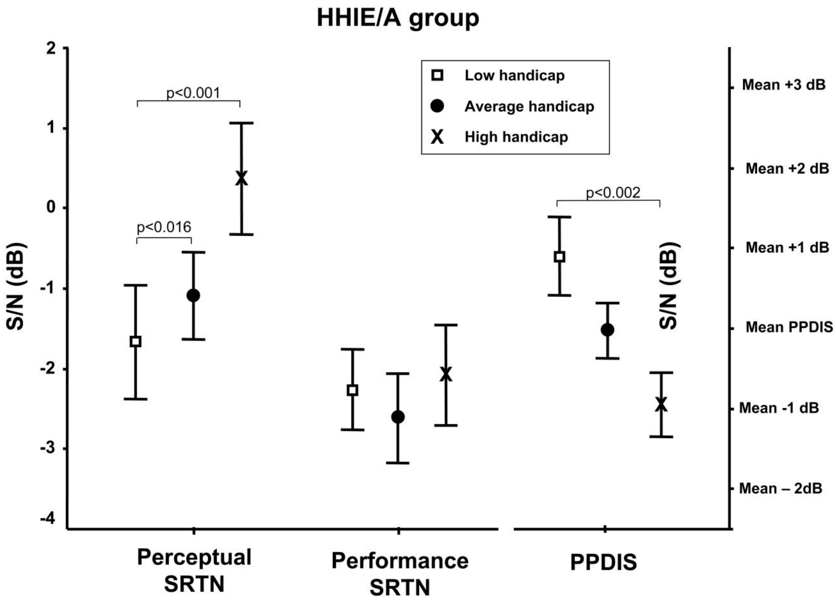


Fig. 5. Group mean Perceptual SRTN, Performance SRTN, and PPDIS values with ±1 standard error bars. Subjects are grouped according to the mean expected HHIE/A score for their audiometric thresholds. Open squares (Low Handicap group) depict subjects with an HHIE/A score more than 0.5 SD above the mean expected score. Filled circles (Average Handicap group) depict subjects with an HHIE/A score that is within ±0.5 SD of the mean expected score, and X (High Handicap group) depicts subjects who have an HHIE/A score more than 0.5 SD above the mean expected score.

of the variance in the SC scales such that participants with a lower PPDIS (less confidence in their hearing ability) perceived that they had received better service.

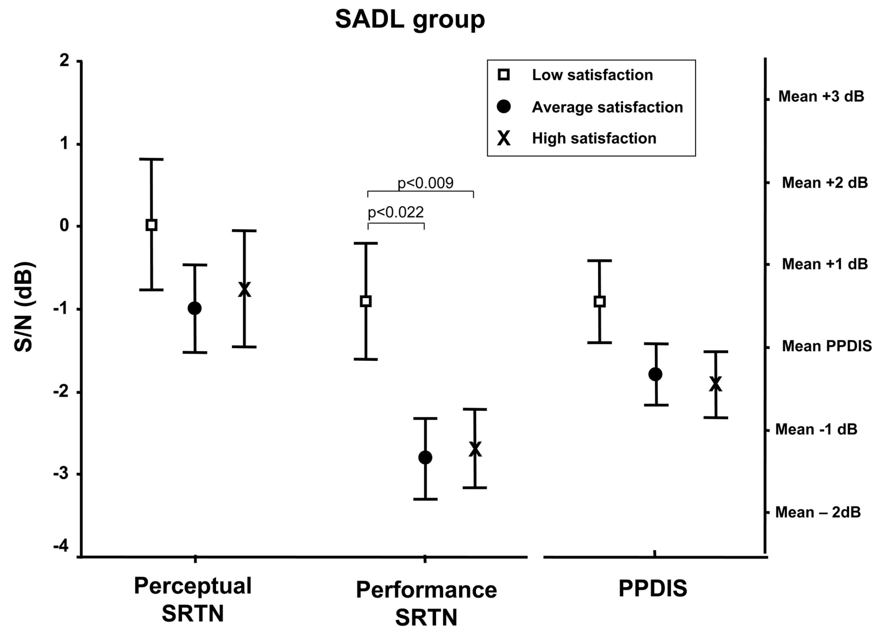
To parallel the analyses conducted for the HHIE data, the participants were classified into three groups, based on their Global SADL score, using the SADL normative data published by Cox et al. (1999) and the data for PPT variables plotted (see Figure 6). The mean and standard deviation values used were for the Global score excluding the reasonable cost item, because many of the study participants were veterans service-connected for hearing impairment who had received their hearing aids at no cost. Participants depicted by an open square (Low Satisfaction group) have a SADL score more than 0.5 SD below the mean published scores. Participants depicted by the filled circle (Average Satisfaction group) have a SADL score that is within ±0.5 SD of average published SADL scores, and participants depicted by the X (High Satisfaction group) have a SADL score more than 0.5 SD above the mean published SADL scores. Univariate ANOVAs, using 4F-PTA as a covariate, showed that there was no relation between Global SADL score and the Perceptual SRTN ($F = 0.4$, $p = 0.682$) but that significant Satisfaction group differences existed for the Performance SRTN ($F = 4.1$, $p = 0.020$). Pairwise comparisons, again shown on the figure, revealed that individuals in the Low Satisfaction

