CONDITIONED RESPONSE AUDIOMETRY: A CLINICAL TECHNIQUE AND INSTRUMENTATION FOR TESTING EXCEPTIONAL CHILDREN

THESIS FOR THE DEGREE OF Ph. D.

MICHIGAN STATE UNIVERSITY
STEWART W. KINDE
1972

This is to certify that the

thesis entitled

CONDITIONED RESPONSE AUDIOMETRY: A
CLINICAL TECHNIQUE AND INSTRUMENTATION
FOR TESTING EXCEPTIONAL CHILDREN

presented by

Stewart W. Kinde

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Special Education

Charles V. Mange

Major professor

Date 2-16-72

O-7639

ABSTRACT

CONDITIONED RESPONSE AUDIOMETRY: A
CLINICAL TECHNIQUE AND INSTRUMENTATION
FOR TESTING EXCEPTIONAL CHILDREN

by Stewart W. Kinde

The purpose of the present study was to develop and evaluate a hearing test method which would have immediate and widespread clinical application for use with young and non-communicative children.

The resultant experimental procedure was labeled Conditioned Response Audiometry (CRA) and utilized a two-phase structure. The first phase was limited to conditioning procedures directed toward establishing stimulus control. A multisensory approach was employed using visual, tactile and auditory stimuli. A portable audiometer was utilized and an adaptor, called a Shaping Adaptor for Portable Audiometers (SHAPA), was developed and designed to permit and facilitate the simultaneous presentation of the three stimuli. Responses were modeled and as stimulus control was established each stimulus, in turn, was faded out until the subjects were responding to the auditory stimulus alone. The second or testing phase utilized this conditioned response with two testers employed to check and maintain the validity of the subjects' responses. One tester, positioned in front of the

subject, maintained attention to the task and signaled for stimulus presentations. The second tester, positioned out of view, presented a schedule of test tones interspersed with no-stimulus presentations which served to point up false-positive responses.

Twenty children were selected, from the total population of a pre-school deaf education program, to serve as subjects in an evaluation of the efficacy of the CRA approach. These subjects were assigned to test groups on the basis of their ability (Group I) or failure (Group II) to respond consistently and appropriately to conventional pure tone hearing test procedures.

Three factors were of interest in this study: (a) the test/retest reliability of CRA procedures; (b) a comparison of thresholds obtained by CRA with those obtained by conventional play audiometry and; (c) the determination of the validity of CRA thresholds obtained from young children demonstrated to be untestable by conventional means. Since the latter two factors were incompatible by definition, the use of two test groups and the inclusion of an objective test measure (ERA) was thought to provide an indirect assessment of these variables.

The test measures employed with Group I included conventional play audiometry (CPA), evoked response audiometry (ERA) and conditioned response audiometry (CRA).

Group II received only CRA and ERA. All tests were administered twice to each subject in each group.

Within the limits imposed by the study the results indicated the following conclusions: 1) CRA test procedures can be employed to obtain consistent threshold measurements over extended test occasions. 2) Threshold levels obtained by CRA appear to be linearly related to those obtained by conventional methods. The difference factor may reflect the involvement of different response-criteria under each test method. 3) Some children cannot be conditioned to respond consistently under CRA-phase I procedures and consequently cannot be tested by phase II test procedures. There is evidence to suggest that such children may be representative of a group for whom hearing loss is not, educationally, their primary problem.

CONDITIONED RESPONSE AUDIOMETRY: A CLINICAL TECHNIQUE AND INSTRUMENTATION FOR TESTING EXCEPTIONAL CHILDREN

By

Stewart W. Kinde

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Special Education

1972

67-1615

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The early detection of hearing loss is of prime importance in the success of most habilitative measures for those with severe auditory disorders. This holds true whether the auditory disorder is the primary problem or whether it is overlaid on a more pervasive handicap. While the preceding statement is widely accepted among those concerned with the education of exceptional children, the means by which it can be accomplished, on any significant scale, are at the present time extremely limited.

The detrimental effects of such diagnostic limitations relative to the needs of the congenitally deaf child with other learning problems was emphasized in the report of the National Research Conference on Day Programs for Hearing Impaired Children published in 1968:

> Frequently educational placement and programing for multiple handicapped (deaf) children is made without first establishing the primary handicap thus compounding the problem...there is an urgent emergency to find ways and means for extending the effectiveness of diagnostic efforts which, on a nationwide scale, are something less than comprehensive.[1]

---

[1]A. M. Mulholland and G. W. Fellendorf, Final Report

Similarly the Sub-Committee on Human Communications and Its Disorders of the National Advisory Neurological Diseases and Stroke Council in 1969 stressed the need for special methods for the identification of hearing loss in infancy and early childhood.[2] It was also pointed out that the methodologies of operant conditioning appear to have high potential when used with children not far removed from infancy but have not been systemically applied to audiometry and developed to the stage of routine procedures. It was also noted that there is a paucity of documented data on such factors as the reliability and validity of auditory tests on infants and young children compared with standardized procedures.

Statement Of Purpose

The early initiation of speech and language training for the deaf child is accepted practice in the United States as well as many other countries. This is evidenced by the fact that while it is not generally mandatory in this country, for tax supported schools to provide pre-school classes for deaf children, there were over five thousand children under

---

of the National Research Conference on Day Programs for Hearing Impaired Children (Washington, D. C.: Alexander Graham Bell Association for the Deaf, 1968), p. 23.

[2]Human Communications and Its Disorders - An Overview,

the age of six years enrolled in public and private residential and day school programs during the 1967-68 academic year.[3]

Silverman and Lane suggested that the impetus for the movement toward early educational intervention has come in large part from indirect evidence accumulated in the fields of neurophysiology and psycholinguistics. This evidence indicates critical early periods when sensory experience influences the development of the nervous system and when syntactical aspects of language are most efficiently acquired.[4]

It is evident that the implementation and growth of pre-school deaf educational programs is dependent upon available clinical procedures for identifying and differentiating auditory disorders in young children. While various testing techniques have been developed over the years, most have failed to provide adequate or reliable information or have been unsuccessful with the very young or severely handicapped

---

A report prepared and published by the Subcommittee on Human Communication and Its Disorders (Bethesda, Maryland: U. S. Department of Health, Education and Welfare, 1969), p. 63

[3]"Type and Size of Educational Programs Attended by Hearing Impaired Students in the United States: 1968-69," Annual Survey of Hearing Impaired Children and Youth, Series D, Number 4 (Washington, D.C.: Office of Demographic Studies, Gallaudet College).

[4]S. R. Silverman and H. S. Lane, Chapter 16, Hearing and Deafness, eds. H. Davis and S. R. Silverman (New York:

child. In addition, while test procedures utilizing conditioning techniques have been demonstrated successful in obtaining threshold measurements with young and mentally retarded children in laboratory settings, they have not been restructured or standardized for general clinical application.

The present study was conducted in an effort to incorporate, adapt, and instrument into a structured technique having potential for immediate and widespread clinical application those methods of behavioral modification and audiometry which have been demonstrated, through research, to be effective and reliable with non-communicative children. The resulting method was labeled Conditioned Response Audiometry (CRA) and met the necessary criteria of minimal equipment needs, minimal administrative skill, maximum information yield and maximum control of interfering variables. In addition, fundamental questions were posed regarding the reliability and validity of the CRA approach.

Importance Of The Study

The majority of conventional hearing test procedures have been developed for and standardized on adult populations. As a consequence such tests have had limited applicability for the identification and measurement of hearing loss among

---

Holt, Rinehart and Winston, 1970), p. 393

young and handicapped children. Frisina stated that the difficulties encountered in testing young children are most often related to the stimulus materials employed, response modes required, instructional modes utilized and subjective maturational factors.[5] These associated variables are generally compounded in the case of the exceptional child. Emotional disturbance, mental retardation and other central nervous system disorders obfuscate the symptoms of auditory problems and complicate or frustrate the measurement process.

Davis reviewed current development in audiometry and stressed the need for special testing techniques with certain children. He stated:

> The children that require special testing are those (1) who are too young to understand, (2) who are mentally retarded, (3) who cannot make appropriate responses because of cerebral palsy even though they may hear something, (4) who are emotionally disturbed, (5) who respond only erratically or not at all to any stimuli, (so-called "autism" or childhood schizophrenia) or, (6) who may suffer from congenital or early acquired hearing loss so severe that they have failed to learn to pay any attention to such sounds as they may be able to hear.[6]

The report of a conference on current practices in the

---

[5]D. R. Frisina, Chapter 4, Modern Developments in Audiology, ed. James Jerger (New York: Academic Press, 1963), p. 128.

[6]H. Davis and R. Goldstein, Chapter 8, Hearing and Deafness, eds. H. Davis and S. R. Silverman (New York: Holt, Rinehart and Winston, 1970), p. 238.

management of deaf infants sponsored in 1968 by the Joint Committee on Audiology and Education of the Deaf stated that while modifications of conventional testing procedures, generally referred to as play audiometry, are reported in the literature to be successful methods of obtaining auditory thresholds for children two years of age and older, for a good number cooperation may be difficult to obtain under the age of four years particularly when hearing loss is profound.[7] It was found that as a consequence most audiologists use sound field examinations with younger children including those from six months to two years of age. However, this approach was described as less than adequate since information is obtained about the better ear only, since no differentiation can be made between air and bone conduction thresholds to detect children with a conductive element and since it is most often successful only with children who have considerable residual hearing.

Initial contact with children exhibiting severe communication disorders often occur in field settings that are minimally equipped and staffed for diagnostic services. Such limitations are usually the unalterable concomitants

---

[7]Freeman McConnell, ed. Proceedings of the Conference on Current Practices in the Management of Deaf Infants (Nashville, Tennessee: The Bill Wilkerson Hearing and Speech Center,Vanderbilt University, 1968).

of the clinic's professional orientation, demographic considerations and financial restrictions. Since as mentioned above, the child without speech or language presents some unique problems which confound the diagnostic process, outside assistance is often sought and habilitative processes are deferred. However, referrals to diagnostic centers from such field clinics often present formidable logistical problems related to delays in obtaining appointments, arranging and scheduling transportation and payment of diagnostic fees. Too often such preliminary obstacles remain unresolved and habilitative efforts are permanently forestalled or delayed beyond critical developmental periods. In other cases when referrals are accomplished and deafness is suspected, few differential data are obtained during the initial appointments since the non-communicative child is seldom able to respond to conventional testing procedures. As a consequence multiple testing sessions are scheduled or the recommendation is made to defer testing until the child has obtained sufficient maturity to ensure cooperation. Both situations involve an undesirable protraction of the diagnostic process and some probability that it may never be completed.

The seriousness of such diagnostic delays was stressed in a 1967 report of the National Conference on Education of the Deaf and related to the deaf child's early need for amplification which would allow him to assimilate language in a manner similar to that of a normal hearing child.[8]

In addition, there is substantial indication that the numbers of multiply handicapped deaf children who require specialized testing techniques is increasing. Hardy reported that over fifty percent of a population of congenitally deaf children seen at the John Hopkins Medical Institution have one or more handicapping conditions in addition to the auditory disorder.[9]

It seems, from this review, that the accepted need for early initiation of training, limitations and restrictions in the means for early identification, combined with an apparent increase in the prevalence of the multiply handicapped deaf adds considerable urgency to the problem.

As a consequence, it is believed that the availability of instrumentation and procedures such as those developed in the present study might do much to eliminate the problems mentioned above. They could provide a hearing screening

---

[8]Education of the Deaf, The Challenge and the Charge, A Report of the National Conference on Education of the Deaf (Washington, D.C.: U. S. Government Printing Office, 1967), p. 72.

[9]W. C. Hardy, "Early Detection and Assessment," Proceedings of International Conference on Oral Education of the Deaf, Volume I (Washington, D.C.: Alexander Graham Bell Association for the Deaf, 1967), p. 8.

technique for the very young or multiply handicapped child who is most often considered untestable in the usual field setting and may facilitate appropriate referrals from such clinics. In addition, such procedures should expedite the diagnostic process since initial conditioning could be established prior to referral to the diagnostic center.

The magnitude and significance of the problem and direction toward a solution were most appropriately summarized by Hardy, in 1967, in a paper presented at the International Conference on Oral Education of the Deaf:

> In these days of further insights into causes of deafness, and of the general extension of effective chemotherapy, and of refined medical practices, it is most important that children with hearing impairment be found early in life, and that their capacities and limitations be assessed as soon as this is feasible. There are more deaf children than ever before. The proportion of those who are born deaf, or become deaf in the neonatal period, is increasing. They are, as well, deaf in more complex ways than was apparently true some years ago. There are many more children who have multiple problems, many of which are at least as severe as the deafness. For all these reasons facilities in health and education are rapidly becoming overtaxed, and probably need extension in ways somewhat different from past practices.[10]

Definition of Terms

The following definitions are employed in this investigation:

---

[10] *Ibid.*, p. 1

1. Sound Levels are the levels of pure tones in dB relative to ISO - 1964 standards.[11]

2. Conventional pure tone hearing test procedures refer to the unilateral presentation of pure tones, through ear phones, employing the modified method of limits described by High, Glorig and Nixon.[12] A pulse stimulus tone with a duration of about .5 seconds was initially presented at a level of 120 dB. The attenuation was then decreased in 5 dB steps until the subject no longer responded to the stimulus tone. The attenuation was then decreased an additional 10 dB below the point where the subject last responded and then increased 5 dB steps until a response was obtained. The lowest of these two measures was accepted as threshold.

3. Play audiometry refers to variations of the response mode which involved gross motor responses considered appropriate to the developmental level of the subjects. These included placing plastic rings on a dowel, placing blocks in a basket, and fitting pegs in a hole.

Limitations Of The Study

Only a limited aspect of auditory perception was examined in this study. This involved the operationally defined threshold response to pure tones of the frequencies 500, 1000, and 2000 cps. Primary variables of interest include intratest reliability corresponding to sessions and frequency and interest validity corresponding to comparisons of the CRA, ERA and conventional procedures. Any variables not designated

---

[11] Standard Reference Zero for Calibration of Pure Tone Audiometers, ISO/R389 - 1964

[12] W. S. High, A. Glorig and J. Nixon, "Estimating the Reliability of Auditory Threshold Measurements," Journal of Auditory Research, 1 (1961) pp. 247-262.

and/or not controlled in the statistical design were assumed to be normally distributed among subjects within each of the two test groups. Any actual differences that may have existed among these groups were assumed to be non-significant statistically and, where this assumption was questionable, appropriate caution was exercised in the interpretation of the observed results.

