Research Roundtable Summary

4

FOURTH

in a Series of Seminars

on MCHB-funded

Research Projects

Simultaneous Screening for Hearing, Speech, and Language

June 24, 1993 • Parklawn Building Chesapeake Room

Raymond Sturner, M.D.*

Associate Professor
Duke University Medical Center
Director
Duke University Child Development Unit

James Heller, M.A., M.S.

Co-Principal Investigator
Duke University Child Development Unit

Reaction:

Lewis Braida, Ph.D.

Professor
Department of Electrical Engineering
Massachusetts Institute of Technology

* Dr. Sturner is now an associate professor in the Department of Pediatrics at the Johns Hopkins University School of Medicine.

U.S. Department of Health & Human Services
Public Health Service
HRSA
Health Resources & Services Administration
Maternal & Child Health Bureau

This series is sponsored by the Maternal and Child Health Bureau
and coordinated by the National Center for Education in Maternal and Child Health (NCEMCH).
This publication was produced by NCEMCH under its cooperative agreement (MCU-119301)
with the Maternal and Child Health Bureau, Health Resources and Services Administration,
Public Health Service, U.S. Department of Health and Human Services.
For more information, please contact Michelle Keyes-Welch, NCEMCH, (703) 524-7802.

Research Roundtable #4 Summary

Simultaneous Screening for Hearing, Speech, and Language

About This Series

The Research Roundtable Series, sponsored by the Maternal and Child Health Bureau (MCHB), disseminates the results of MCHB-funded research to policymakers, researchers, and practitioners in the public and private sectors. The results of these projects influence future service, research, and policy development. The Research Roundtable sessions provide an opportunity for researchers to discuss their findings with policymakers, MCH program directors, service providers, and other health professionals.

The Maternal and Child Health Research Program is directed by Dr. Gontran Lamberty and administered through the Division of Systems, Education and Analysis, Maternal and Child Health Bureau, Health Resources and Services Administration (HRSA). HRSA is a component of the Public Health Service (PHS), U.S. Department of Health and Human Services (DHHS). The purpose of the research program is to support applied research relating to maternal and child health services that shows promise of making a substantial contribution to the advancement of those services.

Introduction

Aaron Favors, Ed.D., an audiologist/speech pathologist with the Division of Systems, Education and Analysis at the Maternal and Child Health Bureau, introduced the speakers and served as moderator for the roundtable.

Raymond Sturner, M.D., director, and James Heller, M.A., M.S., associate in research, Duke Child Development Unit, presented the results from their MCHB-funded work. Two other members of the research team were also present: H. Donnell Lewis, Ph.D., director, Graduate Speech and Hearing Sciences, North Carolina Central University, and Michael Feezor, Ph.D., assistant research professor, Duke Child Development Unit. Lewis Braida, S.M., Ph.D., professor of electrical engineering, Sensory Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, served as reactor.

Presentation of Research and Relevant Findings

Overview

The goal of this project was to develop a time-efficient and cost-effective test for simultaneously screening speech, language, and hearing in young children. A broader purpose was to develop a standard screening tool with high validity and sensitivity for predicting the developmental status of preschool children.

Dr. Sturner noted that, traditionally, the most common reason for seeing a pediatrician has been for well-child care. Regular health checkups provide the opportunity to assess a child's behavioral and developmental functioning. However, a practice survey has shown that while 97 percent of pediatricians say they are committed to developmental screenings, only 10–15 percent use standardized tests. Instead, pediatricians have relied on informal observations that have been shown by several investigators to be unreliable and invalid.

Early Studies

The current research derives from a previous study in which the team restructured a routine vision acuity test to highlight behaviors considered critical to development. This restructured test was conducted on a countywide cohort of children, and the results were cross-validated with another cohort of children. The research team followed the children in school for 3 years using test results, pass/fail status, and special class placement, and found that these outcomes could be predicted using the restructured vision test. This test, known as the Simultaneous Technique for Acuity and Readiness Testing (START), was a significantly better predictor than IQ test results in the same child, although both tests overlook some children.

The researchers then restructured a standard preschool hearing test to test for development, but found that the prediction of cognitive level was not as accurate as with START. In light of these two studies, the team began thinking about developing a simultaneous screening tool for both speech and hearing.

The goal was to develop a more time-efficient combined test using speech stimuli that also would be a relevant and reliable predictor of difficulties in everyday communication. The combined test would measure the natural interaction between hearing and language as well as the expressive language and articulation required for connected speech.

Preliminary Work

In its preliminary work, the research team looked at two basic issues: The use of spoken words for screening hearing and the minimum number of spoken words that could be used in an effective screening.