TABLE 6. Results of stepwise multiple regression analyses for SADL scores

Independent variable entering regression equation	% Variance explained by variable (R^2 change)	Value in final equation	Significance $p <$
SADL PE score			
Measured benefit	5.9	0.244	0.018
Total variance explained	5.9		
SADL SC score			
Performance SRTN	12.6	−0.324	0.001
PPDIS	3.9	−0.200	0.045
Total variance explained	16.5		
SADL NF score			
No variables entered			
Total variance explained	0.0		
SADL PI score			
Performance SRTN	4.4	−0.210	0.042
Total variance explained	4.4		
Global SADL score			
Performance SRTN	9.3	−0.306	0.003
Total variance explained	9.3		

Only variables entering the regression equation with a significance of $p < 0.05$ or less are shown.
SADL, Satisfaction with Amplification in Daily Life; SRTN, Speech Reception Threshold in Noise.

Fig. 6. Group mean Perceptual SRTN, Performance SRTN, and PPDIS values with ±1 standard error bars. Subjects are grouped according to the Global SADL norms. Open squares (Low Satisfaction group) depict subjects with an SADL score more than 0.5 SD below average. Filled circles (Average Satisfaction group) depict subjects with an SADL score that is within ±0.5 SD of average, and X (High Satisfaction group) depicts subjects who have an SADL score more than 0.5 SD above average.



group (open square) performed more poorly (i.e., required a less adverse S/N to understand 50% of the sentences) than individuals in the Average and High Satisfaction groups. For the PPDIS, it is seen that individuals in the High Satisfaction group have higher PPDIS values than the other two groups, suggesting that individuals who overestimate their hearing report more satisfaction with their hearing aids than other individuals. However, ANOVA showed this across group comparison to be nonsignificant ($F = 1.5$, $p = 0.241$).

DISCUSSION

In this paper, we describe the relation between aided PPT data, reported hearing handicap, and hearing aid satisfaction among a group of mostly male hearing aid users. These data are an extension of the data presented in an earlier volume of this journal by Saunders et al. (2004) that described the relation between unaided PPT data and reported hearing handicap.

The scatterplots in Figures 2 and 3 show that the Perceptual and Performance SRTNs of the majority of individuals are very similar. That is, the S/N at which individuals can actually understand 50% of the material, and the S/N at which they think they can understand 50% of the material, are almost the same. For example, 49% of participants have unaided, and 51% have aided Performance and Perceptual SRTNs that are within one test step size of each other (2 dB). Other individuals, however, have fairly large discrepancies between their Perceptual and Performance SRTNs. Some think they need a more advantageous S/N than they really do to understand the material, i.e., they underestimate their hearing ability; whereas others think they can understand the material with a more adverse S/N than they actually can, i.e., they overestimate their hearing ability. It is noteworthy that 57% of the individuals who overestimate their aided hearing ability, such that they fall outside the 68% CIs, also overestimate their unaided hearing ability, and 56.5% of the individuals who underestimate their aided hearing ability also underestimate their unaided hearing ability. Likewise, the unaided and aided PPDIS are highly correlated ($r = 0.781$, $p < 0.001$), and, although there are significant differences between the aided and unaided SRTNs, there is not a significant difference between the aided and unaided PPDIS. We suggest that this is because the PPDIS is not specific to the listening situation but is reflective of the individual's approach to judging his or her auditory ability. We are unsure whether it is related specifically to judgment of hearing or whether it might also extend to self-judgment of other skills. A finding that sheds some light on this was reported by Saunders & Haggard (1993) in a study in which the PPT scores and personality data of individuals with obscure auditory dysfunction (OAD)* were compared with those of individuals who had chronic pelvic pain without obvious organic pathology (CPPWOP). They found that the participants with OAD differed from the participants with CPPWOP on hearing-based measures but not on the PPDIS and other personality test

