Development of the Screening Test for Hearing Problems

Marilyn E. Demorest

University of Maryland, Baltimore County

David J. Wark

University of Memphis

Sue Ann Erdman

ARCCS, Jensen Beach, Florida

Abstract

Purpose: The goal of this study was to develop a brief self-assessment instrument to screen for communication problems and psychosocial adjustment to hearing impairment as part of a rehabilitative needs assessment.

Method: A pseudo-random sample of 1,000 cases was drawn from a large, heterogeneous clinical database containing audiometric data and responses to the Communication Profile for the Hearing Impaired (CPHI; Erdman & Demorest, 1998a). Item response theory (IRT) was used to derive item characteristic curves, and item selection was based primarily on item discrimination. Internal consistency, factor structure, sensitivity, and specificity of two scales, Communication and Adjustment, were evaluated in a holdout sample of 319 cases from the same database.

Results: A 9-item Communication scale and an 11-item Adjustment scale both showed satisfactory internal consistency, and the 20-item test presented a clear two-factor structure. Sensitivity and specificity functions and positive and negative predictive values indicated that the two scales could be used to identify the bottom two quartiles of the clinical population, as defined by factor scores on the CPHI.

Conclusion: The two scales of the Screening Test for Hearing Problems can be used to screen for communication and adjustment problems that warrant a comprehensive rehabilitative assessment.

Development of the Screening Test for Hearing Problems

Screening tests are an expeditious means of identifying individuals who do or do not meet a predetermined criterion or, who do or do not warrant further assessment in order to make such a determination. Pure-tone screening, as an example, is conducted to identify those who may have hearing impairment. Individuals who fail a pure-tone screening are referred for a full audiometric test battery to rule out or to confirm an actual hearing impairment (American Speech-Language-Hearing Association, 1997). Screening is cost-effective in terms of time and resource management because those who pass do not require extensive assessment thus permitting resources to be allocated efficiently towards those who require more comprehensive evaluation.

The validity of screening measures is exceedingly important because of the inherent risk of misclassification. A valid screening test accurately identifies individuals who need a diagnostic test; it does not predict the diagnosis. Therefore, screening measures must have adequate sensitivity and specificity; in other words, they must have low false-negative and false-positive rates. In order to be useful, the test must also have good positive and negative predictive value. Positive and negative predictive value refer to the ratio of true positives and true negatives to the total number of positive and negative test outcomes, respectively. Predictive value is dependent not only on the properties of the test, but also on the prevalence of the target diagnosis in the population.

Hearing screenings are routinely administered to rule out the onset of hearing impairment in at-risk populations such as individuals routinely exposed to hazardous noise levels and patients on ototoxic medications. On an even larger scale, implementation of high risk registries and the Early Hearing Detection and Intervention (EHDI) program (White, Foresman, Eichwald,

& Munoz, 2010) has resulted in increasingly effective infant screening programs (Sininger, Martinez, Eisenberg, Christensen, Grimes, & Hu, 2009). The high incidence of hearing problems associated with aging has also triggered an increase in hearing screenings as a routine part of physical exams for older people (Gates, Murphy, Rees, & Fraher, 2003; Johnson, Danhauer, Bennett, & Harrison, 2009; Johnson, Danhauer, Koch, Celani, Lopez, & Williams, 2008).

Acquired hearing impairment among adults is typically gradual in onset. Individual differences in communication needs and compensatory skills determine the extent to which hearing impairment precipitates communication and associated adjustment difficulties. Indeed, the fact that audiometric data and self-reports of communication problems are only modestly correlated has been documented repeatedly throughout the audiology literature (e.g., Chang, Ho, & Chou, 2009; Erdman & Demorest, 1998b; Hallberg & Carlsson, 1991; Hétu, Lalonde, & Getty, 1987; High, Fairbanks, & Glorig, 1964; Jupiter, 2009; Rowland, Dirks, Dubno, & Bell, 1985; Speaks, Jerger, & Trammel, 1970; Swan & Gatehouse, 1990; Weinstein & Ventry, 1983). Although the strength of the correlations varies depending on the specific variables (Erdman, 1994; Saunders & Cienkowski, 2002; Saunders, Forsline & Fausti, 2004), some conclusions can be drawn from the consistent trends in correlations:

1. Communication difficulties and associated adjustment problems are related to degree of hearing impairment.
2. The relationship between them does not permit the nature or extent of communication and adjustment difficulties to be predicted from an individual's audiogram.

3. The variability in communication and adjustment difficulties for a given hearing impairment suggests that other factors influence the extent to which hearing impairment does or does not become problematic for a given individual.
4. Assessment of hearing impairment alone is an inadequate means of assessing clients' rehabilitative needs (Erdman, 1994; Erdman & Demorest, 1998b).

It is also evident, that pure-tone screening is an inadequate means of identifying those individuals whose hearing ability is precipitating communication problems and/or impacting psychological well-being. The use of self-report measures to assess the subjective experience of hearing problems is, therefore, an essential ingredient of any protocol for adult hearing screening.