Organization Of The Study

Chapter I is organized to give an overall view of the subject matter, the apparent status of knowledge on this subject, questions of interest, populations involved, and the research approach influencing the conduct of the investigation. In Chapter II literature pertinent to this study is reviewed, interpreted and summarized under the general headings of play audiometry, objective techniques and conditioning procedures. The development of an experimental, two-stage, hearing test for young and non-communicative children is described in Chapter III. Details are presented on the instrumentation and the sequential administrative steps of each test phase are outlined. The subject population, the selection of subjects from this population, their assignment to test groups, the test methods employed and the statistical design are described in Chapter IV. In Chapter V results of the investigation are presented and discussed in relation to the questions posed earlier in the present chapter. The study is summarized, the conclusions are stated and the implications for additional research are set forth in Chapter VI.

CHAPTER II

REVIEW OF PERTINENT LITERATURE

Having described the purpose, scope and limitations of this study in Chapter I, literature pertinent to these topics is reviewed in this chapter. The literature reviewed is arbitrarily classified under the general headings of play audiometry, objective techniques and conditioning procedures. A summary of this review is given at the end of the chapter.

The classification system employed was recognized to contain some degree of procedural overlap. As a result test techniques described under any single heading may be found to contain elements of respondent or instrumental conditioning which are common to the remaining categories. This reflects, however, a broad use of terminology to point up distinct and somewhat progressive stages in the development of child audiometrics. The heading "Play Audiometry", for example, has been used to represent an initial period when the focus of test development was directed somewhat exclusively towards the reinforcers of behavior. Fulton and Lloyd described this as a time when it was assumed there was something magical in toys and pictures which enhanced responses.[1] The term "Objective Techniques" delineates a test development stage in which research interests were primarily directed toward reflex responses to acoustic stimuli and the instrumentation and utilization of respondent conditioning principles for detection and measurement. The most current stage of test development has been categorized under the heading "Conditioning Procedures." This period has been characterized by a strict adherence to and utilization of principles of instrumental or operant conditioning.

Play Audiometry

Over the past several decades numerous procedural alterations have been devised and directed toward eliminating or circumventing the intervening variables, which have been previously pointed out to be associated with testing children. Since developmentally, and often pathologically in the case of exceptional children, receptive and expressive language abilities are limited, most such approaches have involved pure tone stimuli. Stimulus-response associations have been simplified and hopefully the total test procedure made more "attractive" through the use of toys and games. There has been considerable diversity in the types of play

---

[1] R. F. Fulton and L. L. Lloyd, Audiometry for the Retarded: With Implications for the Difficult-To-Test

equipment utilized. It is also apparent that most such techniques have to some degree followed the conditioning rubric. The "attractiveness" or reinforcing properties of these approaches were, however, often assumed or believed to be inherent in the toys and games employed or in the approving behavior of the tester (i.e., social reinforcement).

As early as 1930 Ewing described an apparatus, consisting of a toy electric train in a tunnel, which he believed would facilitate the testing of young and exceptional children.[2] The train could be moved from the tunnel only during periods when the child under test perceived a test tone and activated a switch. The function of the child's switch was controlled by the tester and coordinated with the presentation of tones.

During this early period Guilder and Hopkins suggested the use of auditory training before an audiometric test was administered to a young child.[3] They felt that in this way the child could learn to recognize the test sounds and become "accustomed" to giving the proper response. They also

---

(Baltimore: The Williams and Wilkins Company, 1969), p. 13.

[2]A. W. G. Ewing, Aphasia in Children. (London: Oxford Medical Publication, 1930).

[3]R. P. Guilder and L. A. Hopkins, "Program for the Testing and Training of Auditory Function in the Small Deaf Child During Pre-School Years," (Part I) Volta Review, 37 (1935), pp. 5-11

suggested that the response mode could be simplified and motivation provided during the test situation by having the child turn on a small flashlight whenever he heard the tone.

A unique adaptation of stimulus-response modes was developed and described by Bloomer.[4] Pictures of animals or objects were utilized which the author believed made a noise corresponding to the frequency of the pure tone employed in the test. A pretraining period was involved during which the test tone and the representative picture was associated and the child was instructed in making this paired-association by pointing to the picture.

Another of the early, and perhaps most widely known, adaptations of conventional test procedures for use with young children was the "Peep-show", developed by Dix and Halpike.[5] The apparatus consisted of a box with pictures inside which could be illuminated, a viewing hatch in the front of the box and a rather elaborate switching circuit, similar to the one described by Ewing[6], which would permit a child under test to illuminate a picture only at the

---

[4]Harlan Bloomer, "A Simple Method of Testing the Hearing of Small Children." _Journal of Speech Disorders_, 58 (1949), pp. 751-759.

[5]M. R. Dix and C. S. Hallpike, "The Peep Show: A New Technique for Pure Tone Audiometry in Young Children." _British Medical Journal_, 2 (1947), pp. 719-723.

[6]Ewing, _loc. cit._

discretion of the tester. Procedures were also similar and involved the presentation of pure tones to "cue" the child as to appropriate response times along with the simultaneous activation, by the tester, of the child's response switch to permit him to illuminate the box and view the picture.

Following the lead of the "Peep-show" approach, several minor variations in equipment and procedures were suggested. These are exemplified by the reports of O'Neill, Oyer and Hillis[7] who described a modification in which the pictures in a box were replaced with colored slides of fairy tale characters projected on a screen, and Weaver[8] who incorporated filmstrip stories.

Concurrently, a number of major variations in instrumentation were being developed. These involved the use of a variety of animated toys and included such things as a puppet show employed by Waldrop[9] and a mechanical dog used

---

[7] J. O'Neill, H. Oyer and J. Hillis, "Audiometric Procedures Used With Children." Journal of Speech and Hearing Disorders, 26 (1961), pp. 61-66

[8] R. M. Weaver, "The Use of Filmstrip Stories in Slide Show Audiometry." In, The Audiologic Assessment of the Mentally Retarded: Proceedings of a National Conference, eds. L. L. Lloyd and D. R. Frisina, (Kansas: Parsons State Hospital and Training Center, 1965), pp. 71-88

[9] W. F. Waldrop, "A Puppet Show Hearing Test." Volta Review, 55 (1953), pp. 488-489

by Green.[10] The most elaborate of such adaptations was developed and commercialized by Guilford and Haug under the trade name Pediacoumeter.[11] This device consisted of seven, solenoid activated, Jack-in-the-Box heads, each representing a separate audiometric frequency, operated through a switching arrangement and procedures similar to those employed in the "Peep-show."

It is apparent, however, that the success of these concerted attempts to find a toy or picture which would ensure appropriate test participation on the part of young children was very limited since the literature is replete with a progression of additional references which describe procedures and gadgets of similar kind. In addition, some of the most widely used of these techniques, such as the Peep-Show, the Pediacoumeter and the paired-association method suggested by Bloomer,[12] were employed in studies designed to evaluate their effectiveness with groups of children of various age levels. The most relevant of these were studies

---

[10] D. Green, "The Pup-Show: A Simple, Inexpensive Modification of the Peep-Show." _Journal of Speech and Hearing Disorders_, 23 (1958), pp. 118-120

[11] F. Guilford and C. Haug, "Diagnosis of Deafness in the Very Young Child." _Archives of Otolaryngology_, 55 (1952), pp. 101-106.

[12] Bloomer, _loc. cit._

carried out by Haug and Guilford,[13] Statten and Wishart,[14] Myklebust[15] and Barr.[16] The research populations involved in all of the studies consisted of children who did not exhibit any mental or physical handicaps. The results of these studies are listed in the following table:

TABLE 1. Percentage of successful pure tone threshold measurements by age with various procedures of play audiometry.

		Age in years				
Investigator	N	Five	Four	Three	Two	One
Haug & Guilford	968	96	94	82	47	12
Statten & Wishart	204	90	90	73	34	7
Myklebust	61	83	93	68	0	-
Barr	138	92	88	71	41	-

[13]C. Haug and F. Guilford, "Hearing Testing on the Very Young Child: Follow-Up Report on Testing the Hearing of 968 Preschool Patients With the Pediacoumeter," Transactions of the American Academy of Ophthalmology and Otolaryngology, 64 (1960), pp. 269-271

[14]P. Statten and D. Wishart, "Pure-Tone Audiometry in Children: Psychogalvanic-Skin-Resistance and Peep Show." Annals of Otology, Rhinology, and Laryngology, 65 (1956), pp. 511-534

[15]H. Myklebust, Auditory Disorders in Children. (New York: Grune and Stratton, 1954.)

[16]B. Barr, "Pure Tone Audiometry for Pre-School Children," Acta-Otolaryngologica, Supplementum 121 (1955), pp. 1-84

It seems apparent from an inspection of the results in Table 1 that the so-called play techniques have not proven to be highly successful at or below the age of three years. As indicated approximately 25% of the children could not be tested at the age of three years, less than 50% at the age of two years and at best only 12% could be tested at the age of one year.

Since these techniques have not proved to be highly successful with relatively normal young children, there is some reason to believe they would be even less successful with children who exhibit psychological or behavioral disorders. This assumption has been in part substantiated by several investigators. Representative of these are the reports of Ewing[17] who attempted the use of his tunnel test with aphasic children and Thorne[18] who evaluated a number of play techniques with the mentally retarded. Both studies indicated the toys were distracting concrete stimuli which appeared to inhibit rather than enhance attention to the abstract pure tone stimuli.

On the basis of the literature reviewed thus far it is evident that adaptations of conventional hearing test pro-

---

[17]Ewing, _loc_. _cit_.

[18]B. Thorne, "Conditioning Children for Pure Tone Testing." _Journal of Speech and Hearing Disorders_, 27 (1962) pp. _84-85_

cedures which have been labeled as play audiometry have facilitated hearing testing down to the chronological age of four years. They do not, however, provide an adequate means by which reliable and valid measurements can be made to determine the hearing acuity of young or exceptional children during the critical learning stage from one to three years.

Objective Techniques

The development of methods for measuring sensory experience which are free from subjective influences has been a common interest of many professional fields. The advantages of such an approach are self-evident under the tenets of scientific methodology. In audiology, major efforts in this direction have utilized reflex kinds of behavior as response modes. For the most part these have involved the transient changes in ongoing physiological activity mediated in some degree by the autonomic nervous system and induced by relatively intense stimuli. Some of the responses involve overt motor activity, such as startle or orienting responses, while others are covert, but recordable, electrophysiologic events.

Within the field of audiometry the focus on development of such techniques has been motivated in large part by the medico-legal aspects of hearing loss and the demonstrated inefficacy of conventional test procedures with young and exceptional children. The terms "objective audiometry" or objective testing techniques" are commonly employed in the literature to describe any hearing test method that does not require a volitional response on the part of the subject. However, some criticism has been directed towards this usage. Goldstein has pointed out that since audiometry is more than a response, the term 'objective' can only apply if the technique is also objective.[19] This appears to be a valid criticism since, as the following review points out, considerable subjectivity is involved in the assessment and evaluation of responses in most of the techniques described.

Two major approaches have been employed in detecting and quantifying involuntary responses to acoustic stimuli. These have, in the main, utilized visual observations of unconditioned and conditioned reflex behavior and the recording of unconditioned and conditioned electrophysiologic activity.

The most relevant of such procedures are described and discussed under appropriate headings below.

Overt Responses

The auropalpebral (APR), Moro and orienting reflexes

---

[19] R. Goldstein, "Electrophysiologic Audiometry," in Modern Developments in Audiology, ed. J. Jerger, (New York: Academic Press, 1963).

have been terms most widely applied in describing the involuntary, overt response modes employed in the assessment of hearing acuity. There appears to have been considerable latitude taken in assigning various observable behaviors under these reflex categories. In some cases this has been acknowledged by prefixing the classification with the term quasi. In general, however, auropalpabral has been used to refer to eye blinks and a widening, squinting or flickering of the eyelids; Moro to a sudden contraction of the limbs or movement of the body and orienting to arousal or cessation of activity, and lateralization toward or away from the source of the stimulus.

Considerable interest has been directed toward the use of these innate or unconditioned reflex responses to sound in the development of infant screening tests. In the mid 1940's Froeschels and Beebe advocated a simple test to be used to detect hearing impairment in the newborn.[20] Thirty-three infants under ten days of age were tested using a set of whistles, ranging over a scale of six octaves, and four tuning forks. They reported that only two infants failed to respond to the whistles while none responded to the forks. The most common response observed was the APR. These results

---

[20]E. Froeschels and H. Beebe, "Testing the Hearing of Newborn Infants," Archives of Otolaryngology, (1946), pp. 710-714

were purported to indicate that a complex rather than a pure tone was the more effective stimulus although there was little control over the intensity of the two stimuli.

Richmond, Grossman, and Lustman were among the first to attempt to control the intensity of the stimulus employed in infant testing.[21] This involved a cowbell that generated a SPL of 113 dB at the infant's ear when activated approximately one inch away. Their total population of forty-six infants responded with either an APR or startle response. They also reported that the best behavioral state for testing was light sleep while active crying, deep sleep, and active nursing were the poorest states.

One of the better controlled and more comprehensive of the early infant studies was conducted by Wedenberg.[22] An audiometer was employed with an intensity output range from 50 to 115 dB SPL measured slightly over two inches from the loudspeaker. Test frequencies available were 500, 1000, 2000, 3000 and 4000 cps. The purpose of the study involving 20 infants, was to determine threshold levels for APR responses at each of the available frequencies and threshold levels

---

[21]J. Richmond, H. Grossman, and S. Lustman, "Hearing Test for Newborn Infants." Pediatrics, 11 (1958), pp. 634-638

[22]E. Wedenberg, "Auditory Tests on New-born Infants." Acta Otolaryngologica, 46 (1956), pp. 446-461

for startle or orienting (which he classified as waking) responses at frequencies of 500 and 3000 cps. It was found that deep sleep and high activity levels in the awakened state tended to lower response reliability. This was overcome by spacing test sessions over several days with stimulus presentations contingent upon optimal response states. Under these conditions results showed that APR thresholds ranged between 105 and 115 dB SPL. Intensity levels necessary to elicit an awaking response to a 500 or 3000 cps tone, of two to five seconds duration, presented intermittently for a maximum test time of one minute were found to be approximately 75 dB SPL if the infant was in deep sleep and 55 dB SPL if in light sleep.