*Individuals who complain of hearing difficulties and yet have clinically normal hearing.

scores. This supports the hypothesis that the direction of the PPDIS judgment (i.e., under versus overestimation) might extend to self-judgment of other skills. Nabelek et al. (2004) also found no difference between aided and unaided ANLs and hypothesized that this was because the ANL is dependent on the individual person rather than the listening situation.

The results of the Pearson correlations shown in Table 4 confirm the findings of Saunders et al. (2004), showing that both SRTNs have a hearing-related component, i.e., they correlate significantly with the 4F-PTA, but the PPDIS does not. This underscores the difference between the PPT and the ANL measure of Nabelek, Tucker, & Letowski (1991). By having available the Performance SRTN and the PPDIS, the hearing-based component and a subjective component can be examined separately, whereas the ANL provides only a combined value, equivalent to the Perceptual SRTN. As found in previous work with the PPT (Saunders & Cienkowski, 2002; Saunders et al., 2004), the PPDIS is not significantly correlated with age. It is reassuring to confirm this because research has shown that older individuals often become more cautious (e.g., Calhoun & Hutchison, 1981; Deakin, Aitken, Robbins, & Sahakian, 2004). However, this evidently does not apply to the PPDIS.

The finding that the Performance SRTN, the PPDIS, and age are predictors of HHIE/A scores replicates the finding of Saunders et al. (2004) for unaided listening. As discussed in that publication, it is reasonable to expect that individuals who have difficulty understanding speech in noise will report more handicap than individuals who have less difficulty. It is also reasonable to expect that individuals who have little confidence in their hearing ability (underestimate it) will report more handicap than individuals who have more confidence in their hearing ability, although this is not something that is routinely measured or taken into consideration during counseling. As for age, past work has shown that for the same degree of hearing loss, younger individuals report more handicap than older individuals (Merluzzi & Hinchcliffe, 1973; Wiley, Cruickshanks, Nondahl, & Tweed, 2000), perhaps because older individuals expect to encounter disabilities as they age and thus do not report those disabilities on a questionnaire. Another explanation is that older individuals might have less reliance on their hearing than younger individuals and thus hearing impairment has fewer negative effects.

The SADL scores are not well predicted by any of the variables measured in this study. In light of the fact that three of the four scales have little to do with either perceived or actual hearing in noise (PI, SC, and NF), this is perhaps not surprising. Of all the measures examined, however, hearing-based measures play the most consistent role in explaining the variance in the SADL scores. For each scale, individuals who have better aided performance for understanding speech in noise, or who gain more measured benefit for understanding speech in noise, had better SADL scores. This finding is not unexpected in light of the work of Walden & Walden (2004), who found significant correlations between hearing aid success, as measured by the International Outcome Inventory for Hearing Aids (IOA-HA; Cox & Alexander, 2002), and measures of speech in noise (QuickSIN Speech in Noise Test, 2001) for aided listening. However, as in this study, they did not find a significant relation between SADL scores and threshold-based measures or speech in quiet measures.