Ventry and Weinstein (1983) first advocated the use of self-report as an adjunct to pure-tone screening to identify individuals for whom hearing impairment was problematic. Shortened versions of the Hearing Handicap Inventory for the Elderly (HHIE-S) (Ventry & Weinstein, 1983) and the Hearing Handicap Inventory for Adults (HHIA-S) (Newman, Weinstein, Jacobson & Hug, 1990) were developed for screening purposes. A number of studies have demonstrated the usefulness of the HHIE-S as an adjunct to pure-tone audiometric testing in identifying elderly persons with hearing impairment (Jupiter, 2009; Jupiter & DiStasio, 1998; Koike & Johnston, 1989; Lichtenstein, Bess, & Logan, 1988a, 1988b; McBride, Mulrow, Aguilar, & Tuley, 1994; Mulrow, Tuley, & Aguilar, 1990; Weinstein, 1986). In these studies the focus was on screening for hearing impairment per se, and the individual's reports of communication and social-emotional consequences of hearing impairment were used to provide information complementary to that obtained from a pure-tone audiometric screening.

Once a diagnosis of hearing impairment has been made, the focus of assessment is on the individual's rehabilitative needs. Is the client experiencing significant difficulties with communication? How does he or she cope in difficult listening situations? What psychosocial problems are attributable to the hearing impairment? The Communication Profile for the Hearing Impaired (CPHI; Demorest & Erdman, 1987; Erdman & Demorest, 1998a) assesses these aspects of behavioral and psychosocial adjustment to hearing impairment. The 163-item instrument yields 25 scale scores in five areas: Communication Performance, Communication Importance, Communication Environment, Communication Strategies, and Personal Adjustment. The comprehensive scope of the CPHI makes it ideal for identifying the specific rehabilitative needs of individuals with hearing impairment and for assessing outcome (Erdman, 2006). However, just as it is useful to have a screening tool for the presence of hearing impairment, it would also be useful to have a tool to screen for an individual's rehabilitative needs. Failure on the screening test would indicate the need for a more comprehensive and detailed rehabilitative assessment, such as that offered by the CPHI. Rehabilitative services could then be tailored to individual's specific needs. Passing the screening test would suggest that the individual's rehabilitative needs could be addressed according to standard practice in a given clinical setting.

As a first step in developing a screening test for rehabilitative needs, we examined the factor structure of the CPHI. The factor analysis reported by Demorest and Erdman (1989) identified five factors underlying the 25 CPHI scales. The first factor, Adjustment, accounted for 29.0% of the instrument variance and comprised all of the Personal Adjustment scales of the CPHI, along with Attitudes of Others, Behaviors of Others, and Maladaptive Behaviors—the psychosocial aspects of adjustment. Low scores on Adjustment reveal an individual's negative feelings and attitudes, perceptions of negative attitudes and behaviors of others, and use of maladaptive communication strategies. Another important factor was Communication Performance, which accounted for 9.8% of the variance. Individuals who score low on Communication Performance report inability to communicate effectively in social, work, and home situations, under both ordinary and adverse listening conditions. Both Adjustment and Communication Performance are addressed by audiologists offering comprehensive rehabilitative services.

The remaining CPHI factors (Reaction, Interaction, and Communication Importance) accounted for 10.1%, 10.1% and 9.0% of the variance respectively (Demorest & Erdman, 1989). Reaction and Interaction each represent interplay between the individual and his or her environment, and, to our knowledge, have not been the focus of rehabilitative services. Communication Importance, although distinct from Communication Performance, provides additional information to the audiologist for interpreting the Communication Performance scales, but is not itself a target for intervention. For these reasons, we did not elect to screen for these additional factors explicitly.

The Screening Test for Hearing Problems (STHP) was designed for rehabilitative screening: (a) to identify individuals who warrant a comprehensive evaluation of communication and adjustment problems secondary to hearing impairment and (b) to thereby bring efficiency to the process of diagnosing an individual's rehabilitative needs. The STHP was derived from the CPHI, using items that best represent the CPHI factors of Communication Performance and Adjustment. An advantage of using the CPHI as the basis for the screening test is that items can be selected to systematically represent the various facets of communication and adjustment that are represented by the CPHI scales. The STHP consists of two scales, Communication and Adjustment, which are scored separately, but which must both be passed in order to pass the screening; failure on either scale suggests that a more comprehensive evaluation of rehabilitation needs is indicated.

Method

Participants

Erdman and Demorest (1998a) described a multicenter clinical study of adjustment to hearing impairment and reported normative data on the CPHI for a heterogeneous clinical database of 1,008 clients. Five clinical centers participated: Callier Center for Communication Disorders, University of Texas at Dallas; Department of Communication Disorders, Manhattan Eye, Ear and Throat Hospital; San Francisco Hearing and Speech Center; Memphis Speech and Hearing Center, University of Memphis; and Department of Otolaryngology, Washington University Medical Center. These centers incorporated the CPHI into their standard clinical protocol and no specific selection criteria were imposed other than ability to complete the audiological tests and the CPHI. The goal was to obtain a representative sample of the clinical population of each center.

