

RELATIVE EFFICIENCY OF TESTS USED TO DETECT FUNCTIONAL HEARING LOSS

Ira M. Ventry.

Introduction

This paper has two purposes. First, to describe briefly some of the goals and procedures of a large scale multidiscipline investigation of functional hearing loss. The second purpose is to present some preliminary experimental findings on the relative efficiency of certain tests and test relationships used to detect functional hearing loss in our study. There appear to be no published research data on the relative efficiency of the tests and test relationships that I will describe later. Consequently, our findings should be of considerable value to clinicians confronted with the problem of which tests to use in the identification of patients with a functional hearing loss. It should be emphasized that while much of this paper has an audiology focus, the general plan of the study is the result of a collaborative effort between audiology, psychology, psychiatry, otorhinolaryngology, and neurology.

General Plan of the Investigation

In April, 1959, a United States Public Health Service grant was awarded to support a multidiscipline research project titled „An Investigation of Functional Hearing Loss“. The basic plan of this study involves intensive evaluation of an Experimental Group of at least 60 subjects who present a functional hearing problem and a Control Group of at least 30 hard-of-hearing subjects without a functional component to their hearing problem. Both groups are randomly selected from the same population of male veterans referred to the San Francisco Veterans Administration Hospital for compensation examinations. All subjects are evaluated by audiology, otorhinolaryngology, neurology, psychiatry, and psychology. All subjects also receive chest X-Rays, mastoid X-Rays, routine EEGs, blood studies, urinalysis, and a general medical examination.

The following general goals define the scope and focus of the study: (1) To evaluate the efficiency of various procedures currently used to establish the diagnosis of functional hearing loss. (2) To determine whether persons with functional hearing loss demonstrate characteristics which differentiate them from individuals having only an organic hearing loss. (3) To determine what differences exist between persons who quickly resolve their functional hearing loss and those in whom the functional loss persists. (4) To develop from the empirical data of the study a theoretical frame of reference that will facilitate the diagnosis and treatment of functional hearing loss and lead to further research.

For the purposes of the study, functional hearing loss has been defined operationally as the presence of significant intratest or intertest audiometric discrepancies without medical findings adequate to explain the discrepancies. Clearly specified audiometric criteria are used to select the Experimental Group and the Control Group. The strict audiometric criteria were designed to establish clearly defined groups free of the confounding effects of group overlap, thus eliminating one of the most important deficiencies of previous research on functional hearing loss.

Audiologic Tests and Criteria Used in Subject Selection

Experimental Group subjects are selected on the basis of the results of pure-tone audiometry, speech audiometry, conditioned GSR pure-tone audiometry, pure-tone Stenger tests, and the speech Stenger test. I should point out that routinely we do a pure-tone Stenger test whenever there is a threshold difference of 20 db or more between ears for any of the three speech frequencies. A speech Stenger test is performed whenever there is an interaural difference of 20 db or greater for speech reception thresholds. In selecting Experimental Group subjects the following audiometric criteria are used: (1) a test-retest difference of at least 15 db for voluntary thresholds at 1000 cps; (2) a difference of ± 12 db or more between the appropriate pure-tone average and speech reception threshold in the same ear; (3) a conditioned GSR threshold at 1000 cps at least 15 db lower than any voluntary threshold at 1000 cps; (4) a pure-tone Stenger test positive 15 db or more below voluntary threshold at any frequency tested; and (5) a speech Stenger test positive 15 db or more below the voluntary speech reception threshold in the reportedly poorer ear. Subjects in the Control Group must not meet any of these criteria, while Experimental Group subjects must have positive findings on GSR or on 2 or more of the other 4 criteria.

Results

The following analyses are based on an Experimental Group N of 47. These subjects had a functional component at the time of their initial testing. However, in evaluating the Stenger tests, the N is reduced to 19 since only 19 subjects provided data for evaluating both Stenger criteria in addition to providing data for evaluating the other criteria.