Clinical Application and Future Work

The relevance of the PPDIS is best illustrated in Figure 5. In this figure, it is clearly seen that there is a strong relation between estimation of hearing ability and reported hearing difficulties. Individuals who report more handicap than would be expected based on their audiometric thresholds, have a more negative PPDIS than average (i.e., underestimate their hearing), whereas those who report less handicap than expected based on their audiometric thresholds, have a more positive PPDIS than average (i.e., overestimate their hearing). In other words, responses on the HHIE/A are strongly influenced by the degree and direction to which individuals misjudge their hearing ability, independent of their actual ability to hear. There is no reason to believe this would not also be the case for other disability and handicap questionnaires. Clinically, PPDIS data could be used in counseling. The PPDIS could be explained to the patient, their actual PPDIS would be revealed, and a discussion of the ramifications of underestimating or overestimating hearing ability could take place. Such counseling would seem particularly apt for at least two patient groups: (1) individuals who report more difficulties than would be expected, based on their measured performance (sometimes referred to as "complainers"), and (2) individuals who deny hearing difficulties even though they have a substantial hearing loss (sometimes referred to as "deniers"). A study examining the efficacy of PPDIS-based counseling for hearing aid users is currently underway to determine whether such counseling can alter an individual's PPDIS and/or alter responses on a handicap and disability questionnaire.

SUMMARY AND CONCLUSIONS

In summary, this study has shown the PPT to be a reliable outcome measure that can provide more information than a performance measure and/or a questionnaire measure inasmuch as the PPDIS can provide the clinician with an explanation for discrepant objective and subjective reports of hearing difficulties, although it does not appear to be as effective at explaining hearing aid satisfaction, as measured by the SADL. This probably is because the SADL is a measure of the patient's overall "hearing aid experience," reflecting satisfaction with encounters with the clinician through to hearing aid effectiveness and self image. This emphasizes the multidimensional nature of hearing aid outcome and should alert the clinician to the necessity of deciding in advance what aspect of outcome is of interest when selecting a measurement tool.

The key finding of this study is that self-reported handicap is affected independently by both actual ability to hear and the perception of ability to hear and underscores the difficulties clinicians encounter when trying to interpret outcomes questionnaires. That is, when reported difficulties conflict with measured performance, clinicians do not usually have an explanation as to why. These data suggest that underestimation or overestimation of hearing ability plays a very important role.

In current clinical practice, only the hearing-related component is measured, thus leaving unexplained much of the basis for reported hearing handicap. This work suggests that measuring the PPDIS might provide an explanation. A study is currently being undertaken to determine whether providing patients with an explanation of their PPDIS and discussing the ramifications can better adjust an individual's PPDIS or at least alter the handicap/disability they report. We suggest that this should be taken into account when interpreting questionnaires and counseling patients.

ACKNOWLEDGMENTS

This research was supported by the Department of Veterans Affairs Health Administration, Rehabilitation Research and Development Grant C2709I, and the National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, Oregon. We thank ShienPei Silverman for administrative support.

Address for correspondence: Gabrielle H. Saunders, PhD, National Center for Rehabilitative Auditory Research, Portland VA Medical Center, 3710 SW US Veterans Hospital Road, Portland, OR 97207. E-mail: gabrielle.saunders@med.va.gov

Received August 1, 2005; accepted December 20, 2005.

REFERENCES

American Speech-Language-Hearing Association. (1978). Guidelines for pure-tone audiometry. *ASHA*, *20*, 297–301.

Bentler, R., Niebuhr, D, Johnson, T., & Flamme, G. (2000). Impact of digital labeling on outcome measures. *Ear and Hearing*, *24*, 215–224.

Cacciatore, F., Napoli, C., Abete, P., Marciano, E, Triassi, M., & Rengo, F. (1999). Quality of life determinants and hearing function in an elderly population: Osservatoria Geriatrico Campano study group. *Gerontology*, *45*, 323–328.

Calhoun, R., & Hutchison S. J. (1981). Decision-making in old age: cautiousness and rigidity. *International Journal of Aging and Human Development*, *13*, 89–98.

Cox, R., & Alexander, G. (1992). Maturation of hearing aid benefit. *Ear and Hearing*, *13*, 131–141.

Cox, R., & Alexander, G. (1999). Measuring Satisfaction with Amplification in Daily Life: The SADL Scale. *Ear and Hearing*, *20*, 306–320.

Cox, R., & Alexander, G. (2002). The International Outcome Inventory for Hearing Aids (IOI-HA): psychometric properties of the English version. *International Journal of Audiology*, *41*, 30–35.