> *All* adults who were seen for a comprehensive audiological assessment were considered eligible for inclusion in the study. It can therefore be assumed that there is natural variability within each sample of factors such as referral source, etiology of hearing loss, hearing aid use, and so on. Emphasis was placed on avoidance of selection biases by the clinical staff that might compromise the representativeness of the sample. (Erdman & Demorest, 1998a, p. 110)

The design and procedures of the clinical study were approved by the Institutional Review Board at the University of Maryland, Baltimore County and at each of the four clinical centers that had such a board. The clinical protocol was specified as follows:

> (a) air and bone conduction thresholds (re: ANSI S3.21-1978 [R-1986]); (b) speech reception or speech awareness thresholds; (c) word recognition scores based on recorded NU-6 word lists; (d) immittance testing; (e) results of any special tests performed for diagnostic purposes (e.g., ABR, ENG, etc.); (f) results of hearing aid evaluations (if relevant) including real ear gain measurements and description of the amplification system; and (g) history of aural rehabilitation. (Erdman & Demorest, 1998a, p. 110)

Because data collection continued after the database was frozen for the Erdman and Demorest (1998a) study, an additional 311 cases were available for the present project,[1] for a total sample of *N* = 1,319. A pseudo-random subsample of 1,000 cases was selected to serve as the item selection sample, and the remaining 319 cases were held out for subsequent evaluation of sensitivity and specificity of the selected items. The two subsamples were equivalent[2] with respect to all demographic and audiometric variables tested and had comparable scores on all CPHI scales.

The total sample was 55.8% male, and the average age was 65.0 years (*SD* = 15.4; range, 16 - 97). The racial/ethnic composition of the sample was: White, 82.4%; Black, 10.4%; Hispanic, 3.1%; Asian, 3.1%; and Native American, 0.4%. Two-thirds of the sample (66.7%) had at least some college education. The mean pure-tone audiogram of the group is shown in Table 1. Better-ear Speech Recognition Threshold averaged 30.3 dB HL (*SD* = 18.1; range, 0 - 110), and better-ear word recognition for recorded tests[3] averaged 84.5% (*SD* = 16.4; range = 4 - 100). More than half the sample (56.1%) reported no prior hearing-aid use.

Item Scoring

Items of the CPHI are scored on a 5-point scale. However, in a screening test, individual test items function as “symptoms.” It is more important to know whether or not a given problem is present and less important to know exactly how severe the problem is. Scoring each item dichotomously (i.e., pass-fail) yields a total score that is indicative of the number of areas in which difficulties are experienced. It also reduces the possibility of a high score on one item compensating for a very low score on another item. The first step in construction of the STHP was to map responses to CPHI items from a 5-point response scale to a dichotomous pass-fail score.

For Communication Performance items, a passing score was defined as a response of 4 or 5. That is, the respondent had to indicate that he or she could communicate effectively 4 (*Frequently*) or 5 (*Usually, Almost Always*). A failing score resulted when the individual indicated that he or she could only communicate effectively 3 (*About Half the Time*), 2 (*Occasionally, Sometimes*), or 1 (*Rarely, Almost Never*).

All items assessing Adjustment to Hearing Impairment describe difficulties or problems experienced by the individual, and hence they are reversed for scoring. A passing score was defined as a response of 1 or 2. For some items, the response scale is a frequency scale with options of 1 (*Rarely, Almost Never*) and 2 (*Occasionally, Sometimes*). For other items, the response scale is a scale of agreement, with a passing score awarded for responses of 1 (*Strongly Disagree*) and 2 (*Disagree*) and a failing score assigned for responses of 3 (*Uncertain*), 4 (*Agree*), and 5 (*Strongly Agree*).

Item Selection

Item response theory (IRT; Hambleton, Swaminathan, & Rogers, 1991) was used to select items for inclusion in the STHP. As described more fully in the Appendix, item-characteristic curves (ICCs) are plotted showing the probability of passing the item as a function of the total score on the scale. The ICC curve is a logistic function with two parameters:

difficulty and discrimination. Difficulty is related to the overall pass rate for the item, and discrimination is related to how well the item discriminates between those who pass and fail on a criterion measure. Items can be compared with respect to their difficulty and discrimination and selected for inclusion in a test on that basis.

Two screening scales were derived, one for the Communication factor of the CPHI and one for the factor of Adjustment to hearing impairment. For Communication, an item-characteristic curve (see Appendix) was estimated for each of the 18 Communication Performance items of the CPHI using XCALIBRE™ (Version 1.10e). Separate analyses were performed for Average and Adverse situations, with the total score across all items in each scale serving as the criterion. For Adjustment, item-characteristic curves were estimated separately for the items within each scale contributing to the CPHI's Adjustment factor. For example, the eight items that comprise the Self-Acceptance scale were analyzed using the total score on those items as the criterion. One item was then selected from each scale. This ensured that the items of the screening scale would represent all aspects of adjustment included in the Adjustment factor of the CPHI.

The primary selection criterion for items was high item discrimination, but the difficulty of passing an item was also taken into account. Because many items were acceptable in terms of these two item characteristics, it was possible to use additional criteria for item selection, such as item content and wording.

Results

Items Selected

Descriptive statistics and estimated discrimination and difficulty parameters for the items selected are presented in Table 2. The nine Communication items are balanced with respect to the Social, Work, and Home scales of the CPHI (three items each) and also with respect to the the CPHI scales for Average and Adverse conditions (five and four items, respectively). The Adjustment scale contains 11 items, one from each CPHI scale that contributes significantly to the Adjustment factor.

Item discrimination (parameter *a*) values, range from 1.10 to 1.80, with a mean of 1.41. Most items have difficulty (parameter *b*) values near zero. That is, the items tend to discriminate best among those patients who score near the mean of the clinical population.