A review of existing speech and language screening tests revealed that only one test, the Fluhary, had a validation study showing reasonable predictive indices. Because another study showed much lower rates of prediction, the team conducted its own validation studies, which produced results in agreement with the study showing a sensitivity of language outcomes too low for recommended use. After some additional studies exploring the validity of existing instruments, the team decided to invent a new set of stimuli, using sentence repetition. Sentence repetition was chosen because the repetition response in children may depend on inner language processing capacity rather than simple memory (although Dr. Sturner acknowledged that there was some controversy surrounding this). Also, speech in continuous discourse is the most accurate indictor of the ability to articulate.

The team developed a 3-minute Sentence Repetition Screening Test (SRST) and administered it to 2 cohorts (727 prekindergarten and kindergarten children), with stratified samples of 154 children receiving diagnostic batteries. The results were better than any cited in the literature. The language component had a sensitivity range of 0.62–0.76, with a specificity of 0.90–0.92; the articulation component had a sensitivity range of 0.57–0.74, with a specificity of 0.92–0.94. The team then compared its findings to measurements that used other key predictors (observations by parents, teachers, and health professionals; developmental milestones; and another speech and language questionnaire), and the SRST indices were much higher.

The team also investigated the advantages of using spoken words to screen for hearing, because pure tone screening has a number of limitations, including unreliability of response in young children, ambient noise interference, insensitivity to middle ear pathology, and a limited capability to predict the ability to hear actual words. Although other tools have been developed using words to screen for hearing, they have not proved valid in cases of mild conductive hearing loss. The researchers believed that a realistic screening test should consider the auditory demands of the typical hearing environment.

To overcome the problems associated with using speech signals to screen for hearing, the team (1) carefully selected stimuli, including minimal contrasting consonants of those sounds most difficult for children with impaired hearing to hear; (2) made sound discrimination more difficult by using a babble masker; (3) completed psychoacoustic function studies on both children and adults to define ideal intensity levels and signal-to-noise ratios; and (4) developed an interactive computer system to present visual and precisely controlled acoustic stimuli.

Prototype Development and Testing

The team developed a prototype test by modifying SRST test items so that pictures accompanied the concepts and by using some sentences from the Northwestern Sequence Screening Test (NSST). The team used slightly different versions of the stimuli sentences to test for critical listening and understanding of the concepts. The team also developed vocabulary stimuli to avoid the redundancy of sentence context. The team oversampled later-developing phonemes for articulation and high-frequency phonemes for hearing. (Phonemes are the smallest phonetic units capable of conveying a distinction in meaning.) They presented these words in minimum contrast sets (e.g., sick, kick, lick, chick) to test for hearing and language. The researchers then conducted studies of the test prototype with 300 adults and children who had normal speech, language, and hearing.

The research team conducted studies of picture recognition and visual format as well as audio stimuli. Pictures were initially presented to the children in sets of 6, then in smaller sets in order to assess vocabulary (with a level of 80 percent recognition) in a cohort of 80 preschool children (20 in each of the 4 preschool age groups). After studying different backgrounds and child preferences, the team selected the best visual formats. The researchers digitized the pictures and redigitized the audio stimuli until a group of expert listeners was able to respond at a level of 100 percent.

The researchers evaluated the psychoacoustic characteristics of test items with children and adults, comparing babble, white noise, and low intensity. These studies confirmed that babble was superior in discriminating between normal and impaired hearing. The team also determined that the signal-to-noise ratio was more important than the absolute level of the signal.

Components of the Simultaneous Screening Test

The test stimuli were composed of word-picture pairs and sentence-picture pairs, both presented with background noise. Both pairs had receptive and expressive listening/speaking response modes.

The testers presented 28 word-picture pairs with 55 consonants (i.e., every word but 1 had 2 consonants). In this part of the test, a child was shown a picture, the target was presented, the child repeated the target, and the target was scored. Following this, a series of four pictures were presented, the auditory target was presented, and the child was asked to pick one of the four pictures.

In addition, 23 sentence-picture pairs were presented. Within the sentences, 48 consonants were scored for articulation. The sentence-picture pairs also contained 159 morphemes (the smallest word or part of a word that conveys meaning), which were also scored.

Study Design and Methodology

The study sample was composed of 244 children; of these, 32 had hearing impairments, 44 had language impairments, and 42 had speech impairments. Of the total sample population, 53 percent were male and 47 percent female. There were more males in the language-impaired group, because more males have identified language impairments.

The mean age of the study population was 58 months. The mean age of the hearing-impaired group was somewhat older, because the team recruited older children with known hearing impairments to determine the hearing component of the test.

Of the 244 children participating, 78 percent were white, 20 percent were African American, and 2 percent belonged to other ethnic groups. African-American children were somewhat overrepresented in the impaired group. Future research is needed to differentiate between language differences as a result of dialect and disorders.