The most efficient measure in detecting functional hearing loss was the speech reception threshold-pure-tone average relationship, which was positive in 79 % of the Experimental Group subjects. Second in efficiency was GSR audiometry which correctly identified 74 % of the Experimental Group. The third most effective measure was the pure-tone test-retest relationship, positive in 66 % of the cases. The Stenger procedures were the least efficient, with the pure-tone Stenger test positive in 53 % of the cases in which it was applied and the speech Stenger positive in only 26 % of the cases. The difference in efficiency among these 5 measures is statistically significant. However, this overall statistical difference is a function of the low efficiency of the speech Stenger. When the speech Stenger test is omitted from the analysis, there is no significant difference in efficiency among the other 4 measures.

Another way to evaluate efficiency is in terms of false-negative results,

that is, failure of a test to identify a subject as functional. GSR audiometry produced no false-negative classifications; the pure-tone test-retest relationship was negative in 11 % of the cases; and the speech reception threshold-pure-tone average relationship had a false-negative rate of 17 %. The pure-tone and speech Stenger tests produced the highest false-negative rates, 42 % and 63 % respectively.

With the exception of the speech Stenger test, our experimental results are generally consistent with our clinical observations. Despite significant differences in procedure and criteria, our results are in relatively good agreement with the incidental findings on test efficiency reported by Menzel in 1960. The high efficiency of GSR audiometry — 74 % correct identifications and no false negatives — must be qualified by the fact that no GSR thresholds were obtained for 21 % of the Experimental Group subjects. That is, they did not meet the GSR conditioning criterion or a reportable GSR threshold was not established. However, GSR audiometry has the additional advantage of precisely specifying organic hearing thresholds, thus providing an objective measure of the extent of the functional component.

The low efficiency of the speech Stenger test, even when it is applicable, is surprising. There are a number of factors that may account for the test's low efficiency. For example, elevated thresholds in the better ear or a minimal functional overlay for speech reception thresholds despite a relatively large overlay for pure tones. A more detailed analysis of the speech Stenger results, when all data are collected, should help identify the factors that appear to affect the efficiency of this test. One final point: Our findings on the Stenger tests are not consistent with Goetzinger and Proud's statement in 1958 that the pure-tone Stenger is „unbeatable”, or with Hood's comment in 1959 that the speech Stenger is „seldom of value”, or with Newby's opinion that the speech Stenger is "practically infallible". Rather, our results suggest that the Stenger tests are neither as bad nor as good as some authors have suggested.

The Efficiency of the Doerfler-Stewart (D-S) Test

Although the Doerfler-Stewart test is not used in our study to select Experimental Group subjects, it is a test that has gained wide acceptance as an efficient screening test for functional hearing loss. It is surprising, then, that there is so little objective evidence on the efficiency of the Doerfler-Stewart test. Our purposes in evaluating this test were to determine its efficiency in detection of functional hearing loss and to compare its efficiency to the efficiency of other tests for functional hearing loss. The Doerfler-Stewart was administered to 29 Experimental Group subjects and 33 subjects who had no functional component at the time of their research evaluation. The test procedure utilized was that described by Doerfler and Epstein in 1956. The number of positive measures was determined by comparing the results obtained for each subject to the norms specified by Doerfler and Epstein.

Results

In evaluating the efficiency of the Doerfler-Stewart test in identifying Experimental Group subjects, the following classifications were utilized. If a

subject had 2 or more positive signs, the test was considered positive; one positive sign was interpreted as equivocal; and zero signs was a negative result. The Doerfler-Stewart test correctly identified 48 % of the Experimental Group, was equivocal in 24 % of the cases, and was negative in 28 % of the Experimental Group.

The same analysis for subjects without a functional hearing loss indicated that the Doerfler-Stewart test was negative in 49 % of the cases, equivocal in 27 %, and positive in 24 % of the subjects without a functional hearing loss. A chi-square analysis revealed that the differences between the groups were not statistically significant. The Doerfler-Stewart test does not appear very efficient when considered by itself or when considered in relationship to some of the tests and test relationships discussed earlier.