Internal Consistency

The two scales of the Screening Test for Hearing Problems were evaluated for internal consistency in the holdout sample using coefficient alpha. Cases with missing data were excluded from the analysis. For the Communication scale, $\alpha = .842$ ($n = 284$), and for the Adjustment scale, $\alpha = .770$ ($n = 299$). These values are quite satisfactory, but they are lower than the values reported for many scales of the CPHI (Erdman & Demorest, 1998a) for two reasons. First, dichotomous scoring of the items tends to reduce the interitem correlations upon which coefficient α is based. Second, the scales of the screening test are, by design, more heterogeneous in content than scales of the CPHI.

Factor Structure

To verify that the screening test measures the two factors, Communication and Adjustment, principal factors with varimax rotation were derived using TESTFACT for Windows (version 4.0). Because the items were dichotomously scored, the interitem correlations were first converted to tetrachoric correlations. The eigenvalues for unrotated Factor 1 and Factor 2 were 8.06 and 3.26, respectively. All remaining eigenvalues were less than 1.50. A scree plot of the eigenvalues clearly indicated a two-factor structure. The factor loadings for the two principal factors, after rotation, are shown in Table 3. As expected, the two factors represent the Communication and Adjustment scales, respectively.

Sensitivity and Specificity

Sensitivity is the probability that a test correctly identifies those *with* a particular characteristic (true positive rate or hit rate); specificity is the probability that a test correctly identifies those *without* that characteristic (true negative rate). Sensitivity and specificity are inversely related, and the appropriate balance between them depends on the consequences of the two corresponding types of screening errors: misses and false positives. In the absence of compelling reasons to prefer one type of error over the other, it is desirable to make sensitivity and specificity approximately equal to one another.

To evaluate sensitivity and specificity it is necessary to have a dichotomous (or dichotomized) criterion variable and a cutoff score that defines the screening test outcome as pass or fail. For the Communication scale, the criterion variable was the individual's score on the Communication Performance factor of the CPHI; for the Adjustment scale, the criterion variable was the score on the (psychosocial) Adjustment factor of the CPHI. Factor scores were calculated for each participant in the holdout sample using the factor-score coefficients reported by Demorest and Erdman (1989).

When tests are used to screen for hearing impairment, considerable consensus has been achieved concerning the definition of the criterion as present or absent (American Speech-Language-Hearing Association, 1997). There is no such consensus in the audiologic rehabilitative literature. In clinical practice, it may be desirable to have more than one definition of pass/fail on the criterion variable, depending on the severity of problems (and hence the percentage of the clinical population) that one wishes to identify or that one has the resources to serve. Accordingly, we used two norm-based approaches to dichotomize the criterion variables, one that defined the bottom quartile (i.e., Quartile 1, the lowest 25%) of the population as the screening target and one that defined Quartile 1 and Quartile 2 (i.e., the lower 50%) of the population as the screening target. Sensitivity and specificity functions were generated by evaluating sensitivity and specificity for each possible cutoff score on the corresponding screening scale. The two pairs of functions that resulted are shown in Figure 1.

Because the target characteristics for these screening scales involve hearing difficulties and problems in psychosocial adjustment, failure on the screening test means that the individual is identified as having those problems. When the cutoff score for passing is set at 0, all individuals pass the screening test, and hence no one is identified as having problems. This results in sensitivity of 0. Specificity equals 1, however, because there are also no false positives. As the cutoff is raised, sensitivity increases and specificity decreases.

The four functions shown in Figure 1 indicate excellent sensitivity and very good specificity for these screening scales in the holdout sample. For each possible cutoff score, there is a corresponding value of sensitivity and specificity. Selection of a cutoff involves decisions about the relative importance of these two outcomes. Because we are most interested in identifying individuals who are experiencing significant difficulties with communication and personal adjustment, we have elected to place the cutoffs so that a slight preference is given to sensitivity over specificity.

The cutoff scores shown in Figure 1 were selected with these considerations in mind. The goal was to obtain sensitivity greater than or equal to 90% and then to maximize specificity. When screening for the bottom 25% of the clinical population, the resulting cutoffs yield high values of sensitivity and moderately high values of specificity. For Communication (Figure 1a), a cutoff score of 3 yields sensitivity of 91.4% and specificity of 77.3%; for Adjustment (Figure 1c), a cutoff score of 5 yields sensitivity of 93.2% and specificity of 77.7%. Screening for the bottom 50% of the clinical population yields similar results. For Communication (Figure 1b), a cutoff score of 4 produces sensitivity of 93.3% and specificity of 85.9%; for Adjustment (Figure 1d), a cutoff score of 7 produces sensitivity of 93.4% and specificity of 68.7%.

Predictive Value

Because the target for the STHP is a fixed percentage of the population (i.e., the bottom 25% or 50%), the prevalence of problems with communication and psychosocial adjustment is known and does not have to be estimated. Estimates of positive and negative predictive value can be obtained directly from the corresponding ratios in the sample data. Screening for the bottom 25% of the clinical population and using the cutoffs illustrated above, we estimate positive predictive values of 61.7% for Communication and 58.0% for Adjustment. The estimates for negative predictive value are 95.7% and 97.2%, respectively. If the target is the bottom 50% of the population, the positive predictive values are 88.0% and 75.4%; the negative predictive values are 92.1% and 91.0%.

Discussion

The two screening scales developed in this study make it possible, and clinically feasible, to screen for communication and adjustment problems in a clinical population of adults with hearing impairment. An individual must pass both the Communication and Adjustment screening scales in order to pass the test. Failure on either scale constitutes failure on the screening test as a whole and indicates that a more comprehensive assessment of rehabilitative needs is warranted. The values of sensitivity, specificity, and positive and negative predictive value reported here are high enough to support the validity of the STHP as a rehabilitative screening tool.