Hardy, Dougherty, and Hardy[23] tested 2000 infants under two days of age employing a wooden clacker with a broad frequency range, five milliseconds duration, and 65 dB SPL twelve feet from the infant. Approximately 98% gave a startle reflex or some modification of it. The 2% who failed the test fell into two general categories. One group consisted of premature infants who eventually responded normally and the other group was made up of infants who were fed just prior to the test and were difficult to arouse.

---

[23]J. Hardy, A. Dougherty, and W. Hardy, "Hearing Responses and Audiologic Screening in Infants." Journal of Pediatrics, 55 (1959), pp. 382-390

A study of similar design involving 2000 infants within the first hour after birth was carried out in 1960 by Froding.[24] The test stimulus was a gong with a broad frequency spectrum and intensity of 125 to 133 dB SPL. A positive APR was elicited from 96.1% of the test population. Sixty-six infants (3.3%) failed to respond with an APR and 0.6% gave doubtful responses. A follow up study on these children at an age when they could be tested by conventional means indicated that seven of those who had given positive reflex responses had losses ranging from 40 to 70 dB.

A mass screening technique for neonates was developed by Downs and Sterritt. Since mass screening on a large scale would involve considerable cost in professional man hours, the authors advocate the use of trained volunteers from service clubs and hospital auxillaries. The stated goal of such screening was:

> "to identify new born infants who may have a congenital hearing loss of severe degree (65 - 100 dB average), so that their physicians will be alerted to look for hearing problems that can be treated or habilitated early."[25]

Two instruments were developed for use in this technique. The Vicon Apriton produces a narrow band white noise peaking

---

[24]C. Froding, "Acoustic Investigation of Newborn Infants." Acta Otolaryngologica. 52 (1960), pp. 31-40

[25]M. Downs and G. Sterritt, "A Guide to Newborn and Infant Hearing Screening Programs." Archives of Otolaryngology, 85 (1967), pp. 15-22

at 3000 cps with most energy between 2800 and 3800 cps. This instrument also produces a wide band white noise (100-10,000 cps frequency range). Intensity outputs for both signals are 70, 80, 90, and 100 dB SPL four inches from the speaker. The second instrument is the Rudmose Warblet 300 which emits a warbled 3000 cps tone at intensities of 80, 90, 100 dB SPL ten inches from the speaker. Both instruments were reported to be equally effective. It was stated, however, if the child is asleep when tested, the Warblet elicits slightly greater response intensity.

A highly structured test procedure was advocated. Testing should be scheduled approximately 45 minutes prior to feeding to ensure the infant is in an optimum response state. Responses to be observed are categorized as eye blink, eye widening, cessation of activity, head turn, sucking, and movement of some part of the body. Intensity of responses are to be rated as follows: (1) No response; (2) Questionable; (3) Slight but definite; (4) Strong; and (5) Paroxysmal (Moro response). The infant's response must be rated as "3" or greater for the infant to pass the test. Prior to presentation of the stimulus, the pretest state of the infant - awake, sleeping, drowsy, active - is reported.

The 3000 cps stimulus is presented to the infant at 90 dB SPL. Two or more observers in addition to the tester immediately record the type and intensity of the observed response. If no response is observed initially, up to six presentations can be given before terminating the test. After a later retest if response intensity is rated below "3", the infant is marked as "suspect" and is scheduled for weekly evaluations by an audiologist.

Using this technique the authors screened 10,000 infants and identified 150 as failing the test. They experienced considerable difficulty in obtaining follow up on this "suspect" group and were able to confirm hearing loss in only four infants.

Since the instrumentation employed in this study has been made commercially available, infant screening programs have been initiated in a number of hospitals throughout the country. However, the apparent inadequate test/retest reliability of such procedures and the lack of sufficient follow-up data to establish the validity of these techniques has caused considerable professional concern. This concern was given official sanction through the following joint committee statement issued by the American Academy of Ophthalmology and Otolaryngology, the American Academy of Pediatrics, and the American Speech and Hearing Association:

> In recognition of the need to identify hearing impairment as early in life as possible, auditory screening programs have been implemented in new born nurseries throughout the country. Review of the data from the limited number of controlled studies which have been reported to date has convinced us that the results of mass screening programs are inconsistent and misleading.

> To determine whether mass screening programs for new born infants should indeed be instituted, intensive study of a number of variables is essential. These should include stimuli, response patterns, environmental factors, status at the time of testing, and behavior of observers. Furthermore, confirmation of results obtained in the nursery must await data derived from extended follow-up studies which involve quantative assessment of hearing status.
>
> In view of the above considerations and despite our recognization of the urgent need for early detection of hearing impairment, we urge increased research efforts, but can not recommend routine screening of new born infants for hearing impairment.[26]

While the use of such techniques with infants has been somewhat circumscribed by the variables described above, they have been employed most extensively as a means of obtaining some index of auditory sensitivity with young or noncommunicative children who appear untestable because of their failure to respond to conventional testing procedures. This type of informal assessment has evolved from the use of mechanical noisemakers to the utilization of electronically controlled sound stimuli presented through loud speakers and is popularly referred to as Behavior Observation Audiometry (BOA).

In 1944 Ewing and Ewing reported the results of an extensive study of the behavioral responses of children to a

---

[26]American Academy of Ophthalmology and Otolaryngology, American Speech and Hearing Association, and American Academy of Pediatrics, "Joint Committee Statement On Infant Hearing Screening," ASHA, 13 (1971), p. 79

variety of mechanically produced and voice sounds.[27] They described and classified responses relative to hearing status, age, and stimulus employed and suggested a testing technique based on distraction of the child prior to the presentation of the stimulus.

Several investigators followed the Ewings' lead in utilizing mechanical noisemakers with a distraction technique. Hardy et al extended their neonatal testing techniques to children up to one year of age employing sound stimuli produced by a wooden clacker along with bells, rattles, squeakers, and unvoiced consonants.[28] Response behavior most often elicited were those classified under the orienting reflex.

These approaches involved the use of an assistant to the tester who could focus the child's attention away from the sound source. The additional personnel was thought to be a limiting factor by Murphy who suggested a variation of this distraction technique which could be carried out by a single tester.[29] He suggested the use of rattles held in both

---

[27] I. Ewing and A. Ewing, "The Ascertainment of Deafness in Infancy and Early Childhood." Journal of Laryngology and Otology, 59 (1944), pp. 309-333

[28] J. Hardy, A. Dougherty, and W. Hardy, "Auditory Screening of Infants." Annals of Otology, Rhinology, and Laryngology, 71 (1962), pp. 759-766

[29] K. Murphy, "Ascertainment of Deafness in Children."

hands of the tester who was positioned directly in front of the child and who then gradually separates the rattles in an arch toward the child's ears. When the child visually tracts one of the rattles the other is shaken to elicit an orienting response.

The limitations inherent in the use of mechanical noisemakers have been pointed up by numerous audiologists. Barr described the difficulties inherent in such an approach and some of the sources of error which would invalidate the results.[30] Primary among these were difficulty in specifying the intensity level of the sound stimuli due to variations in the strength of production, distance from the child's ear, and difficulties in specifying the frequency range presented and the intensity level at each frequency present.

Since it is possible to control these variables through the use of electronic equipment available in most audiology clinics and since BOA procedures could provide some gross but useful information regarding the hearing sensitivity of difficult-to-test children, this approach has generated considerable research activity directed toward refining and extending the technique and has become an accepted and widely used technique.

Audecibel, 14 (1966), pp. 89-93

[30]Barr, loc. cit.

Ewertsen followed the distraction technique described above and controlled the stimuli by placing small loud speakers in the stomachs of identical teddy bears.[31] The bears are then to be held in front of the child and slowly separated with the acoustic stimuli presented from the bear the child has not visually followed. He suggests the use of warbled tones to overcome the problems of standing waves, created by pure tones. In addition he feels they more easily attract the child's attention.

A technique called "Free Field Startle Response Audiometry" was described by Suzuki and Sato.[32] Four loud speakers, arranged in a square two meters per side, are mounted on a ceiling of a test room. The child is seated at a midpoint directly beneath the speakers. Recorded cows mooing, cocks crowing, and cuckoos singing are presented for ten seconds each with two seconds between individual stimuli and three seconds between each series. The sound is presented through one of the front speakers or through the speaker on the side of the child's better ear, if this is known. Intensity is increased until an observable response occurs and then the sound is immediately switched to a different speaker. The

---

[31]H. Ewertsen, "Teddy-bear Screening Audiometry for Babies." Acta Otolaryngologica, 61 (1966), pp. 279-280

[32]T. Suzuki and I. Sato, "Free Field Startle Response Audiometry." Annals of Otology, Rhinology and Laryngology,

authors state the most frequent response for children under one year of age is eye movement and over one year of age is localizing. Using this technique, 181 normal children between 18 days and three years six months were studied. The results are summarized in the following table:

TABLE 2. Summary of results obtained using free field startle response audiometry as reported by Suzuki and Sato, 1961.

Age (months)	Mean response in dB*
5	36.6
6-11	26.0
12-17	18.6
18-23	13.8
24+	10.2

*dB re adult threshold for same condition

The decibel levels were relative to the average subjective thresholds of adults for the same sounds, from the same source, and in the same position as the children. The authors recognized the limitations of using such a broad band sound but found that pure tones were not suitable since

70 (1961), pp. 997-1007

they did not elicit reflex responses at such low intensity levels.

DiCarlo and Bradley tested 50 children between the ages of ten months and three years who had been referred for suspected hearing loss employing a "simplified auditory test" that had been developed to meet the criteria of short duration, easy administration and sufficient reliability to be practical as a screening technique.[33] Instrumentation for the test involved four speakers placed in a test chamber with 90° azimuth difference. Music, recorded pure tones ranging from 250 through 6000 cps, white noise, and recorded voice were employed as stimuli. These were presented in random order, with short duration using a descending attenuation technique with the initial presentation at 90 dB. Although the authors did not report the results of the original test, the results of follow up testing when the children were four or five years old indicated 31 had normal hearing and 19 had a hearing loss. It was also reported that two children, under the age of two years, failed to respond with any form of the orienting reflex, although shown to have normal hearing on the follow up test. They were subsequently diagnosed as mentally retarded plus central nervous system disorder. These results

---

[33]L. DiCarlo and W. Bradley, "A Simplified Auditory Test for Infants and Young Children." Laryngoscope, 71 (1961), pp. 628-646

tend to point up the fact that the absence of a reflex response to sound can be due to disorders other than hearing loss.

The difficulty in differentially assessing the absence of reflex response to sound was pointed up in a conference on the audiological assessment of the retarded held in Parsons, Kansas, in 1965. A panel of audiologists agreed that data obtained through sound field evaluations employing BOA techniques can be useful, the audiologist must be continually mindful of interferring variables. Relative to reflex behavior they stated:

> There are critical periods in the life of the developing organism at which times reflexive responses are expected and normal. However, in many central nervous system (CNS)-impaired youngsters this time table of reflexes may be significantly altered and of importance to the neurologist in his assessment of the child. The possibility of these developmental alterations confounding the audiological assessment should not be underestimated.[34]

In addition to the above limitation, as previously mentioned, the report of a conference on current practice in the management of deaf infants pointed out the inadequacies of BOA procedures in differentiating between conductive and sensori-neural impairments and being limited to providing information only on the better ear.[35]

---

[34]L. Lloyd and D. Frisina (eds), _The Audiologic Assessment of the Mentally Retarded_ (Parsons State Hospital and Training Center, 1965)

[35]McConnell, _loc. cit._

A number of investigators have attempted to overcome the problems of eliciting reflex response to sound through the use of respondent conditioning techniques. While a majority of these studies have involved the conditioning of electrophysiological activities, some interest has been directed toward observable reflex behavior with varying degrees of success.

Galambos, Rosenberg and Glorig suggested conditioning a pure tone to an eye blink response elicited by directing a sudden puff of air to the child's eye.[36]

A more conveniently administered and instrumented technique called "conditioned orientation reflex audiometry (COR)" was reported by Suzuki and Ogiba.[37] This procedure utilized the lateralizing or localizing aspect of the orienting response to a visual stimulus. Instrumentation consists of two plastic dolls which can be illuminated from inside the bodies placed on loud speakers on opposite sides of the test room. The test procedure consists of placing the child midway between each doll and speaker and presenting a tone

---

[36]R. Galambos, P. E. Rosenberg, and A. Glorig, "The Eyeblink Response as a Test of Hearing." Journal of Speech and Hearing Disorders, 18 (1953), pp. 373-378

[37]T. Suzuki and Y. Ogiba, "Conditioned Reflex Audiometry: A New Technique for Pure Tone Audiometry in Children Under Three Years of Age." Archives of Otolaryngology, 74 (1961), pp. 192-198

from one speaker, followed within one second by the illumination of the doll atop it. This is followed by the same sequence to the opposite speaker. The authors state that after three or four repetitions conditioning is usually established and the child turns to the dolls when the tone is presented alone. Appropriate responses are then reinforced by illumination of the doll. As responses occur, the intensity of the tone is decreased in 5 dB steps until no response is obtained. This level is considered to be threshold. To test the efficacy of this procedure, the authors tested 115 children under the age of three years who were thought to have normal hearing. The table below summarized the results of this investigation:

TABLE 3. Successful and unsuccessful tests on infants using conditioned orientation reflex audiometry as reported by Suzuki and Ogiba, 1961.