Doerfler and Epstein point out that „The number of positive measures... does not seem to be as critical as the specific measures on which a positive result is obtained... The measures involving Noise Detection and Noise Interference... are most sensitive to the presence of functional involvement and should receive a greater weight...” A review of our Experimental Group data reveals that indeed there was a significant difference in sensitivity among the 5 measures. But contrary to Doerfler and Epstein's statement, Noise Detection Threshold and Noise Interference Level are 2 of the 3 **least** sensitive measures. Neither measure correctly identified more than 17 % of the Experimental Group. Only one measure — Speech Reception Threshold (2) minus Noise Detection Threshold — differentiated the functional group from the nonfunctional group at a statistically significant level.

Although our data are not complete, the trends described above suggest to us that until further evidence is available, the Doerfler-Stewart test should be used cautiously. If the present trends continue to the conclusion of the study, we will probably propose a revision of the Doerfler-Stewart test to increase its efficiency. It should be noted, finally, that the Doerfler-Stewart test's inefficiency may be related to measurement problems inherent in the test as well as methodological variables involved in our use of the test.

Summary

In summary, I have presented a brief outline of a major multidiscipline investigation of functional hearing loss. In addition, I have presented some preliminary results on the efficiency of some tests and test relationships commonly used in the detection of functional hearing loss. These results suggest that conditioned GSR pure-tone audiometry, the speech reception threshold-pure-tone average relationship, the pure-tone test-retest relationship and the pure-tone Stenger tests are of high (and comparable) efficiency in detecting the presence of functional hearing loss. The Doerfler-Stewart test and the speech Stenger test are the least efficient of the 6 tests (or test relationships) evaluated. I should point out that my colleagues and I are anxiously awaiting the completion of our study at which time we will be analyzing audiologic data relevant to over 40 research hypotheses. Finally, I wish to express my appreciation to Dr. Joseph Chaiklin who shares the project's audiological duties with me and to Dr. Thomas Trier, the project's research psychologist. They both provided considerable assistance to me in the preparation of this paper.

UNE EVALUATION DE L'EFFICACITE DES TESTS UTILISES POUR LE DEPISTAGE D'UNE PERTE FONCTIONNELLE AUDITIVE

Cette présentation a deux objectifs: premièrement décrire brièvement quelques buts et méthodes d'une vaste enquête scientifique comprenant plusieurs champs d'études sur la perte fonctionnelle auditive, deuxièmement présenter quelques premiers résultats expérimentaux qui permettent d'évaluer l'efficacité de certains tests et les rapports entre différents tests utilisés pour le dépistage d'une perte fonctionnelle auditive.

Les recherches multidisciplinaires comprennent principalement une évaluation intensive de sujets faisant partie d'un groupe expérimental (GE) — ceux-ci présentaient une perte fonctionnelle auditive — et de sujets faisant partie d'un groupe de contrôle (GC) — ceux-ci ne présentaient pas d'élément fonctionnelle à leur perte auditive. Les points généraux suivants indiquent l'étendue et le but principal de cette étude: 1) évaluer l'efficacité de plusieurs méthodes utilisées à l'heure actuelle pour établir le diagnostic perte fonctionnelle auditive; 2) déterminer si les personnes qui souffrent d'une perte fonctionnelle auditive se différencient par certaines caractéristiques de celles qui souffrent exclusivement d'une surdité organique; 3) déterminer les différences qui existent entre les personnes qui résolvent leur perte fonctionnelle et celles pour lesquelles la perte fonctionnelle auditive persiste; et 4) développer à partir des données empiriques de cette étude un cadre théorique susceptible de faciliter à l'avenir le diagnostic et le traitement de perte fonctionnelle auditive et la poursuite d'autres recherches.