Because the Communication and Adjustment factors of the CPHI are relatively independent (*r* = .22), it is important to screen for both. Table 4 shows the numbers of individuals in the holdout sample who fell in each quartile on each criterion variable. Eighty-seven participants fell in the bottom quartile on Communication only, 79 fell in the bottom quartile on Adjustment only, and 29 fell in the bottom quartile on both factors. These data suggest that if the target is the lowest quartile on each factor, then about 43.4% of the clinical population will be targeted for further assessment. If the target is the bottom half of the clinical population on each factor, then 71.5% of the population would be targeted. These values suggest that the screening test can be used most efficiently when it is used to target the bottom 25% of the clinical population on each factor. If it is considered important to identify individuals with less severe communication and psychosocial adjustment problems, then it is more efficient to provide all clients with a comprehensive rehabilitative assessment, such as the CPHI, than to engage in a two-stage procedure that will ultimately result in conducting the more complete assessment on a substantial majority of those who are screened.

Use of a holdout sample in the present study has permitted us to obtain estimates of internal consistency, factor structure, sensitivity, and specificity that are independent of the sample upon which item selection was based. In that sense, the results may be considered as cross-validated. However, there is a more important aspect of cross-validation that was not part of the present study. All participants in this study were administered the full CPHI. There is no guarantee that the item characteristics exhibited in the context of a 163-item questionnaire will be invariant when only the 20 items of the screening test are administered. A more rigorous and more realistic evaluation of the Screening Test for Hearing Problems requires that it be administered clinically, by itself, and that subsequent testing with the CPHI, preferably on another occasion, provide independently assessed factor scores.

It is important to note that the Screening Test for Hearing Problems is *not* to be considered a “short form” of the CPHI. The two instruments have different purposes and are scored in different ways. The CPHI is designed to provide a comprehensive assessment of adjustment to hearing impairment. The profile of scores that it provides permits a detailed description of the particular areas in which individuals are experiencing difficulties, and the individual scales within the profile have each been evaluated psychometrically. Each item is scored on a five-point scale and the resulting scale scores may be considered reliable and valid quantitative measures. The screening test, in contrast, is only concerned with two factors that are measured by the CPHI. In the screening test, each item is scored on a pass-fail basis, and each of the two scales is likewise scored as pass/fail. The recommended cutoff scores are those that balance sensitivity and specificity, with a slight preference for sensitivity over specificity. The focus in the screening test is to determine whether there is sufficient evidence of difficulties in either area to warrant an in-depth comprehensive assessment.

Another important caveat concerns the interpretation of passing a screening test for hearing problems. Although failing the test justifies and suggests the need for further evaluation of specific rehabilitative needs, passing the test does not imply that further assessment or intervention is not warranted. Indeed, a legitimate, if not critical aspect of rehabilitation is the prevention of communication and adjustment difficulties that can develop quite insidiously as communication needs, environmental demands, or personal circumstances change (Erdman, 2000).

The STHP is included as a supplement to this article in a format suitable for reproduction and administration. The test has a number of potential clinical uses. For example, it could be used in conjunction with a screening test for hearing impairment per se, providing a functional perspective to complement that provided by pure-tone screening. It could also be used with patients who do not have a hearing impairment or who are not candidates for a hearing aid, to identify those who nevertheless report significant hearing problems and who might benefit from a more comprehensive assessment and referral for rehabilitative services. Finally, as was the original intent, it could be used with individuals who have a documented hearing impairment to determine whether they may be candidates for a more targeted intervention or prevention program than is generally offered in a given setting. Although there is no conceptual reason why a comprehensive test such as the CPHI could not be routinely used for this purpose, practical considerations frequently result in no such evaluation at all. Administering the STHP requires a minimal investment of both client and clinician time, and if the individual fails either scale, and therefore the test as a whole, the clinician would have clear evidence that further assessment and evaluation are indicated.

Acknowledgements

The original clinical study was supported by a grant from the National Institutes of Health, R01DC01091. The authors gratefully acknowledge the following collaborators and the centers with which they were affiliated for their contributions to the clinical database on which the present study is based: Robert D. Madory, San Francisco Hearing and Speech Center; Joseph J. Montano, Manhattan Eye and Ear Hospital; the late Margaret W. Skinner, Washington University, St. Louis; and P. Lee Wilson, Callier Center for Communication Disorders, University of Texas at Dallas.

References

American National Standards Institute. (1978). *Methods for manual pure-tone threshold audiometry* (ANSI S3.21 [R-1986]). New York: ANSI.

American Speech-Language-Hearing Association. (1997). *Guidelines for Audiologic Screening* [Guidelines]. Available from www.asha.org/policy. Index terms: screening. doi: 10.1044/policy.GL1997-00199

Chang, H., Ho, C., & Chou, P. (2009). The factors associated with a self-perceived hearing handicap in elderly people with hearing impairment--results from a community-based study. *Ear and Hearing, 30*, 576-583.

de Ayala, R. J. (2008). *The theory and practice of item response theory*. New York: Guilford Press.

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

Demorest, M. E., & Erdman, S. A. (1989). Factor structure of the Communication Profile for the Hearing Impaired. *Journal of Speech and Hearing Disorders, 54*, 541-549.