Age (years)	Successful	Partially Successful	Failure	Total
Less than 1	5 (38.5%)	0	8	13
1 to 1-5	16 (84.2%)	3	0	19
1-6 to 1-11	12 (80.0%)	3	0	15
2 to 2-5	17 (73.9%)	5	1	23
2-6 to 2-11	28 (96.6%)	1	0	29

It can be seen from these results that over 80% of normal hearing children between the ages of one and three years could be successfully conditioned under this procedure. However, since this is merely an extension of the general BOA procedures it suffers from the same deficiencies of differential diagnosis. In addition, since the test population included normal hearing children, the validity of the technique is still to be established with a handicapped population.

In an attempt to improve the diagnostic capabilities of the COR procedure, Reddell and Calvert developed a modification which they called "conditioned audio-visual response audiometry."[38] They employed a single luminous toy placed on a wall of the test room at a 45° angle lateral to the child's forward vision. A distraction technique is employed with an assistant holding the child's attention on some object or activity. Tones are presented through earphones on a bone conduction transducer and is immediately followed by illumination of the toy. They also point out that using this technique it is possible to obtain additional diagnostic information in the form of speech detection thresholds. In addition, the authors reported test-retest data on 20 children between

---

[38]R. Reddell and D. Calvert, "Conditioned Audio-Visual Response Audiometry." The Voice, 16 (1967) pp. 52-57

the ages of 15 to 47 months who were tested initially at a diagnostic clinic by means of CA-VR audiometry and retested with conventional audiometry after a mean lapse time of 11 months. Although the authors reported mean difference scores for the speech frequencies of 9.7, 13.2 and 6.2 which they felt was within the "clinical error standard" of 15 dB, the individual differences for the various frequencies ranged from a minus 30 to a plus 30 dB.

Covert Responses

Considerable interest and research has been directed toward the development of instrumentation for recording the electrophysiologic activity of the autonomic nervous system (ANS) and changes in the ongoing electrical activity of the central nervous system induced by acoustic stimuli. A number of ANS functions, such as heart rate, pulse volume, pupil dilation and constriction, respiration rate, and sweat gland activity have been evaluated as potential response vehicles of hearing. To date only changes in sweat gland activity have had any widespread clinical application.

Sweating occurs in response to thermal changes relative to the normal range of body temperature and to stimuli that induce a generalized emotional response. The latter has been most widely employed as a response mode in audiology. Most popular methods of measuring changes in sweat gland activity have involved instrumentation for recording the degree of resistance offered by the skin in conducting an externally applied electrical current. Since there is an inverse relationship between skin resistance and the degree of sweating, a measurement of the absolute and relative changes in electrical resistance provide an indirect measure of sweat gland activity. Goldstein and Derbyshire suggested the term electrodermal response (EDR) to be applied to identify the electrical manifestation of induced changes in sweat gland activity.[39]

The principles underlying EDR and the procedures most commonly employed in utilizing such responses in audiometry have been well stated by Newby:

> The skin of an individual has a resistance to the flow of electric current. As this resistance increases the amount of current that can flow across the skin decreases, and vice versa. Under circumstances of quiet and relaxation, an individual skin resistance remains relatively consistent. Stimuli that excite any kind of an emotional response, however, will cause the resistance of the skin to decrease. The increase flow of current resulting from this resistance drop may be amplified and displayed on a graphic recorder. The technique of utilizing the skin response in hearing testing is to condition the patient to respond to sound as an emotion-producing stimulus. This is done by pairing the test stimulus from an audiometer with a mild electric shock, but strong enough to produce momentary discomfort. The shock is sufficient to

---

[39]R. Goldstein and A. J. Derbyshire, "Suggestions for Terms Applied to Electrophysiologic Tests of Hearing." Journal of Speech and Hearing Disorders, 22 (1957), pp. 696-697

induce an emotional response, which causes a momentary drop in skin resistance. During the conditioning process, a spurt of tone and an almost simultaneous shock are presented to the patient, alternated in a random fashion with the presentation of tone alone. When conditioning has been achieved, presentation of tone alone will cause a drop in skin resistance as the patient anticipates the unpleasantness of the shock that might occur. Once the patient has been satisfactorily conditioned to tones above his voluntary thresholds, the "sampling" process being: attempts to specify the minimum hearing levels at which consistent EDR's can be obtained. Conditioning is maintained by random reinforcements, that is presenting the tone and shock together as in the conditioning process.[40]

One of the first, and most extensive, reports on the use of EDR audiometry with children, as described above, was made in 1949 by Bordley and Hardy.[41] Since that time, an instrument has been made commercially available to control and administer the acoustic and conditioning parameters and a number of studies have been carried out to determine the reliability and validity of EDR audiometry and the applicability of the technique as a measurement of hearing in young and difficult to test children.

Hardy and Bordley conducted one of the earliest studies comparing EDR audiometry with standard techniques on 59 children between the ages of 4 to 18 years with a mean age

---

[40]H. Newby, Audiology, (New York: Appleton-Century-Crofts, 1964), p. 155

[41]J. E. Bordley and W. G. Hardy, "A Study in Objective Audiometry With the Use of Psychogalvanometric Response." Annals of Otology, Rhinology, and Laryngology, 58 (1959), pp. 751-760

of 8.1 years.[42] Their results indicated that threshold differences between the two tests were within an acceptable plus or minus 5 dB in 67% of the cases at 500 cycles per second, 61% of the cases at 1000 cps, 73% of the cases at 2000 cps and 43% of the cases at 4000 cps. They considered these results to indicate that EDR audiometry was a reliable and valid technique that could be used in testing infants and very young children. While they experienced some difficulty resulting from a lack of cooperation, they felt that this was most prevalent among older children. These results have not, however, been substantiated in a number of similar comparative studies undertaken with young children.

Barr conducted a comparative study on 324 children under the age of seven years.[43] Play audiometry was compared with EDR to determine which was most applicable for use with children in different age groups. The results indicated that EDR audiometry was unsuccessful in about 1/2 of the children under 2 1/2 years of age. Barr's general conclusion was that the younger the child, the greater the risk of failure regardless of the method employed.

Similar results were obtained by Statten and Wishart

---

[42]W. Hardy and J. Bordley, "Special Techniques in Testing the Hearing of Children." Journal of Speech and Hearing Disorders, 16 (1951), pp. 122-313

[43]Barr, *loc. cit.*

who used EDR to test 101 children who could not be tested by play audiometry and 204 children who could be tested with a play technique.[44] In their results they reported that only 50% of the children were successfully conditioned with the EDR technique. Of this group that could be conditioned only 12% of the tests were considered successful, 72% partially successful, and 16% were unsuccessful.

O'Neill, Oyer and Hillis tested 50 children with a mean age of 65.05 months employing EDR and several forms of play audiometry.[45] They concluded that in no case was EDR found to be successful when play audiometry had been unsuccessful.

Although a number of studies appear in the literature which support the validity and reliability of EDR techniques with adult populations, it appears that there is general agreement in the reported research with children that it is no more effective than conventional or play techniques.

Changes in the ongoing electrical activity of the brain in response to auditory stimulation was first reported by Davis in 1939.[46] Considerable interest and research in the audiometric application of electroencephalographic (EEG)

---

[44]Statten, and Wishart, _loc_. _cit_.

[45]O'Neill, Oyer, and Hillis, _loc_. _cit_.

[46]P. A. Davis, "Effects of Acoustic Stimulation on the Waking Human Brain." _Journal of Neurophysiology_, 2 (1939), pp. 494-499

techniques followed this initial report.

One of the first uses of EEG for diagnosis of hearing impairment in children was reported by Marcus, Gibbs, and Gibbs.[47] Mechanical noisemakers were employed at "powerful intensity levels." A few children were tested while under either natural or induced sleep and the arousal response to these stimuli were recorded on the EEG tracings. The authors concluded that hearing was present in those children in whom the arousal response was elicited by acoustic stimuli.

A follow up study was carried out by Marcus on 71 children using the same stimuli and response criteria.[48] An arousal response to the auditory stimulation was obtained in 44 of the children tested. In addition, correlation between the arousal response and regular clinical evaluation was obtained in all but 8 of the 44 children studied. It was also reported that 19 of 53 patients with hearing loss had abnormal EEG tracings. On the basis of these results the author concluded that this arousal response to auditory stimulation provided an adequate index of hearing and could be used in the

---

[47]R. Marcus, E. Gibbs, and F. Gibbs, "Electroencephalography in the Diagnosis of Hearing Loss in the Very Young Child." Diseases of the Nervous System, 10 (1949), pp. 170-173

[48]R. Marcus, "Hearing and Speech Problems in Children: Observations and Use of Electroencephalography." Archives of Otolaryngology, 53 (1951), pp. 131-146

early diagnosis of hearing loss in infants and preschool children.

Derbyshire, Fraiser, Dermit, and Bridge employed an audiometer to produce the auditory stimuli presented to 22 preschool and hard of hearing children tested under induced sleep.[49] Their results indicated that thresholds obtained with the EEG method did not differ by more than 18 dB from standard audiometric findings.

Somewhat better results were reported by Withrow and Goldstein in a comparative study conducted with hard of hearing children.[50] Thresholds obtained by EEG procedures were found to be within 10 dB of thresholds obtained in a routine audiometric manner.

Thus the validity of the cortical-evoked response was demonstrated by a number of investigations. However, variability in the amplitude of the response coupled with high levels of recorded "noise" (ongoing brain activity) made visual detection of the response by the unaided eye a tenuous and often unsuccessful task.

---

[49]A. J. Derbyshire, _et. al._, "Audiometric Measurements by Electroencephalography." _Journal of Electroencephalography and Clinical Neurophysiology_, 8 (1956), pp. 467-478

[50]F. B. Withrow and R. Goldstein, "An Electrophysiologic Procedure for Determination of Auditory Thresholds in Children." _Laryngoscope_, 68 (1958), pp. 1674-1699

One of the earliest and most successful techniques for improving the "visibility" of the evoked response was suggested by Dawson.[51] This involved a photographic superimposition of short segments of the EEG trace following successive stimulation. An averaging process was involved based upon the assumption that the response was time-locked to the stimulus while the ongoing EEG activity would occur in random patterns relative to the same stimulus. Thus those parts of successive traces which were related to the stimulus were superimposed in a relatively exact manner while the ongoing EEG activity was randomly dispersed.

Although the superimposition technique was widely adapted and utilized, some discrepancy was noted among various investigators as to the effectiveness of the technique in improving the "visibility" of the response at or near threshold levels. For example: Suzuki and Asawa reported that they were able to detect evoked responses to pure tones, presented at threshold levels, in 20% of the records, at 30 dB above threshold in 65% of the records, and at 60 dB above threshold in 80% of the records.[52] However, Webb, Kinde, Webber, and

---

[51]G. D. Dawson, "A Summation Technique for the Detection of Small Evoked Potentials." Journal of Electroencephalography and Clinical Neurophysiology, 6 (1954), pp. 65-84

[52]T. Suzuki and I. Asawa, "Evoked Potential of Waking Human Brain to Acoustic Stimuli." Acta Otolaryngolica, 24 (1962), pp. 217-224

Beedle found that with such a visual detection technique it was not possible to utilize the evoked auditory response as a screening test with mentally retarded subjects when the pure tone stimuli were presented at a hearing level of 20 dB.[53]

A most important advancement in evoked response audiometry (ERA) came with the incorporation of the averaging computer into the instrumentation of this technique. While a number of different computers have been employed, they generally all function in a manner similar to that described by Price:

> A temporally varying voltage (such as the EEG) is made available as input. When the computer is triggered, a gating system sweeps across the inputs of a series of storage bins (memory) in a present time. This allows the fluctuating input voltage to be systematically sampled as a function of time relative to the onset of the sweep and each sample stored in the computers memory. If the onset of the sweep happens to be locked to the onset of the auditory stimulus, the voltage is sampled at set intervals following each auditory stimulus. For example, if the computer has 100 storage bins and the sweep is set for 500 milliseconds, a sample of the input voltage will be taken at 5 millisecond intervals for 1/2 second. Each sample will be stored in a separate memory bin. The opening and closing of the input to each bin is time-locked to the onset of the stimulus. Thus, all samples taken at a given time, relative to stimulus onset, are stored in the same bin and repeated samples result in an algebraic summation of the voltages occurring at that particular time following onset of the stimulus. Since the computer is summing

---

[53]C. Webb, et. al., "Procedures for Evaluating the Hearing of the Mentally Retarded." Cooperative Research Project No. 1731, U.S. Office of Education, 1964

> voltages, each sample either makes a voltage in the memory more positive or less positive depending upon whether the sample is positive or negative relative to the voltage already stored in the memory. After a given number of sweeps a display of the voltages stored in the memory bins shows the average voltage as a function of time.[54]

Much of the early research employing computer techniques was directed toward defining the parameters of the evoked response and delineating the effects of variations in stimulus parameters and subject variables.

While there is general agreement as to the shape and amplitude of the evoked response to sound, arbitrary variations in the polarization of recording equipment, and subsequently in the assignment of symbols to the various components of the wave form, have created some difficulty in relating the results of various investigators. In other words, depending upon the polarity of the recording system utilized, an upward deflection of the wave form may be described as the P1 component (first positive component) or as the N1 component (first negative component).