A validation study was performed to identify the set of test items that best distinguished children with normal articulation, language, and hearing abilities from those with impairment. First, univariate analysis was performed to identify children with impaired hearing, speech, or language. Second, summary scores were calculated (i.e., the sum of significant phonemes in hearing, the percentage of correct morphemes in sentences for language, and the Developmental Sentences Score of expressive sentences for speech.) Third, discriminant function analysis was performed for the variables identified in the first two steps.

Study Findings

Hearing impairment. Univariate analysis of 314 possible items (receptive, expressive, morphemes, and phonemes) revealed that 119 items had a probability (*p*) value of 0.05 and 57 had a *p* value of 0.01. Discriminant function analysis used 21 items in the final analysis, which correctly identified 26 of 32 (81 percent) of the children who were hearing impaired and 103 of 105 children (98 percent) with normal hearing. Thus, the reduced set of 21 items correctly classified 95 percent of all the children.

Articulation impairment. Univariate analysis of 103 phonemes revealed that 68 could identify a significant difference in performance between normal and impaired groups. Discriminant function analysis used 32 items. In the final analysis, these 32 items correctly classified 37 of the 40 children (93 percent) with impaired articulation and 128 of 129 children (99 percent) with normal hearing. Thus, 98 percent of all the children were correctly classified.

Language impairment. Univariate analysis of 210 possible items (all receptive items in all morphemes) revealed that 67 had a *p* value of 0.05, and 33 had a *p* value of 0.01. Discriminant function analysis used 31 items in the final analysis. These 31 items correctly classified 32 of the 40 children (80 percent) who were language impaired and 122 of 127 children (96 percent) with normal hearing. Thus, the test correctly classified 92 percent of all the children.

Missed cases of hearing impairment. Six of 32 language-impaired children in the study were incorrectly identified as having normal hearing. In one of these children (age 3-1/2 years), behavioral factors seemed to have influenced test results. Of the five remaining children, three had very mild sensory neural hearing loss, and two had middle ear fluid that the team failed to identify.

The analysis conducted as a part of the validation study accomplished three purposes. It explored the potential of the items to discriminate impaired from normal groups; it emphasized correct classification of each of the three impairments with a limited set of items; and it identified a set of test items for each impairment group.

The research team concluded by citing the benefits of developing this new communication screening system. First, the test system is user-friendly for young children. Currently, many children under age 4 are not screened for communicative abilities (especially hearing); this screening test reduced the age level to 3-1/2 years without any difficulty and can probably be used in children as young as age 3. Second, this screening test appears to have overcome the previous obstacles to the use of spoken words as audiometric screening test signals. Third, preliminary data suggest that this new screening system is more time efficient. Fourth, the system uses suprathreshold test signals with a masker background and thus is more likely to overcome much of the ambient noise problems seen in typical screening settings.

Ongoing Studies

Additional analyses are being conducted to combine all three sets of items into a single test set for the simultaneous screening of articulation, language, and hearing. Once this single set is identified, it will be tested and refined further. The team is currently conducting three studies: (1) A validation in a large representative population, with subsequent refinement of the test; (2) a replication and comparison study; and (3) a preliminary dialect scoring of African-American vernacular.

The purpose of the validation study is to set a “cut” score and shorten the test based on the best prediction of hearing, articulation, and language. A secondary purpose is to assess temporal stability and intertester reliability. The population sample is composed of 400 African-American children and 400 white children. The study design involves two test sessions: Screening/hearing and diagnostic speech and language.

The purpose of the replication and comparison study is to compare (1) the prediction results using the “cut” score derived from the earlier study; (2) this test with other hearing and speech/ language screening tests; and (3) the results in a typical ambient noise environment. The population sample will be similar to the cohort in the earlier validation study. The design also will be the same as the one used in the earlier validation test.

In the study of language disorder versus language difference, an independent test battery will be administered to a subsample of children to determine the probability of dialect usage. The goal of this study is to develop an alternative scoring system for dialect and/or secondary testing to determine dialect use. There are additional benefits to developing dialect scoring and testing. Alternative dialect scores will be available for diagnostic tests, and new diagnostic procedures for dialect testing will gain preliminary validation.

Reaction

Dr. Lewis Braida commended the work of the research team in developing the test instrument. He noted that the specific mechanisms (e.g., interactive computer with touch-screen display, earphones) and the testing environment were very positive accomplishments that could eventually contribute to the long-term goal of the project. Dr. Braida also commended the research team for recognizing the need for large-scale validation and replication studies. Dr. Braida expressed concern, however, about three fundamental issues: (1) The use of discriminant analysis in a heterogeneous population to diagnose specific types of hearing loss, (2) the costs involved in missing cases, and (3) the fundamental philosophy underlying combined testing for hearing, speech, and language.