Erdman, S. A. (1993). Self-assessment in audiology: The clinical rationale. *Seminars in Hearing*, *14*, 303-313.

Erdman, S. A. (1994). Self-assessment: From research focus to research tool. In J.-P. Gagné & N. Tye-Murray (Eds.), *Research in audiological rehabilitation: Current trends and future directions*, *Journal of the Academy of Rehabilitative Audiology*, *27* (Mongr. supp.), 67-90.

Erdman, S. A. (2000). Counseling adults with hearing impairment. In J. G. Alpiner & P. A. McCarthy (Eds.), *Rehabilitative Audiology: Children and Adults* (3rd ed., pp. 435-470). Baltimore: Lippincott Williams & Wilkins.

Erdman, S. A. (2006). Clinical interpretation of the CPHI. *Perspectives on Aural Rehabilitation and Its Instrumentation, 13*(1), 3-18. doi: 10.1044/arii13.1.3

Erdman, S. A., & Demorest, M. E. (1998a). Adjustment to hearing impairment I: Description of a heterogeneous clinical population. *Journal of Speech, Language, and Hearing Research, 41*, 107-122.

Erdman, S. A., & Demorest, M. E. (1998b). Adjustment to hearing impairment II: Audiological and demographic correlates. *Journal of Speech, Language, and Hearing Research, 41,* 123-136.

Gates, G. A., Murphy, M., Rees, T. S., & Fraher, A. (2003). Screening for handicapping hearing loss in the elderly. *The Journal of Family Practice, 52*, 56–62.

Hallberg, L. R.-M., & Carlsson, S. G. (1991). Hearing impairment, coping and perceived hearing handicap in middle-aged subjects with acquired hearing loss. *British Journal of Audiology, 25,* 323–330.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Hambleton, R. K., & Swaminathan, H. (2009). *Item response theory: Principles and applications.* New York: Springer-Verlag.

Hétu, R., Lalonde, M., & Getty, L. (1987). Psychosocial disadvantages associated with occupational hearing loss as experienced in the family. *Audiology, 26,* 141–152.

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

Johnson, C. E., Danhauer, J. L., Bennett, M., & Harrison, J. (2009). Systematic review of physicians' knowledge of, participation in, and attitudes toward hearing and balance screening in the elderly population. *Seminars in Hearing, 30*, 193-206.

Johnson, C. E., Danhauer, J. L., Koch, L. L., Celani, K. E., Lopez, I. P., & Williams, V. A. (2008). Hearing and balance screening and referrals for Medicare patients: A national survey of primary care physicians. *Journal of the American Academy of Audiology, 19*, 171-190.

Jupiter, T. (2009). Screening for hearing loss in the elderly using distortion product otoacoustic emissions, pure tones, and a self-assessment tool. *American Journal of Audiology*, *18,* 99-107.

Jupiter, T., & DiStasio, D. (1998). An evaluation of the HHIE-S as a screening tool for the elderly homebound population. *Journal of the Academy of Rehabilitative Audiology, 31,* 11-21.

Koike, K. J. M., & Johnston, A. P. (1989). Follow-up survey of the elderly who failed a hearing screening protocol. *Ear and Hearing*, *10*, 250-253.

Kramer, S. E., Kapteyn, T. S., Festen, J. M., & Tobi, H. (1996). The relationships between self-reported hearing disability and measures of auditory disability. *Audiology, 35,* 277–287.

Lichtenstein, M. J., Bess, F. H., & Logan, S. A. (1988a). Diagnostic Performance of the Hearing Handicap Inventory for the Elderly (Screening Version) against differing definitions of hearing loss. *Ear and Hearing*, *9*, 208-211.

Lichtenstein, M. J., Bess, F. H., & Logan, S. A. (1988b). Validation of screening tools for identifying hearing-impaired elderly in primary care. *Journal of the American Medical Association, 259*, 2875–2878.

McBride, W. S., Mulrow, C. D., Aguilar, C., & Tuley, M. R. (1994). Methods for screening for hearing loss in older adults. *The American Journal of the Medical Sciences*, *307*, 40–42.

Mulrow, C. D., Tuley, M. R., & Aguilar, C. (1990). Discriminating and responsiveness abilities of two hearing handicap scales. *Ear and Hearing, 11*, 176–180.

Rowland, J. P., Dirks, D. D., Dubno, J. R., & Bell, T. S. (1985). Comparison of speech recognition-in-noise and subjective communication assessment. *Ear and Hearing, 6,* 291–296.

Saunders, G. H., & Cienkowski, K. M. (2002). A test to measure subjective and objective speech intelligibility. *Journal of the American Academy of Audiology, 13*(1), 38.

Saunders, G. H., Forsline, A., & Fausti, S. A. (2004). The performance-perceptual test and its relationship to unaided reported handicap. *Ear and Hearing*, *25*(2), 117-126.

Schow, R. L. (1991). Considerations in selecting and validating an adult/elderly hearing screening protocol. *Ear and Hearing*, *12*, 337–348.

Sininger, Y. S., Martinez, A., Eisenberg, L., Christensen, E., Grimes, A., & Hu, J. (2009). Newborn hearing screening speeds diagnosis and access to intervention by 20-25 months. *Journal of the American Academy of Audiology*, *20*, 49-57.