The averaged evoked response (AER) to auditory stimulation as described by Price, Rosenblut, Goldstein, and Sheppard and numerous other investigators is characterized by peak components occurring at 50, 85, 160 and 260 milliseconds

---

[54] L. L. Price, "Cortical-Evoked Response Audiometry." In, <u>Audiometry for the Retarded</u>, eds. R. Fulton and L. Lloyd. (Baltimore: The Williams & Wilkins Co., 1969) pp. 211-212

following the onset of the auditory stimulus.[55] Figure 1 below illustrates the triphasic wave form generally attributed to the AER with peak components labeled relative to a positive, upward, pen deflection:

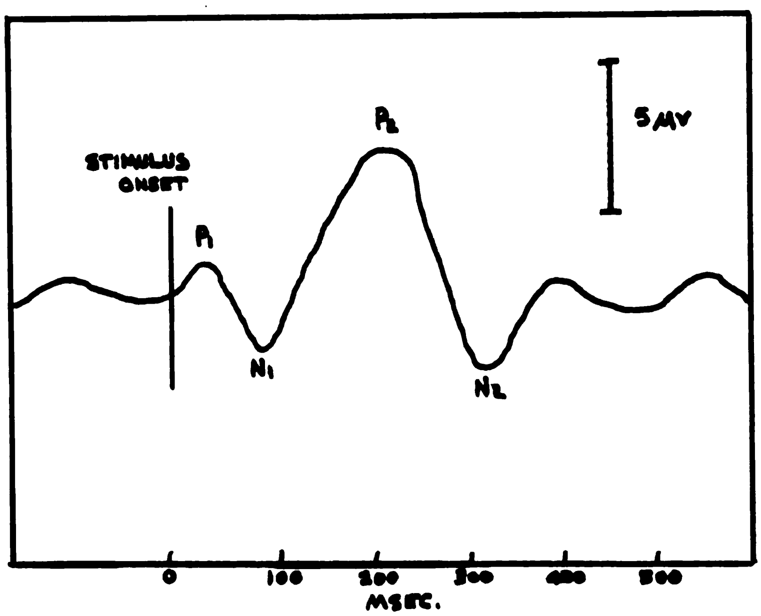

Figure 1---A representative AER waveform elicited by pure tones presented at sensation levels 60 dB or greater.

Although the above illustration is representative of the AER both the latencies and amplitudes of the various components are somewhat variable being influenced by stimulus parameters such as the type, rise time, intensity, and

---

L. Price, et. al., "The Averaged Evoked Response to Auditory Stimulation." Journal of Speech and Hearing Research, 9 (1966), pp. 361-370 repetition rate. Some of the variability in these response parameters results from subject-related factors such as age, physiological state, mental set, and CNS pathology.

Since the AER is recorded from the non-specific frontal cortex, it can be evoked by several types of sensory stimuli including auditory, visual and tactile. In order to assess hearing status it is of course necessary to use an auditory stimulus. Several different types of acoustic stimuli have been used in studies of the evoked auditory potentials. These have included clicks, pure tones, and speech. Williams and Graham found that it was easier to evoke a response to clicks than to pure tones.[56] They hypothesize that this is the case because the rise time of a click is shorter and thus the cortical activity evoked by a click is more diffuse than with a pure tone.

McCandless and Best found that when an averaging response computer was used, it was as easy to obtain an evoked response with pure tones as with clicks.[57] They also pointed out that when pure tone stimuli are used, it is possible to obtain

---

[56]W. G. Williams and J. T. Graham, "EEG Responses to Auditory Stimuli in Waking Children." Journal of Speech and Hearing Research, 6 (1963), pp. 57-62

[57]G. A. McCandless and L. Best, "Evoked Responses to Auditory Stimuli in Man Using A Summing Computer." Journal of Speech and Hearing Research, 7 (1964), pp. 193-202

information concerning the sensitivity of the auditory mechanism at different frequencies. In addition, they stressed that reference levels are available for pure tone stimuli and the attenuation of pure tones is more effectively accomplished with conventional devices.

Derbyshire, Palmer, Elliot, Cassidy and Kapple[58] and Walter, Aldridge, Cooper, McCallum and Cohen[59] have reported using words to evoke an auditory response and have attempted to relate such responses to the semantic content of the stimuli. However, speech has not been widely employed as an auditory stimulus since it is subject to some of the limitations described above and would, of course, have limited application with very young children.

Since pure tones are the preferred stimuli in threshold audiometry, a number of investigations have been directed toward determining the effects of frequency and intensity variations on the form of the evoked response. While McCandless and Best[60] reported that they were unable to observe any differences in overall pattern or latency of the evoked

---

[58] A. J. Derbyshire, et. al., "EEG Responses to Words." ASHA, 6 (1964), p. 392

[59] W. G. Walter, et. al., "Responses to Semantic Stimuli in the Human Brain." Journal of Electroencephalography and Clinical Neurophysiology, 19 (1965), p. 314

[60] McCandless and Best, loc. cit.

response to 250, 1000 and 4000 cps tones, Rapin et al[61] reported decreased amplitudes associated with stimuli of 6000 cps. In addition, they reported that the amplitude of the evoked response of the tones was regularly and definitely related to changes in stimulus intensity in an almost linear fashion. It was noted, however, that this relationship did not hold for all subjects. It was suggested that a change in psychophysical state might explain the variability. McCandless and Best reported similar results under test-retest conditions and observed that responses to the same stimuli were similar for a single subject, but different from subject to subject.[62] Similar results have been reported by other investigators such as Teas[63] and Davis and Zerlin[64] who reported that for most, but not all subjects, increase in the intensity of test tones resulted in an increase in

---

[61]I. Rapin, _et. al._, "Evoked Responses to Clicks and Tones of Varying Intensity in Waking Adults." _Journal of Electroencephalography and Clinical Neurophysiology_, 21 (1966), pp. 335-344

[62]McCandless and Best, _loc. cit._

[63]D. C. Teas, "Analysis of Evoked and Ongoing Electrical Activity at the Scalp of Human Subjects." _Journal of Speech and Hearing Research_, 8 (1965), pp. 371-387

[64]H. Davis and S. Zerlin, "Acoustic Relations of the Human Vertex Potential." _Journal of the Acoustic Society of America_, 39 (1966), pp. 109-116

the amplitude of the evoked response. It has also been observed by some investigators that the latency of the evoked response increases significantly at near threshold levels. In a study by Davis and Yoshie it was reported that the latencies of response components did not change to tones in the frequency range between 300 and 4800 cps when presentation levels were between 20 to 75 dB re ASA 1951 audiometric reference levels but that as the stimulus tone approached the patient's threshold level increased response latencies of up to 50 milliseconds were obtained.[65]

Because the averaging method of analyzing EEG responses to sound involves the relatively rapid presentation of sizeable numbers of acoustic stimuli, considerable research interest has been directed towards the possible adaptation of the evoked response under these conditions. However, Rose and Ruhn reported that responses to 0.1 msec. clicks at 20 and 40 dB SL showed no significant adaptation for signals presented during a single day or over six days.[66] In addition, Roeser and Price presented pure tone stimuli every 2.3 seconds for two hours and found that while there was a decrease in

---

[65] H. Davis and N. Yoshie, "Human Evoked Cortical Responses to Auditory Stimuli," Physiologist, 6 (1963), p. 164

[66] D. E. Rose and H. B. Ruhn, "Some Characteristics of the Peak Latency and Amplitude of the Acoustically Evoked Response." Journal of Speech and Hearing Research, 9 (1966), pp. 412-422

the amplitude of the responses as a function of time for about the first 30 minutes of stimulation (a total decrement of approximately 4 microvolts) the response remained present and detectable throughout a two hour session.[67]

However, several studies have demonstrated that stimulus repetition rate appears to be a critical variable in maintaining maximum amplitude of the evoked response. According to Davis, et al, when inter-stimulus intervals are shorter than 0.5 seconds the latency of the evoked response is altered and the amplitude reduced.[68]

McCandless and Best studied the effects of presenting stimuli in a range from 3 per second to 1 per 2 seconds and found that with a repetition rate faster than 1 per 2 seconds some components of the evoked response decreased in amplitude.[69] However, Davis and Zerlin found that while ten seconds are required for full recovery of the vertex potential, the greatest voltage output per minute can be recorded at a repetition rate of 1 per second.[70]

---

[67] R. Roeser and L. Price, "Effects of Habituation on the Auditory Evoked Response." ASHA, 10 (1968), p. 396

[68] H. Davis, et. al., "The Slow Response of the Human Cortex to Auditory Stimuli: Recovery Process." Journal of Electroencephalography and Clinical Neurophysiology, 21 (1966),

[69] McCandless and Best, loc. cit.

[70] Davis and Zerlin, loc. cit.

In spite of the subject and instrumentation variables, which appear to affect the elicitation and detection of the evoked response, there appears to be considerable research evidence to indicate that under carefully controlled conditions averaged-evoked response audiometry can provide a valuable adjunctive technique to the audiologists clinical armamentarium.

Rapin and Graziana stressed the importance of recognizing and dealing with the intra-subject response variability.[71] They found in a study of 18 normal babies that while 23% of the initial test runs at intensities between 50 and 79 dB did not yield detectable responses, they were able to reduce the no-response to less than 2% by repeating each series of stimuli four times.

Several studies have indicated that the amplitude of the evoked response and concomitantly its detectability are enhanced if the subject's attention is focused on the auditory stimuli. Gross, Begliter, Tobin and Kissin compared the evoked response to clicks while subjects read and while the subjects counted the clicks.[72] They found that the ampli-

---

[71] I. Rapin and L. Graziana, "Auditory-Evoked Responses in Normal, Brain Damaged and Deaf Infants." Neurology, 17 (1967), pp. 881-894

[72] M. Gross, et. al., "Auditory Evoked Response Comparison During Counting Clicks, and Reading." Journal of Elec-

tude of the evoked response was greater when the subjects directed their attention to the clicks by counting them. Similarly Davis observed that averaged evoked auditory response amplitude was increased when subjects were instructed to signal their detection of an increase in stimulus intensity and to do nothing when intensity appeared unchanged.[73]

At the present time, the clinical application of the AER technique has been primarily limited to the determination of hearing sensitivity. While there has been some indication from the work of Walter[74] that there is a particular wave form which has been labeled a "contingent negative variation" that is related to the integrative or associative cortical processes, the most widespread clinical interpretation of the evoked response related it only to a passive change in CNS activity resulting from sensory stimulation. It is, in fact, generally considered that an evoked response does not necessarily indicate that the subject is consciously aware of the presence of stimuli since such responses can be obtained

---

troencephalography and Clinical Neurophysiology, 18 (1965), pp. 451-454

[73]H. Davis, "Enhancement of Evoked Cortical Potentials in Humans Related to a Task Requiring a Decision." Science, 145 (1964), pp. 182-183

[74]W. G. Walter, "The Contingent Negative Variation: An Electrocortical Sign of Significant Association in the Human Brain." Science, 146 (1964), p. 434

while the subject is in a state of sleep. Consequently, thresholds obtained by AER techniques are not generally considered as an equivalent of psychophysical thresholds. However, a number of studies have been carried out which indicate that thresholds obtained by evoked response and psychophysical techniques are sufficiently close to permit the use of evoked responses to predict psychophysical thresholds.

Earlier studies employing visual detection techniques, such as those carried out by Derbyshire and McDermott[75] and Suzuki and Asawa, [76] reported that thresholds obtained by evoked response and standard audiometric techniques were within 20 dB of agreement. Employing averaging techniques, McCandless and Best reported that some evoked responses were obtained for a group of normal hearing adults tested at levels within 10 dB of psychophysical thresholds.[77] Similarly, Price, Rosenblut, Goldstein and Sheppard found that with normal hearing adults the main threshold difference between these techniques was 4.4 dB with a maximum difference of

---

[75]A. J. Derbyshire and M. McDermott, "Further Contributions to the EEG Method of Evaluating Auditory Functions." Laryngoscope, 68 (1958), pp. 558-570

[76]T. Suzuki and I. Asawa, "Evoked Potential of Waking Human Brain to Acoustic Stimuli." Acta Otolaryngologica, 48 (1957), pp. 508-515

[77]McCandless and Best, loc. cit.

10 dB.[78] A slightly wider range of variance was pointed up in a study by Cody and Bickford who compared thresholds obtained on 20 normal adults by means of conventional and evoked response audiometry.[79] Employing test tones of 500, 1000, and 2000 cps, they found that thresholds for 43% of the tones obtained by the two methods were identical, for 40% a 5 dB difference, for 12% a 10 dB difference and for 5% a 15 dB difference was obtained. Davis reported similar results in which he found a main difference of 0.1 dB between behavioral and evoked thresholds on 162 hearing impaired children and contrasted this to results with normal hearing adults in which it was necessary to go as high as 25 dB SL in order to elicit an evoked response.[80] Suzuki and Taguchi reported comparable variability in a study employing 40 infants under three months of age, 15 young children age 1 to 5 years, and ten adults who were all presumed to have normal hearing.[81] Their primary findings appear to be that

---

[78] L. Price, _et_. _al_., _loc_. _cit_.

[79] D. Cody and R. G. Bickford, "Cortical Audiometry: An Objective Method of Evaluating Auditory Acuity in Man." _Mayo Clinic Proceedings_, 40 (1965), pp. 273-287

[80] H. Davis, "Slow Cortical Responses Evoked by Acoustic Stimuli." _Acta Otolaryngologica_, 59 (1965), pp. 179-185

[81] T. Suzuki and K. Taguchi, "Cerebral Evoked Response to Auditory Stimuli in Young Children During Sleep." _Annals of Otology, Rhinology and Larynology_, 77 (1968), pp. 103-110

responsitivity improved with age since, for example, evoked thresholds were elicited for 30 dB SPL tones from 50% of the infants, from 67% of the young children and from 80% of the adults.

Rose and Rittmanic studied the relations between psychophysical and evoked response thresholds with 35 retarded males between the chronological ages of 21 and 51 years.[82] Subjects were selected on the basis of their ability to give voluntary responses to pure tone audiometry and their ability to pass a 30 dB hearing screening at a test frequency of 1000 cps which was also the stimulus used in the ERA testing. Their findings indicated that thresholds for the two techniques were identical for 17% of the subjects, within 5 dB for 54%, within 10 dB for 80% and within 15 dB for 86%. One subject failed to respond by ERA at any testing level and the remaining three subjects yielded ERA at 20, 25 and 30 dB SL.

Lentz and McCandless performed repeated AER and behavioral tests at one, three, six and twelve months.[83] Infants were categorized in three groups; normal, preterm and high risk, on the basis of birth weight. Those in the normal group

---

[82] D. Rose and P. Rittmanic, "Evoked Response Tests With Mentally Retarded." Archives of Otolaryngology, 88 (1968), pp. 495-498

[83] W. E. Lentz and G. A. McCandless, "Averaged Electroencephalic Audiometry in Infants." Journal of Speech and Hearing Disorders, 36 (1971), pp. 17-28

had birth weights in excess of 2500 grams, and those in the high risk group had birth weights less than 1500 grams. AER and behavioral tests were administered at each of the four test periods. These authors found a trend toward improved AER responsitivity with increasing age and found behavioral thresholds obtained at 12 months to be on the average of 20 dB lower than AER thresholds.