Les sujets du groupe GS furent sélectionnés d'après les résultats d'audiométrie tonale, d'audiométrie vocale, d'audiométrie tonale de réflexe psychogalvanique conditionnée (R.P.G.), de l'épreuve tonale de Stenger, et de l'épreuve verbale de Stenger. Ces mêmes sujets furent sélectionnés d'après les critères audiométriques suivants: 1) ils montraient une différence entre le test et la répétition du même test d'au moins 15 db. pour les seuils volontaires à 1000 c/s; 2) une différence de ± 12 db. ou davantage entre la perte auditive mesurée au moyen des fréquences (pure-tone average-PTA) et le seuil vocale (speech reception threshold—SRT) dans la même oreille; 3) un seuil tonale de RPG conditionnée à la fréquence 1000 au moins de 15 db. plus faible que n'importe quel seuil volontaire à 1000 c/s; 4) une épreuve positive tonale de Stenger de 15 db. ou plus au-dessous du seuil volontaire à n'importe quelle fréquence donnée au cours du test; 5) une épreuve verbale positive de Stenger de 15 db. ou plus au-dessous le seuil vocale (SRT) volontaire dans l'oreille reconnue la plus faible.

Les analyses suivantes sont basées sur les résultats d'un groupe expérimental de 47 sujets; l'évaluation des épreuves de Stenger est faite sur un nombre restreint de 19 sujets. Par ordre d'efficacité les résultats sont comme suit: la méthode la plus efficace pour dépister une perte fonctionnelle auditive s'avéra être la relation qui existe entre la perte tonale et la perte vocale (PTA-SRT), cette relation était positive chez 79 % des sujets du groupe GE. Ensuite vint l'audiométrie tonale qui identifia correctement 74 % des sujets du groupe GE. En troisième vint le rapport entre un test tonale et sa

répétition, positif dans 66% des cas. Les tests de Stenger étaient les moins efficaces méthodes utilisées: l'épreuve tonale de Stenger positive dans 53 % des cas et l'épreuve verbale de Stenger positive seulement dans 26 % des cas. Une autre manière d'évaluer l'efficacité est en fonction des résultats faux-négatifs, c'est à dire, l'incapacité d'une épreuve à déterminer un sujet comme fonctionnel. L'audiométrie RPG ne présenta aucune classification fautive-négative; le rapport entre deux épreuves tonales était négatif dans 11 % des cas; et la relation PTA-SRT était négative dans 17 % des cas. Les épreuves de Stenger tonales et vocales donnèrent les pourcentages les plus élevés des cas faux-négatifs, 42 % et 63 % respectivement.

Bien que l'épreuve de Doerfler-Stewart (D-S) ne soit pas utilisée dans cette étude pour la sélection des sujets du groupe expérimental, nous évaluons cette épreuve afin de déterminer son efficacité dans la dépistage de perte fonctionnelle auditive. L'épreuve D-S fut administrée à 29 sujets GE et 33 sujets qui ne présentèrent aucune caractéristique de perte fonctionnelle au moment de l'évaluation. L'épreuve D-S identifia correctement 48 % des cas dans le groupe expérimental, fut équivoque dans 24 % des cas, et négative dans 28 % des cas. La même analyse faite pour les sujets sans perte fonctionnelles auditive révéla le test négatif dans 49 % des cas, équivoque dans 27 %, et positif dans 24 % des cas. Il a été dit que les mesures du seuil de détection de bruit et du niveau d'interférence de bruit sont les plus susceptibles de dépister les pertes fonctionnelles auditives. Toutefois les résultats de notre groupe GE indiquent que ces deux méthodes sont parmi les trois moins efficaces, puisque ni l'une ni l'autre n'identifient plus de 17 % des cas du groupe expérimental.

ACKNOWLEDGEMENT

This study is supported, in part, by U.S. Public Health Service Grant No. B-2741 (C2R1) from the National Institutes of Neurological Diseases and Blindness. The study is also supported by the U.S. Veterans Administration.

Ira M. Ventry, Ph.D.

Auditory Research Project
Veterans Administration Hospital,
San Francisco, U.S.A.