Speaks. C., Jerger, J., & Trammell, J. (1970). Measurement of hearing handicap. *Journal of Speech and Hearing Research*, *13*, 768-776.

SPSS Statistics (Version 17.0). (2008). Chicago: SPSS, Inc.

Swan, I. R., & Gatehouse, S. (1990). *British Journal of Audiology*, *24*, 155-60.

TESTFACT for Windows (Version 4.0) [Computer software]. Lincolnwood, IL: Scientific Software International.

Ventry, I. M., & Weinstein, B. E. (1983). Identification of elderly people with hearing problems. *Asha*, *25*(7), 37–42.

Weinstein, B. E. (1986). Validity of a screening protocol identifying elderly people with hearing problems. *Asha, 28*(5), 41–45.

Weinstein, B. E., & Ventry, I. M. (1983). Audiometric correlates of the Hearing Handicap Inventory for the Elderly. *Journal of Speech and Hearing Disorders, 48,* 379–384.

White, K. R., Forsman, I., Eichwald, J., & Munoz, K. (2010). The evolution of early hearing detection and intervention programs in the United States. *Seminars in Perinatology, 34*, 170-179.

XCALIBRE^{TM} for Windows (Version 1.10e) [Computer software]. St. Paul, MN: Assessment Systems Corporation.

Appendix

Item response theory (IRT; Hambleton, Swaminathan, & Rogers, 1991) is a mathematical formulation of the relationship between a single test item and the latent trait (θ) the item is designed to measure. IRT has several advantages over classical true score theory. Of particular importance are that estimates of the latent trait are not test-dependent and estimates of item characteristics are not population-dependent. IRT also does not assume that measurement error is the same for all individuals, but instead provides different estimates of measurement error at different values of the latent trait. Once the characteristics of individual items have been estimated, it is possible to construct tests with predictable characteristics.

For dichotomously scored items, the probability of passing an item is modeled as a function of two item parameters,[4] difficulty and discrimination. Although IRT was developed to measure abilities, where it is natural to speak of "passing" an item and of item "difficulty," it can also be applied to dichotomously scored items in other domains. In the present context, "passing" an item means giving a response that is *not* indicative of problems with communication or psychosocial adjustment.

The probability of passing an Item i (p_i) is usually modeled as a logistic function of θ with two parameters, b_i, item difficulty, and a_i, item discrimination:

$$p = \frac{1}{1 + e^{-Da(\theta - b)}}$$

D is a scaling constant that is arbitrarily set equal to 1.7.[5] It is assumed that the set of items used to measure the latent trait is *unidimensional*, that is, that only one latent trait is affecting item responses. It is also assumed that test items exhibit *local independence*: when the latent trait is held constant, responses to different items are statistically independent. This means that the joint probability of passing any two items is the product of the probabilities of passing each item. (Items that are statistically independent when θ is held constant may, however, be positively correlated when θ is allowed to vary.)

The graph of p for a particular item is called an *item characteristic curve*. Figure 2 shows an item characteristic curve for $a = 1.270$, $b = -.178$. Values of the latent trait are scaled in standard normal units and are therefore expected to range from about -3.0 to +3.0, although in principle the range is $-\infty$ to $+\infty$. The parameter b is the value of θ for which $p = .5$. The difficulty parameter therefore affects the location of the curve on the latent trait continuum. When b is negative most of the curve lies to the left of zero, and the item is relatively easy. That is, relatively low values of the latent trait are needed to pass the item. When b is positive, most of the curve lies to the right of zero, and the item is relatively difficult. Values of b usually range between -2.0 and +2.0. The parameter a represents item discrimination because a affects the steepness of the curve: it is proportional to the slope of the curve at $\theta = b$. A low positive value of a produces a curve that rises slowly as θ increases, whereas a high value of a, such as that shown for Item 28 in Figure 3, yields a steeply rising curve. Values of a are expected to range between 0 and 2.

When item response theory is used to develop a scale, it is necessary to estimate the a and b parameters for each item. As in classical item analysis, it is customary to use the total score on all items as an estimate of the latent trait. Items can then be selected based on their statistical characteristics. To illustrate, consider the two empirically estimated item characteristic curves shown in Figure 3 for Item 23 and Item 28 of the Attitudes of Others scale of the CPHI. Both items exhibit good discrimination, but Item 28 is the more discriminating item of the two. It is also a slightly “easier” item. Item selection could be based on either or both of the item parameters, but item discrimination is usually given more weight, provided that the value of *b* is not inappropriate.

Author Note

Marilyn E. Demorest, Department of Psychology, University of Maryland, Baltimore County; David J. Wark, School of Communication Sciences and Disorders, University of Memphis; Sue Ann Erdman, ARCCS, Jensen Beach, Florida.

Correspondence concerning this article should be addressed to Marilyn E. Demorest, Department of Psychology, UMBC, 1000 Hilltop Circle, Baltimore, MD, 21250. E-mail: demorest@umbc.edu

Footnotes

[1]For consistency with the analyses reported by Erdman and Demorest (1998a), no cochlear implant patients/candidates were included among the additional 311 cases.

[2]Although a Bonferroni correction would have been appropriate because of the large number of tests performed and the high power associated with large sample size, this was not necessary. Only one of the 25 statistical tests for the individual scales met a conventional criterion of $\alpha = .05$, a result easily attributable to Type I error.