On the basis of the literature cited, it appears that the use of average evoked response audiometry in the hearing assessment of young and exceptional children is limited to an adjunctive rather than a primary source of diagnostic data. AER responsitivity appears to be the lowest and variability the highest with this group. Because of these factors and the concomitant difficulties in the interpretation of the results, the accuracy of the technique is dependent upon the training and experience of the tester and thus involves a considerable degree of subjectivity. Price was one of the first to point up this limitation and to caution against the widespread application of this technique as the answer in differential diagnosis with the young or difficult to test child.[84] He suggested that the subjectivity involved in the evaluation of AER responses could be reduced through the use of a no-response

---

[84] L. Price, "Evoked Response Audiometry: Some Considerations." Journal of Speech and Hearing Disorders, 34 (1969) pp. 137-141

baseline measurement against which responses would be evaluated. He advocated obtaining such controlled series at repeated intervals throughout the test by averaging a series of EEG samples when no auditory stimulus was presented. These are then considered as the noise level of the individual and stimulus samples are accepted as positive responses only when they are clearly different from the control.

CONDITIONING PROCEDURES

As previously pointed out, while some of the principles of respondent and instrumental conditioning have been employed throughout all of the historical stages of hearing test development, contemporary research and clinical interests have focused primarily on the stringent adherence to reinforcement principles. Unlike earlier approaches that have been classified as "play audiometry" and which were primarily concerned with a search for attractive or interesting response modes, the contemporary phase as described by Lloyd[85] has been concerned with such critical factors as the selection of an appropriate reinforcer, reinforcement contingencies, immediacy of reinforcement, reinforcement schedules and reinforcement shifting. He points out that sensitivity and the

---

[85]L. Lloyd, "Behavioral Audiometry Viewed As An Operant Procedure." *Journal of Speech and Hearing Disorders*,

use of these variables differentiates the skilled from the unskilled audiologist.

Basic understanding of the concept of reinforcement is a necessary prerequisite to the selection of an appropriate reinforcer. Most introductory texts which deal with reinforcement theory point up the circular or functional relationship between response and reinforcement and caution against the tendency to identify or accept particular stimuli (candy, pop, ice cream) or events (games, social praise, affection) as "universal reinforcers." These facts are emphasized by Smith and Moore in the following definition of reinforcement:

> Whether a stimulus event is reinforcing is an empirical question; if a stimulus event strengthens (increases the probability of) a class of responses or class of stimulus response connections by its presentation or by its termination, it is said to be a reinforcer.[86]

Tangible reinforcers have been the most popular choice in test development with young and handicapped children. This has been in part due to an interest in automating test procedures. A study reported by Meyerson and Michael described an operantly structured test program utilizing response instrumentation and mechanical dispensation of edibles and

---

31 (1966), pp. 128-136

[86]W. I. Smith and J. W. Moore, Conditioning and Instrumental Learning. (New York: McGraw-Hill Book Co., 1966) p. 13

trinkets.[87] Although they did not describe any pretest method which might have been used to determine the effectiveness or strength of their reinforcers, they implied that the use of over 300 separate items ensured the availability of a sufficient number of appropriate reinforcers.

A more individualized approach to determining appropriate reinforcement was described by Lloyd, Spradlin, and Reid.[88] Prior to initiating test procedures with MR children, a variety of solid foods were presented to each child in the hand of the examiner. If the child reached for any item the examiner closed his hand; and if the child struggled to obtain it, this was considered to be an effective reinforcer. A similar "closed fist" test procedure could be employed with small non-edible objects. In the case of large toys and/or games a short period of pretest observation in an area where a number of such items are available to the child might be an effective means of measuring reinforcement properties.

The temporal relationship between response and reinforcement is another critical factor to be considered in the application of these principles to audiometry. The acquisi-

---

[87]L. Meyerson and J. Michael, "Assessment of Hearing by Operant Conditioning Techniques." In Proceedings International Congress on Education of the Deaf. (Washington, D. C.: Gallaudet College, 1963)

[88]L. Lloyd, J. Spradlin and M. Reid, "An Operant

tion rate of a conditioned operant depends upon the time interval between the execution of the response and presentation of the reinforcement. This principle, demonstrated many times in the laboratory with animals and human subjects, is typified by Perins' demonstration that the acquisition of bar-pressing behavior by rats is fastest when reinforcement is immediate, and with a 30 second delay between response and reinforcement the conditioned response is not acquired.[89] The import of this principle on hearing test design has been a tendency on the part of researchers to develop and espouse the use of elaborate mechanical reinforcement dispensers. However, mechanical dispensers are expensive and, while a necessary part of a totally automated program, are not essential to successful conditioning.

Considerable care must be exercised in structuring and controlling reinforcement contingencies. In this regard, Lloyd suggests that a descending method of stimulus presentations to be made at levels which are assumed, on the basis

---

Audiometric Procedure for Difficult-to-Test Patients." _Journal of Speech and Hearing Disorders_, 33 (1968), pp. 236-245

[89]C. T. Perin, "A Quantitative Investigation of the Delay-of-Reinforcement Gradient." _Journal Experimental Psychology_, 32 (1943), pp. 37-51

of clinical observation, to be above the subject's threshold.[90]

Fulton and Spradlin pointed up the need to guard against establishing accidental or inappropriate response contingencies.[91] They stressed the machine programs are especially vulnerable to the unintentional reinforcement of responses which are initiated just prior to the tone presentation but which carry over into the stimulus period. They suggest that this deficiency can be overcome by arranging the apparatus so each response occurring during the no-stimulus period delays the occurence of the stimulus period by a few seconds. In addition, they stress that the equipment should be set up so there is a fraction of a second delay between the onset of the tone and the availability of reinforcement to prevent the subject from receiving reinforcement before he has heard the tone. The manual presentation of stimuli and reinforcers permits greater versatility in anticipating inappropriate responses and interjecting such delays. For example, when the response mode involves a protracted play activity, it is possible for the tester to reinforce only those responses that are initiated after the tone is presented and to continue the tone presentation up to the moment of reinforcement.

---

[90] L. Lloyd, _loc_. _cit_.

[91] R. T. Fulton and J. E. Spradlin, "Operant Audiology With Severely Retarded Children." _Parsons Demonstration Project Report No. 90_. (Kansas: Parsons State Hospital

Since the advent of the Skinner box and Skinners'[92] purported accidental discovery of the effects of partial reinforcement considerable research effort has been directed toward determining the behavioral effects of various methods of administering reinforcement and some basic principles have been established. As a result, it is well known that an operant response is most effectively established under a continuous schedule of reinforcement. However, such a schedule is not the most effective means of maintaining high rates of responding; and because of the high number of reinforcers administered over time, it is most susceptible to satiation. Among the most commonly employed partial or intermittent schedules of reinforcement the highest rates of responding occur under the variable ratio schedules along with the greatest resistence to extinction. As a consequence, this is the schedule most often selected in operant audiometry. The average ratio of interval to be employed is, in the case of the young or exceptional child, one which must be determined empirically. The importance of applying appropriate schedules of reinforcement was indicated by Lloyd:

---

and Training Center, 1968)

[92]B. F. Skinner, "A Case History in Scientific Method." In Koch, S. (ed.), Psychology: A Study of a Science, Vol. 2 (New York: McGraw-Hill, 1959).

> Once the pattern of responding to sound is established the clinician usually reduces the frequency and amount of reinforcement. This reduction during the testing session results in greater testing efficiency. The skilled audiologist attempts to apply a sufficient schedule and amount of reinforcement to maintain a high rate of responding, but does not waste time administering excessive reinforcement, which is not only inefficient in terms of measurements per unit of test time but also increases the chance that the subject will be satiated and cease responding.[93]

The concept of stimulus generalization has considerable import for operant audiometry since conditioning is initially established for a single test tone and must then be transferred to the other test frequencies. While on the basis of the generalization phenomenon it might be assumed that sufficient similarity exists between the audiometric test tones employed for the transfer to be somewhat automatic. Spradlin, et al caution against changes in excess of one octave in either direction.[94] In addition, Meyerson and Michael suggest that after conditioning is established to a single frequency and before threshold testing begins that a period of "generalization training" be undertaken and conditioning established for each of the frequencies to be tested, from low to high.[95]

---

[93]L. Lloyd, loc. cit.

[94]J. E. Spradlin, B. J. Locke and R. T. Fulton, "Conditioning and Audiological Assessment." In Fulton, R. T. and Lloyd, L. (eds.), Audiometry for the Retarded (Baltimore: The Williams and Wilkins Company, 1969)

[95]Meyerson and Michael, loc. cit.

The study reported by Meyerson and Michael was one of the earliest studies in which operant learning theory principles and behavior modification techniques were applied in a highly structured procedure that was automatic, except for the changing of frequency and intensity of the stimulus.[96] Employing automated instrumentation which they called the "Audiomat" they tested a small number of severely and profoundly retarded children and concluded that the most effective procedure involved reinforcement of both tone-on responses and tone-off responses. They reported their method was successful with all of the children tested and only three thirty minute sessions were necessary to obtain threshold measurements for each of the children. The first session was directed toward developing the differential-discrimination of tone-on/tone-off and bringing the appropriate responses under the control of light and sound stimuli. The second session shaped the responses to sound alone and introduced a generalization training of all frequencies. During the third session intensity levels were gradually reduced and thresholds were identified as the level where the childs' previously accurate discrimination broke down. These authors stated that the "Audiomat" would be of special value with

---

[96]Meyerson and Michael, Ibid

difficult-to-test cases such as those resulting from immaturity, speech or language difficulties and brain damage. They also pointed up the need for widespread clinical use of such procedures to avoid the necessity of dismissing the difficult-to-test child with vague instructions to the parents to try to sensitize him to sound or to bring him back for another test when he is older.

An operantly structured hearing test technique which utilized edible reinforcers was employed by LaCrosse and Bidlake to evaluate the hearing of mentally retarded children whom they had found to be untestable by the "classical clinical method."[97] They reported they were able to successfully evaluate the hearing of 92% of their test population of 384 children.

Lloyd, Spradlin and Reid tested the hearing of 50 profoundly retarded children who were unable to respond to conventional forms of audiometry. They utilized an operantly structured test procedure which they referred to as tangible reinforcement operant conditioning audiometry (TROCA).[98] A button-pushing response mode was employed and reinforcement

---

[97]E. L. LaCrosse and H. Bidlake, "A Method to Test the Hearing of Mentally Retarded Children." Volta Review, 66 (1964), pp. 27-30

[98]Lloyd, Spradlin, and Reid, loc. cit.

was automatically machine-dispensed. The test procedure consisted of a bilateral screening at 20 dB ISO for the frequencies 250-8000 cps, followed by threshold testing for any child who failed to respond appropriately to any one of the seven test frequencies. The authors reported that they were able to successfully test 78% of the children using these procedures and, in addition, were able to employ masking in the better ear to identify three cases of unilateral losses. They also stated that "tone-control" was established with 84% of the children, but they did not elaborate the reasons for their failure to obtain pure tone data on the additional 16%. A range of from four to fifty sessions were necessary to obtain definitive audiometric information on the children, with sessions varying in length from ten to twenty minutes. The authors reported that their study demonstrated the validity of the TROCA procedure on the basis that reasonable audiometric configurations were obtained, their results were in agreement with other data such as otologic findings, and the fact that they were able to identify three cases with unilateral hearing impairment. They state this procedure has application to other difficult-to-test populations, such as infants, and reported that they were able to obtain thresholds measuring 20 dB or better (ISO) for the octave frequency range 250-8000 cps with three normal hearing infants aged 7, 15 and 18 months.

Bricker and Bricker used a group of 36 low-functioning, mentally retarded children to investigate temporal and rate aspects of reinforcement as well as the simultaneous presentation of light and tone stimuli.[99] The results of their research indicated there was no significant difference in the rate or acquisition of tone control as a result of prior light discrimination or rate building schedules of reinforcement. The authors stated that the results of their investigation added support to the use of operant audiometry as an effective technique for assessing the hearing of low functioning retardates since they were able to establish reliable thresholds for the total population. On the basis of the successful results of their research these same authors published a step-by-step program for using operant audiometry with low functioning children.[100] They advocate the use of such a procedure on the basis that it involves a sequence of reinforcement and behavior control techniques that can be useful in subsequent speech and language training. They also stress that the successful use of operant audiometry

---

[99] W. Bricker and D. Bricker, "Four Operant Procedures for Establishing Auditory Stimulus Control With Low-Functioning Children." <u>American Journal of Mental Deficiency</u>, 73 (1969), pp. 981-987

[100] W. Bricker and D. Bricker, "A Programmed Approach to Operant Audiometry for Low-Functioning Children." <u>Journal of Speech and Hearing Disorders</u>, 34 (1969), pp. 312-319

is dependent upon a strict adherence to and an understanding of the principles involved in determining an appropriate reinforcer, non-reinforcement of inappropriate behavior, reinforcement of successive approximations, discriminative stimuli, and schedules of reinforcement.

Spradlin, Locke, and Fulton,[101] tested six severely mentally retarded children, between the ages of 8 and 16 years using procedures described by Lloyd, Spradlin and Reid,[102] with the single variation that each subject received ascending and descending schedules of tone presentation, alternately over six sessions. Thresholds obtained over all experimental sessions did not vary more than 10 dB which the authors interpreted to be essentially equivalent thresholds for ascending and descending procedures.

It is apparent from this review of the literature regarding the use of operant principles and the assessment of hearing that the major research interest has been confined to the application of these techniques to the assessment of hearing problems in mentally retarded children. The successful application of such techniques with individuals functioning at a profoundly retarded level points up the need to extend this approach to other difficult-to-test populations.

---

[101]Spradlin, Locke and Fulton, loc. cit.

[102]Lloyd, Spradlin and Reid, loc. cit.