Deakin, J., Aitken, M., Robbins, T., & Sahakian, B. (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society*, *10*, 590–598.

Haggard, M., Foster, J., & Iredale, F. (1981). Use and benefit of postaural aids in sensory hearing loss. *Scandinavian Audiology*, *10*, 45–52.

Hetu, R., Jones, L., & Getty, L. (1993). The impact of acquired hearing impairment on intimate relationships: implications for rehabilitation. *Audiology*, *32*, 363–381.

Horwitz, A., Turner, C., & Fabry, D. (1991). Effects of different frequency response strategies upon recognition and preference for audible speech stimuli. *Journal of Speech and Hearing Research*, *34*, 1185–1196.

Kochkin, S. (2005). MarkeTrak VII: hearing loss population tops 31 million people. *The Hearing Review 12*, 16–29.

McClymont, L., Browning, G., & Gatehouse, S. (1991). Reliability of patient choice between hearing aid systems. *British Journal of Audiology*, *25*, 35–39.

Merluzzi, F., & Hinchcliffe, R. (1973). Threshold of subjective auditory handicap. *Audiology*, *12*, 65–69.

Nabelek, A., Tampas, J., & Burchfield, S. (2004). Comparison of speech perception in background noise with acceptance of background noise in aided and unaided conditions. *Journal of Speech, Language and Hearing Research*, *47*, 1001–1011.

Nabelek, A., Tucker, F., & Letowski, T. (1991). Toleration of background noises: relationship with patterns of hearing aid use by elderly persons. *Journal of Speech, Language and Hearing Research*, *34*, 679–685.

Newman, C., Weinstein, B., Jacobson, G., & Hug, G. (1990). The Hearing Handicap Inventory for Adults: psychometric adequacy and audiometric correlates. *Ear and Hearing*, *11*, 430–433.

Nilsson, M., Gelnett, D., Sullivan, J., Soli, S., & Goldberg, R. (1992). Norms for the Hearing In Noise Test: the influence of spatial separation, hearing loss, and English language experience on speech reception thresholds. *Journal of the Acoustical Society of America*, *94*, 2291–2385.

Preminger, J., & Cunningham, D. (2003). Case-study analysis of various field study measures. *Journal of the American Academy of Audiology*, *14*, 39–55.

Preminger, J., Neuman, A., & Bakke, M. (2000). An examination of the practicality of the simplex procedure. *Ear and Hearing*, *21*, 177–193.

QuickSIN Speech in Noise Test. (2001). Elk Grove Village, IL: Etymotic Research.

Ricketts, T., & Bentler, R. (1992). Comparison of two digital programmable hearing aids. *Journal of the American Academy of Audiology*, *3*, 101–112.

Saunders, G., & Cienkowski, K. (2002). A test to measure subjective and objective speech intelligibility. *Journal of the American Academy of Audiology*, *13*, 38–49.

Saunders, G., Field, D., & Haggard, M. (1992). A clinical test battery for obscure auditory dysfunction (OAD): development, selection and use of tests. *British Journal of Audiology 26*, 33–42.

Saunders, G., Forsline, A., & Fausti, S. (2004). The Performance-Perceptual Test (PPT) and its relationship to unaided reported handicap. *Ear and Hearing*, *25*, 117–126.

Saunders, G., & Haggard, M. (1993). The influence of personality-related factors upon consultation for two different 'marginal' organic pathologies with and without reports of auditory symptomatology. *Ear and Hearing*, *14*, 242–248.

Valente, M., Fabry, D., Potts, L., & Sandlin, R. (1998). Comparing performance of the Widex Senso digital hearing aid with analog hearing aids. *Journal of the American Academy of Audiology*, *9*, 342–360.

Ventry, I., & Weinstein, B. (1982). The Hearing Handicap Inventory for the Elderly: a new tool. *Ear and Hearing*, *3*, 40–46.

Walden, T., & Walden, B. (2004). Predicting success with hearing aids in everyday living. *Journal of the American Academy of Audiology*, *15*, 342–352.

Wiley, T., Cruickshanks, K., Nondahl, D., & Tweed, T. (2000). Self-reported hearing handicap and audiometric measures in older adults. *Journal of the American Academy of Audiology*, *11*, 67–75.