[3]Of the 999 cases for which the speech test was specified on the audiogram or in the clinician's report, a substantial majority of the word recognition scores (n = 910; 91.1%) were obtained using the NU-6 monosyllabic word lists, as specified in the study protocol. CID W-22 lists were administered to 71 patients (7.1%), and coding on the audiogram was not interpretable for the remaining 18 cases (1.8%)

[4]A three-parameter model is used when the minimum probability of passing an item is greater than zero.

[5]This value of D yields a logistic function that is virtually identical to the normal ogive, an alternative model that is conceptually appropriate but mathematically less desirable.

Table 1

*Pure-Tone Thresholds(dB HL[a]) for the Total Sample (*N *= 1,048*)

	Frequency (Hz)					
Ear	250	500	1000	2000	4000	8000
Left						
Mean	28.5	30.2	35.6	45.0	57.7	67.2
SD	18.4	19.0	19.6	19.4	20.1	22.2
Minimum	0	0	0	-5	-5	-5
Maximum	105	10	105	110	115	115
Right						
Mean	28.4	29.9	35.2	42.7	56.1	66.7
SD	17.8	18.1	19.2	19.6	20.9	23.2
Minimum	0	-5	0	0	-5	0
Maximum	90	95	105	105	120	110

Note. Sample size was reduced because of missing data at some frequencies, particularly at 8000 Hz. Excluding 8000 Hz would increase the sample size to 1,259 and raise thresholds 4-5 dB at each frequency.

[a]Re: ANSI S3.21-1978 [R-1986].

Table 2

*Descriptive Statistics and Item Parameters for Selected Items in the Derivation Sample (*N *= 1,000)*

CPHI Item Number	*n*	Mean[a]	SD	Discrimination[b]	Difficulty[c]
			Communication		
3	987	.66	.47	1.33	-0.54
5	989	.43	.50	1.25	0.22
6	957	.57	.50	1.18	-0.25
8	988	.25	.43	1.51	0.88
10	969	.63	.48	1.76	-0.41
12	989	.33	.47	1.54	0.54
14	987	.39	.49	1.37	0.39
15	987	.19	.39	1.25	1.21
16	994	.24	.43	1.46	0.94
			Adjustment		
28	977	.55	.50	1.30	-0.17
38	990	.64	.48	1.10	-0.56
60	980	.67	.47	1.20	-0.62
70	991	.60	.49	1.46	-0.34
92	988	.35	.48	1.31	0.53
103	989	.39	.49	1.29	0.39
108	992	.70	.46	1.63	-0.67
112	993	.54	.50	1.80	-0.12
129	987	.40	.49	1.32	0.38
131	993	.36	.48	1.48	0.48
139	992	.42	.49	1.69	0.27

[a]Items are dichotomously scored, so the item mean equals the proportion of the sample that passed the item. [b]Discrimination is an item parameter (*a*) that is proportional to the slope of the item characteristic curve when the probability of passing the item is .5. [c]Difficulty is an item parameter (*b*) that equals the value of the latent trait for which the probability of passing the item is .5.

Table 3

*Factor Structure of the Screening Test for Hearing Problems in the Holdout Sample (*n = *271*)

Item		Communication	Psychosocial Adjustment
1	You're at the dinner table with your family.		.47
2	You're at a restaurant ordering food or drinks.		.48
3	You're talking on the telephone when you're at work or a place of business		.50
4	Someone's talking to you while you're watching TV or listening to the stereo/radio.		.47
5	You're talking with someone in an office.		.48
6	You're at a dinner party with several other people.		.35
7	You're at a meeting with several other people.		.39
8	You're at home and someone is talking to you from another room.		.48
9	You're having a conversation at a social gathering while others are talking nearby.		.46
10	I tend to avoid social situations where I think I'll have problems hearing.	.43	
11	People don't speak clearly enough when they're speaking to me.	.43	
12	I feel foolish when I misunderstand what someone has said.	.41	
13	People think I'm not paying attention if I don't answer them when they speak to me.	.48	
14	When I can't understand what's being said, I feel tense and anxious.	.47	
15	I hate to ask others for special consideration just because I have a hearing problem.	.48	
16	I get discouraged because of my hearing loss.	.45	
17	I find it difficult to admit to others that I have a hearing problem.	.46	
18	I get angry when others don't speak up.	.37	
19	Sometimes I miss so much of what's being said that I feel left out.	.43	
20	Others should be more understanding about my hearing problems.	.44	

Note. Factor loadings less than .30 are not printed.

Table 4

Crosstabulation of Participants in the Holdout Sample by Quartile on the Communication and Adjustment Factors of the CPHI

Communication Quartile	Adjustment Quartile				
	1st	2nd	3rd	4th	Total
1st	29	23	18	17	87
2nd	20	25	19	14	78
3rd	20	15	22	16	73
4th	10	16	21	31	78
Total	79	79	80	78	316

Figure Captions

Figure 1. Sensitivity and specificity functions in the holdout sample for the Communication and Adjustment scales using two screening criteria: detection of the lowest 25% of the population (2a and 2c) and detection of the lowest 50% of the population (2b and 2d). Results for Communication are shown in 2a and 2b; results for Adjustment are shown in 2c and 2d.

Figure 2. Item characteristic curve for a hypothetical item with difficulty (*b*) equal to -0.178 and discrimination (*a*) equal to 1.270.

Figure 3. Item characteristic curves for two items from the CPHI Attitudes of Others scale.