Lloyd, Spradlin, and Reid suggested that the use of an operant discrimination paradigm may be viewed as an effective way of communicating to the non-verbal child the task he should perform without the confusions and misconceptions of verbal instructions.[103] Since the literature indicates that in some cases the conditioning process involved protracted time involvement, in the form of repeated and lengthy sessions, this may represent a factor which has inhibited widespread clinical application. This may be explained on the basis that most clinical settings have stringent limitations on space and personnel available for audiological evaluations.

SUMMARY

Since most hearing test procedures were developed for and standardized on adult populations, they have had limited efficacy in testing young or handicapped children. Over the past forty years there has been increasing interest and research directed toward the development of procedures and techniques that would facilitate the testing of children. The clinical and research literature in this area can be conveniently grouped according to three major periods or stages of developmental interest.

---

[103] Lloyd, Spradlin and Reid, _loc_. _cit_.

The first of these periods has been classified under the heading of "Play Audiometry". It was characterized by a major effort to adapt conventional test procedures through the use of toys and games which would hold the interest of children and provide an appropriate degree of motivation to ensure their participation. A variety of mechanical and electrical gadgets were incorporated or developed, as well as pictures, puppets, film strips and slides. While these approaches gained some initial popularity, several studies carried out to evaluate the effectiveness of the procedures most widely used found that they did not facilitate testing below the age of four years and were relatively ineffective with children who exhibited psychological handicaps.

A second period of test development has been classified under the heading of "Objective Techniques". During this stage the interest in test development had turned toward the utilization of reflex responses to sound. Two major approaches were employed. The first utilized visual observations of unconditioned and conditioned reflex behavior. The eye blink, startle and orienting reflexes have been most widely employed in testing infants. Specific instrumentation and procedures have been developed for infants screening programs and a number of such programs, some of sizeable scale, have been initiated in hospital nurseries throughout the country. In

1971, however, the report of a joint committee from three of the concerned professional associations pointed up the paucity of controlled data regarding the results of such programs, and cautioned against their implementation.

Observed, involuntary response to sound has been widely employed as a gross measure of hearing status with non-communicative children who did not respond to any available test procedures. This is called behavioral observation audiometry (BOA) and generally involves conditioning some part of the orienting reflex. While this procedure has been successful with children down to the age of one year, a number of limiting factors have restricted its usefulness. A major drawback has been the fact that such reflex behavior may be altered or inhibited by disorders other than hearing loss.

During this "objective" period, considerable interest was also centered upon the development of instrumentation and procedures for detecting and quantifying of electro-physiological changes in response to sound. While a number of such functions have been investigated, only two have received widespread clinical application. The first of these involves the conditioning of sweat gland activity and is referred to as electro-dermal audiometry. Although this technique has been somewhat useful with adult populations, several studies have shown it to be no more effective with children than play audiometry.

Involuntary or evoked changes in brain wave activity in response to sound has evolved, through the development and incorporation of appropriate instrumentation to the status of a useful and effective adjunctive technique for testing children. However, at the present time, it suffers from several limitations, not the least of which are inter- and intra-subject variability in the parameters of the response, difficulties in objectively evaluating responses and the high cost of required instrumentation.

The third and most contemporary period of test development has been classified under the heading "conditioning procedures". This period has involved a strict adherence to the principles of operant conditioning and the application of behavior modification techniques to hearing assessment. This involves the appropriate use of reinforcers to develop and maintain relevant responses and the systematic sequencing of environmental events necessary to bring the desired response under the control of a tone stimuli. The validity of this approach has been established by numerous studies which have successfully applied these techniques to assess the hearing of profoundly retarded children. The potential application of this approach to other difficult-to-test populations is apparent.

CHAPTER III

EXPERIMENTAL TEST INSTRUMENTATION AND METHODS

Described in this chapter is the development of a two-stage hearing test procedure for use with young and non-communicative children. The instrumentation and methods employed were devised and adapted from the techniques (cited in Chapter II) which were found to be successful in obtaining threshold measurements with the mentally retarded. The resultant procedure was structured in strict accord with the principles of operant conditioning and labelled Conditioned Response Audiometry (CRA).

CRA was constructed to meet the need, pointed up by the National Conference on Education of the Deaf,[1] to extend diagnostic efforts for children in the 0-3 year age group beyond those which presently require sophisticated personnel and instrumentation and which are currently found only in the research or medically oriented clinic. It was designed to provide a greater degree of definitive and quantified information than can presently be obtained from such a

---

[1]A. M. Mulholland and G. W. Fellendorf, Final Report of the National Research Conference on Day Programs for Hearing Impaired Children (Washington, D. C.: Alexander Graham Bell Association for the Deaf, 1968), p. 23

population employing any available informal testing technique.

Prototype instrumentation and sequential training and test procedures were developed to be used experimentally in an attempt to evaluate, in part, the efficacy of such a proposed approach.

Instrumentation

A number of electronic instruments are available for use in conventional pure tone threshold testing. Excluding the automatic devices from consideration, the various types differ from one another primarilly in size and function. Smaller models usually provide only for air and bone conduction testing and are intended for portable use. The larger instruments, generally catergorized as clinical models, are equipped for specialized testing techniques and are most often wired into a sound attenuated test room.

Since the primary purpose of this study was to develop and evaluate a test procedure which would have widespread clinical applicability, the selection of equipment was based upon the criteria of maximum availability and maximum versatility.

Audiometer

The portable type audiometer was selected as the basic instrument for CRA on the basis that it best met the criteria of availability and versatility. This decision was based upon the fact that the relatively low cost of portable audiometers make them the most widely available pieces of testing equipment. In addition, the fact that the portable audiometer can be transported to and utilized in any available space provides an essential element of versatility to the CRA approach. Since the initial phase of the test procedure was concerned only with establishing stimulus control and did not involve hearing testing, it did not need to be carried out in a sound controlled environment.

Since the available basic pure tone testing functions do not differ from the portable to the clinical type instrument, the CRA procedures could be carried out in part, or in total, using a clinical model.

Conditioning Adaptor

As previously stated, the first stage of the CRA procedure involved conditioning sessions directed toward establishing appropriate stimulus control. The primary or appropriate stimulus relative to the CRA program was an auditory signal. However, stimulus-discrimination subsumes perception at some CNS level; and since the auditory sensitivity of the population to be tested by these techniques is an unknown quantity, a multisensory conditioning procedure was employed.

That is, because the hearing acuity of the population to be tested was unknown, the initial conditioning procedures were designed to include the simultaneous presentation of visual, tactile and auditory stimuli.

Few, if any, portable audiometers are designed or equipped to provide a visual stimulus to the person under test or to permit the simultaneous presentation of tactile and auditory stimuli. While some clinical models do provide the necessary circuits, it was necessary to develop an accessory device to accomplish this function with the basic instrumentation of this study, the portable audiometer.

A schematic diagram of a prototype adaptor that was developed and constructed for use in this study is shown in Figure I.

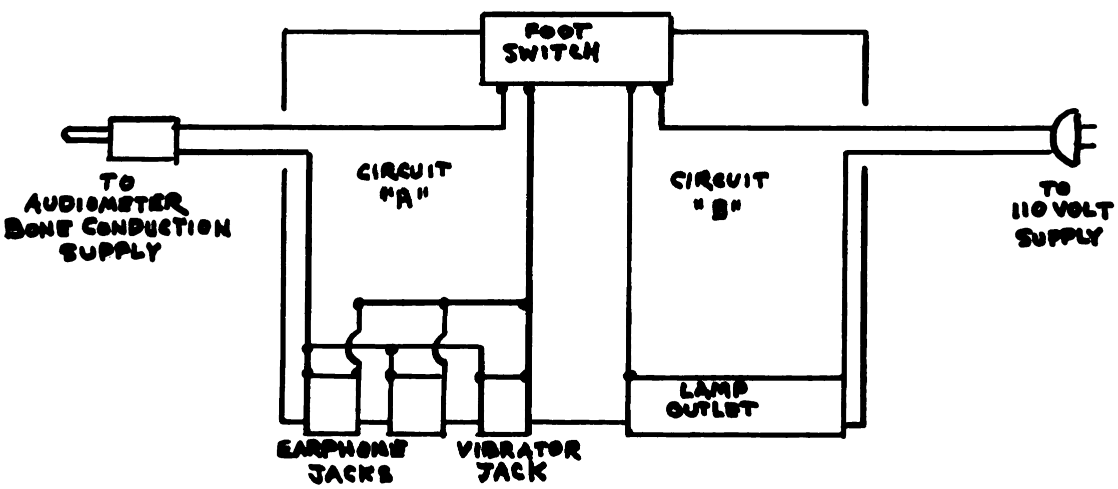

Fig. 1---Shaping adaptor for portable audiometers (SHAPA). Permits simultaneous presentation of visual, tactile and auditory stimuli.

The SHAPA was designed to present a strong vibratory sensation through the bone conduction oscillator of the audiometer, an appropriate pure tone through earphones at a sound pressure level in excess of 100 dB, and to simultaneously present a visual stimulus.

An examination of three portable audiometers, from different manufacturers (Beltone Electronics Corporation, Maico Electronics Incorporated, and Zenith Radio Corporation) revealed that maximum tactile sensation from a hand-held oscillator was induced by maximum output at 500 cps. In addition, it was found that with appropriate accessory circuits sufficient power could be generated to simultaneously produce a 500 cps tone, of approximately 110 dB SPL, through a set of earphones. This was considered to be an acceptable source of power for the tactile stimulus and to provide a most expedient auditory stimulus since low frequency residual hearing is a common characteristic of the profoundly deaf.

As can be noted in Figure I, the SHAPA has two circuits, each with a separate power supply, that are activated by a single switch. The power supply line for circuit "A", which extends outside the adaptor, is terminated by a standard quarter inch 2 conductor phone plug. This was designed to connect to the bone conduction output jack of any audiometer.

Three recessed standard phone jacks are wired in parallel into circuit "A" and provide a connection and power source for the earphones and bone conduction oscillator.

The power supply line of circuit "B" is terminated by a standard 110 volt electrical plug. This circuit contains a standard recessed 110 volt receptacle that provides a connection and power supply for any floor or table lamp.

As previously mentioned both circuits are activated by a single switch. This is a double pole, single throw, foot switch with a 10 ampere rating.

All wiring was done using 18 gauge lamp cord. All of the wiring, jacks, and receptacles were mounted in a bakelite case four inches long, two and a quarter inches wide and two and a quarter inches deep. The total cost of the materials was slightly less than ten dollars and approximately three hours were required to construct the adaptor.

The relative simplicity of the circuit and the fact that components are available which do not require soldering would permit the adaptor to be built by persons with no more than elementary electrical knowledge and skills.

METHODS

The CRA approach was comprised of two separate and distinct stages. Each contains a series of sequential steps structured in strict accord with the principles of operant conditioning. The initial stage was directed toward discrimination and response conditioning whereas the second was concerned with hearing threshold testing.

This divarified structure was utilized in an attempt to maximize the efficiency and applicability of the overall CRA approach. As a result of this division, the initial conditioning procedures, which in some cases may require a considerable expenditure of time, would not be restricted to clinical settings or instrumentation. In addition, they could be carried out by personnel with relatively little knowledge or training in the administration of hearing test procedures. Finally, successful conditioning expedites the testing stage and, since a multisensory approach is employed, failure to achieve conditioning through one or more sensory avenues could provide significant diagnostic information.

Testing procedures in the second state of CRA were structured to promote and facilitate hearing threshold testing with minimal personnel and equipment requirements. Controls were provided to ensure accurate and stable measurements.

Conditioning Procedures

The following procedures were designed to be administered employing a portable audiometer and the shaping adaptor (SHAPA) described above. The equipment was assembled and prepared for use through the following steps.

1. Provide power to the audiometer
2. Remove earphone and bone conduction oscillator
3. Set controls on audiometer
   a. power switch to on position
   b. tone interrupter switch to continuous presentation
   c. output selector to bone conduction mode
   d. frequency selector to 500 cps
   e. hearing level control to maximum output for bc
   f. masking switch and attenuator to off position
4. Connect SHAPA phone plug to audiometer bc output jack
5. Connect earphones and bc oscillator to appropriate jacks on SHAPA
6. Connect SHAPA electrical plug into wall receptacle
7. Connect floor or table lamp in electrical receptacle on SHAPA
8. Position foot switch beyond view of subject

The area to be utilized was cleared of all distracting accouterments and furnished with a small table and appropriate seating. Although an auditory signal was to be employed in the conditioning procedures it was not necessary to measure or control ambient noise levels. The relatively high level of the signal ruled out the possibility of masking by other than extreme levels of noise which would be obvious and uncomfortable to the tester.

A number of tangible items were selected for use as reinforcers in this study. These included a variety of trinkets and solid edibles which could easily be hand-dispensed. However, prior to the initiation of conditioning procedures with any subject, an attempt was made to determine which items were most likely to be effective reinforcers.

This was accomplished through the "closed fist" technique suggested by Lloyd, Spradlin and Reid.[2] Whenever it was apparent from the subject's response that none of the tangible items appeared to have sufficient reinforcement value, liquids such as Kool-aid, orange juice and soda pop, were tried. These were administered by being squirted into the subject's mouth from a small plastic syringe. In some cases semi-solid foods, such as prepared baby foods, were found to be most effective and were dispensed by "loading" a small portion on a number of plastic spoons for rapid delivery. If none of the available items proved to be effective, a deprivation system was employed in which the subject was scheduled for early morning sessions and parents were instructed to withold all food and beverages prior to the clinic visit.

A variety of response modes were available for this study. These consisted of tasks ranging in complexity from those requiring only gross motor movement, such as pushing a block into a box, to those requiring coordinated eye-hand movements. The latter involved such things as placing pegs in a hole, placing graduated plastic rings on a pole, or pushing a button. All subjects were assigned response tasks

---

[2] L. Lloyd, J. Spradlin and M. Reid, "An Operant Audiometric Procedure for Difficult-to-Test Patients." _Journal of Speech and Hearing Disorders_, 33 (1968), pp. 236-245

which appeared, by demonstration, to be within their level of sensory-motor functioning. In addition, an attempt was made to determine which pieces of the response equipment appeared to be most attractive to the subject. This was accomplished by giving each child access to all of the paraphernalia and observing any apparent preference. Such an approach had considerable relevance to the total conditioning procedure through the incorporation of a principle of behavior called the Premack Principle.[3] This states that if one activity occurs more frequently than another, it will be an effective reinforcer for that other activity. For example, children are more apt to inspect or manipulate toys than to sit quietly. However, frequency of sitting behavior may be increased by making access to the toys contingent upon it. The literature reviewed in the previous chapter indicates that the discrimination of pure tones by young and noncommunicative children is a relatively low frequency behavior. Thus it appeared to be most important to make some effort to find and employ high frequency behaviors as response modes.