Dr. Braida expressed concern about using discriminant analysis to arrive at specificity and sensitivity levels when testing for different types of hearing loss caused by very different problems (e.g., sensory neural hearing loss compared to hearing loss caused by otitis media or malfunction of the middle ear). He queried the team on how the test should be designed and discriminant analysis completed to get the specificity and sensitivity that are needed for the different populations.

Dr. Braida expressed concern that, in a heterogeneous population, the costs of misdiagnosing different problems vary considerably across that population. Sensory neural hearing loss, which is fairly rare, has a high cost if it is missed. On the other hand, otitis media, if it is missed, has a fairly low cost. The cost of misdiagnosing a child with normal hearing compared to one with impaired hearing is fairly modest, simply requiring a visit to an audiologist.

In summary, Dr. Braida said that the cost and frequency of occurrence should help determine the balance between specificity and sensitivity. He feels that this important component is not being adequately addressed in the research.

Dr. Sturner explained that the final analysis will not rest solely on discriminant equations. The team is selecting high-frequency items that will be sensitive to sensory neural loss and plans to include those items in the screening test even though they may not have been identified through the discriminant equation. The team is also examining some contrasts between the sensory neural group and the group with normal hearing to see what items can be identified through backward elimination. Dr. Sturner emphasized that the discriminant equation looks for the best percentage agreement and the best categorization. It is not geared to weight sensitivity more than specificity. The team can move up the “cut” score and sacrifice specificity for sensitivity.

Dr. Braida had difficulty in accepting the premise that one can efficiently test hearing, articulation, and language abilities in the same test. He felt that there may be a natural hierarchy among the three elements and that it might be better to separate the tests into parts. Dr. Braida raised the issue of whether one test that simply identifies the child as having a problem in either speech or hearing or language is the right kind of screening test. According to Dr. Braida, as the test is now designed, hearing problems are likely to show up not only as hearing problems, but also as articulation and language problems.

Dr. Braida referred to study data that indicated that, on average, the groups whose scores indicated hearing failure scored 16–19 points lower on the word portion of the test and 20–36 percent lower on the sentences portion. Dr. Braida suggested that this was because the repetition items of the test were performed at a different signal-to-noise ratio than the receptive items. Dr. Braida questioned why noise was used at all in these tests. He suggested simply presenting the items at very high signal-to-noise ratios and reasonable intensities. There is a risk that a child who fails the language or articulation component of the screening test may simply have a temporary hearing loss on that test day. The cost would involve sending those children to audiologists and to diagnostic experts in articulation and language, which would double or triple the cost of false positives. Dr. Braida suggested that the language and articulation components of the test use clearly audible speech or pictures rather than sound. He expressed concern that only 80 percent of the words were known by the children in one of the tests, because that raises the question of whether a test for hearing also involves a test of vocabulary.

Dr. Sturner explained that simultaneous screening frees up to 80 percent of children from concern and further testing, and the team hopes that figure will reach 90 percent with some refinements in the screening test. Several members of the research team mentioned the distinction between first-level testing and secondary testing. At this point in the research, Dr. Sturner emphasized the importance of identifying the smallest number of items that can differentiate between normal and impaired hearing, speech, and language.

Dr. Lewis emphasized that the test is not being designed for speech and language professionals; it is being designed for first-line providers of health care for children. Dr. Braida concluded by saying that, while he believes there are opportunities to make the research stronger, he is impressed with the work of the team. He emphasized the importance of developing more sensitive tests that can replace the casual observations now being used by health providers.

Publications

Heller JH, Sturner RA, Funk SG, Feezor MD. 1993. The effect of input mode in vocabulary identification performance at low intensity. *Journal of Educational Computing Research* 9:509–518.

Sturner RA, Funk SG, Green JA. 1994. Simultaneous technique for acuity and readiness testing (START): Further concurrent validation of an aid for developmental surveillance. *Pediatrics* 20:82–88.

Sturner RA, Heller JH, Funk SG, Layton T. 1993. The Fluharty speech and language screening test: A population based validation study using sample independent decision rules. *Journal of Speech and Hearing Research* 36:738–745.

Sturner RA, Kunze L, Funk SG, Green JA. 1993. Elicited imitation: Its effectiveness for speech and language screening. *Developmental Medicine and Child Neurology* 35:715–726.

Sturner RA, Layton T, Heller JH, Funk SG, Evans AW, Machon M. 1994. A review of available speech and language screening tests for preschool children. *American Journal of Speech-Language Pathology* 2:25–36.

Worley G, Sturner RA, Green JA, Frothingham TE. 1994. A relationship between brachycephaly and middle ear effusion in children. *Clinical Anatomy* 7:84–89.