When the preceding steps had been completed earphones were positioned over the subject's ears, the oscillator was secured to a knuckle on a subject's hand that would not be

---

[3]D. Premack, "Prediction of the Comparative Reinforcement Values of Running and Drinking." Science, 139 (1963), pp. 1062-1063

used in responding, and a light was positioned to illuminate the subject's face. Response paraphernalia were arranged so as to be easily accessible to both the subject and the tester. Since tangible reinforcers were used, it was considered important to place them well beyond the reach of the subject so they were accessible only to the tester. To prevent the sight of the reinforcers from distracting the subject or acting as reinforcers for inappropriate behavior they were also kept from view.

When all preliminary procedures had been completed the training was carried out through the following steps:

Step 1. Conditioning was initiated by depressing the footswitch of the SHAPA to simultaneously present the cueing stimuli. The subject's attention was then directed to each of the stimuli in turn and with some indication of approval from the tester he was lead through the appropriate response. Reinforcement was administered, on a continuous schedule, to coincide as closely as possible to the completion of the response. The footswitch was then released terminating the presentation of the stimuli. This step was repeated five times without change unless there was a positive indication that the subject could carry out the response unaided. If this occurred, training moved immediately to the third step.

Quantified records of stimulus presentations and responses were kept for each subject session that occurred throughout the remaining steps in the training stage.

Step 2. Commencing with the sixth presentation, a five second delay was interjected between the initiating of the stimuli and the response assistance to be provided by the tester. During this period the tester directed the subject's attention to the presence of the three stimuli and in addition to the response mode. If the subject did not make any attempt to initiate the response during the delay period he was lead through by the tester and all reinforcement was withheld. However, if the subject made any partial-response movement, approval was indicated by the tester, assistance provided to complete the response and reinforcement administered. Extensions of the delay period and gestural prompting were then employed to progressively condition a complete response.

Any response-movements that occurred after the stimuli-presentations were terminated were rejected by the tester. This refusal was conveyed to the subject by vigorous negative gestures and reinstating any response equipment which may have been dispositioned. The procedures in this step were continued until the subject initiated and completed the response unaided.

If it became apparent that stimulus control could not be established, training reverted to step one and the reinforcement changed or a deprivation system employed as outlined above.

Step 3. This step was directed toward establishing a measure of successful stimulus control. Stimuli presentations were limited to a maximum of ten second duration in the absence of a response. Prompting was limited to attempts to establish a set-to-respond in the subject. Primarily, this consisted of directing the subject's attention to the response equipment prior to stimuli presentations. The criterion for successful establishment of stimulus control was set at ten consecutive responses. A variable ratio schedule of reinforcement (3:1) was employed in this and succeeding steps.

Step 4. In this step the only variance from the preceding one involved the elimination of the visual stimulus. This was accomplished by unplugging the light from the SHAPA and removing it from the subject's view. Training continued until the criterion measure was reached. If the response broke down step three was repeated.

Step 5. Variance in this step involved the removal of the tactile stimuli. At this point the SHAPA was eliminated, the oscillator removed from the subject's hand and the earphones were connected directly to the audiometer. Training was continued until the criterion measure was reached employing the 500 cps tone. The total procedure was then repeated, first with a 1000 cps and then a 2000 cps tone. When the response was maintained for the additional frequencies the training was terminated and phase I of the CRA procedure was considered completed.

Testing Procedures

The testing phase of the CRA procedure was structured to utilize the conditioned response developed in the training phase to determine hearing threshold levels. Consequently, the consideration of ambient noise levels in the testing environment was considered a primary factor. While this could be most efficiently accomplished employing electronic sound measuring equipment, such instrumentation is not widely available. Consequently a more pragmatic approach was employed which was described by Glorig as the "Adequate Listener" procedure.[4] This involved measuring, within the area to be used, the thresholds of several persons whose thresholds had previously been measured in a sound controlled environment. Significant deviations between tests were recorded and appropriate adjustments made of the intensity levels employed.

Prior to testing, the earphones of the audiometer were positioned on the subject's head and the particular response equipment utilized in Phase 1 was placed before him. The audiometer was situated behind the subject with the controls set for testing by air conduction. A remote control extension cord (available from most variety stores) was plugged into

---

[4]A. Glorig, Audiometry: Principles and Practices. (Baltimore: The Williams and Wilkins Company, 1965)

an electrical outlet close to the audiometer with a small lamp (commercially available night light) connected to the socket. The cord was extended so that the switch could be positioned in front of the subject. This equipment was used as a signal device in the procedures described below.

The validity of subject responses were maintained through the use of specific procedures carried out by two testers.

Tester #1 was equipped with ear muffs and plugs and positioned in front of the subject. His function was to direct the subject's attention to the testing task, to signal for the presentation of test tones by depressing the switch on the extension cord and illuminating the light beside the audiometer, and to administer reinforcement on a variable ratio schedule. The sound attenuating devices were necessary to prevent this tester from hearing test tones presented to the subject at high intensity levels.

Tester #2 was positioned at the controls of the audiometer and in view of the signal lamp. His function was to select the frequencies and intensities to be tested, to present tones and interject a random pattern of no-stimulus presentations whenever the signal lamp was illuminated by Tester #1. This latter function was incorporated to ensure that response contingencies were established and maintained solely to the auditory stimuli. In other words, any response occurring during a no-stimulus period indicated that the subject was responding to something other than the test tone and invalidated all results obtained to that point for the frequency under test.

When such false-positive responses occurred, it was assumed they resulted from visual cues provided by tester #1 or because complete stimulus control had not been established. Testing was immediately interrupted by tester #2 and remedial procedures were instituted. These ranged from observations and restructuring of tester-subject interaction to the termination of testing and repetition of conditioning sessions.

SUMMARY

A hearing test procedure, called Conditioned Response Audiometry (CRA), was designed and developed in an attempt to meet a need for testing methods applicable to very young and noncommunicative children and which would have the potential for widespread availability.

A two-phase structure was utilized to maximize the efficiency and applicability of the overall approach. The first phase was limited to the conditioning procedures directed toward establishing stimulus control while the second phase was concerned with the measurement of hearing thresholds.

This arbitrary division of training and testing was created in an attempt to alleviate factors which may have limited the use of conditioning procedures in obtaining hearing threshold measurements with difficult-to-test patients. Since an operant approach to audiometry requires establishment of stimulus control prior to the initiation of threshold testing, two separate procedures are, in fact, involved. More importantly, major differences in space and personnel requirements and availability exist between these procedures. For example, the existence of high ambient noise levels in space employed for threshold testing may adversely effect the results obtained. However, the initial conditioning procedures, which often involve protracted sessions, can be carried out employing auditory signals of sufficient intensity to override existing noise levels. In addition, threshold audiometry requires highly trained and experienced personnel whereas the conditioning procedures can be carried out with little, if any, training in hearing testing. Since the availability of noise-attenuated space and professional audiologists is extremely limited or heavily scheduled in most field settings, the separation of these procedures should enhance and extend their application.

The conditioning procedures carried out in the initial phase of CRA involved a multisensory approach using visual, tactile and auditory stimuli. Response modes were selected on the basis of their appropriateness to the child's level of motor functioning and a demonstrated "attractiveness." The technique consisted of the simultaneous presentation of the three signals, modeling the response, and the administration of a tangible reinforcer. The modeling was faded out as stimulus control was established. At that point the visual stimulus was removed. If the response was maintained, the tactile stimulus was eliminated. When conditioning was maintained to the auditory stimulus, the second or testing phase of CRA was initiated. This multisensory approach was employed because the hearing status of the population to be tested was unknown and differences in the success of establishing a conditioned response between the various stimuli could provide some preliminary diagnostic information.

To facilitate the widespread use of the CRA approach, the techniques used in both phases were structured to be administered employing a portable audiometer. However, since the portable type of instrument is not generally equipped to permit the simultaneous presentation of visual, tactile, and auditory stimuli, an adaptor was developed to accomplish this function. This accessory device was called a shaping adaptor for portable audiometers (SHAPA) and was designed to permit construction by persons with limited knowledge of electrical circuits and to be made up of parts and materials which could be acquired at relatively low cost.

The second or testing phase of the CRA approach was structured to utilize the conditioned response developed during the initial or training phase to determine hearing threshold levels of the same subjects. Since the immaturity of the population to be tested often required someone to be in view of the subject to focus and maintain interest in the test, the possibility existed of inadvertently providing a visual stimulus which could cue responses. Consequently it was necessary to develop specific procedures to maintain the validity of subject responses. This involved the use of two testers. One positioned in front of the subject who functioned to maintain subject attention, signal for stimulus presentations and administer reinforcement. The second tester, with the audiometer, was positioned behind the subject. His function was to present a schedule of test tones randomly interspersed with no-stimulus periods. Any response occurring during a no-stimulus period indicated the subject was responding to something other than the test tone and signaled the termination of testing. Remedial procedures were then instituted to establish appropriate stimulus control.

CHAPTER IV

RESEARCH PROCEDURES

In this chapter specifics are given regarding the selection of subjects, their assignment to two hearing test groups, the test procedures administered to them, the instrumentation employed, and the overall research design.

In brief, subjects were selected from the total population of a community preschool deaf education program. These subjects were assigned to the groups on the basis of their ability (Group I) or failure (Group II) to respond consistently and appropriately to conventional pure tone hearing test procedures. The groups were matched on age, educational experience, and category of loss.

Three test measures were employed with Group I. These included conventional pure tone play audiometry (CPA), evoked response audiometry (ERA), and conditioned response audiometry (CRA). The results of these tests were employed as criterion measures for the results of ERA and CRA measurements obtained from Group II. All tests were administered twice to each subject to assess the reliability of the procedures within the two groups.

Subjects

The purpose of this study was to evaluate the efficacy of an experimental hearing test procedure designed and developed for use with young and noncommunicative children. Consequently, interest was focused upon the kinds of subjects usually identified by the audiologist as difficult-to-test. This categorization is sometimes applied to adults functioning at low levels of intellectual performance. However, it is as often used to refer to children, below school age, who have little, if any, functional speech or measurable language.

It is this latter group that is most often seen in the usual speech and hearing field clinic and found to be untestable by conventional means. In many cases further testing is deferred to await maturational development. All too frequently such children are diagnosed as deaf, autistic, or retarded on the basis of gross and sometimes intuitive kinds of evaluations. Unfortunately, habilitative measures are many times initiated on the basis of such perfunctory labeling. When this occurs, further evaluative efforts may be indefinitely, if not permanently, forestalled. Such delays do not result from negligence or oversight but appear instead to involve some circular kinds of circumstances which are analogous to the behavioral phenomenon referred to by Rosenthal[1] as a "self-fulfilling prophecy." In the case of the untestable child this involves training initiated to fit a label; behavior conforming to the training model presented; the resultant behavior confirms the label and obviates the urgency or need for further testing.

Children are to be found in many special education programs who have been enrolled for periods of one year or longer on the basis of limited and often unquantified information obtained in a single diagnostic session.

Selection of Subjects

Subjects were selected from the population of the Tampa Oral School for the Deaf, Tampa, Florida. This is a preschool, deaf education, day program that had a total enrollment of 60 children at the time of this study.

The files of the school were surveyed and data on each pupil were recorded. These data included information about chronological age, length of time in school, categorization of hearing loss, pure tone air conduction threshold, and reliability estimates of obtained threshold measurements.

On the basis of these data two lists of names were compiled. List one contained the names of 24 children:

---

[1]R. Rosenthal, "Self-Fulfilling Prophecy," Psychology Today, 2 (1968), pp. 44-51

(a) whose chronological age was three years or below, (b) who had been in school for two years or less, (c) whose hearing status had been categorized as a severe or profound bilateral impairment and, (d) whose recorded pure tone threshold levels for the speech frequencies, on repeated measures, did not vary more than ±5 dB.

List II contained the names of 19 children: (a) whose chronological age was three years or below, (b) who had been in school for two years or less, (c) whose hearing status had been categorized as a severe or profound bilateral impairment and, (d) whose recorded pure tone threshold levels for the speech frequencies were unobtainable or consisted of response levels obtained by gross sound field techniques.

Initial contact was made, in person or by telephone, with the parents of all the children whose names appeared on each of the lists. The overall purposes and procedures of the study were explained and they were asked whether their child, if selected, could participate in the study. Parental agreement was obtained for 23 of the children on List I and for all 19 of the children on List II.

The names of ten children were then selected from each list and assigned to corresponding test groups I and II. The means, medians and ranges were matched between the groups. This was done to establish "equivalency" between the groups so that the results of tests could be attributed to the procedures utilized rather than to any possible differences between groups. Substantially more confidence could then be given to conclusions drawn regarding the effects of the specific treatments under experimental control.

The results of the above matching-between-groups procedure are shown in Table 1. The mean age of all twenty subjects was 2.65 years and the mean educational level was 1.12 years. There were seven profound and three severe losses in one group and eight profound and two severe losses in the other. As can be noted, there were no major discrepancies between the groups on these measures. However, these measures now having been specified can be consulted if necessary and results reported in Chapter V qualified accordingly.

The primary variable of difference between the test groups was the ability of the subjects within the groups to respond appropriately and consistently to conventional pure tone hearing test procedures.

Three of the subjects initially assigned to Group II were replaced before the study was completed (Table 1 reflects the replacements). Although these original three subjects had received twice as many conditioning sessions as any other subject within the group, it was not possible

Table 1. Means, medians, ranges and descriptive data resulting from group matching procedures (N=20, n=10)

	Group I	Group II	Total
Mean Age (years)	2.69	2.62	2.65
range:	2.4-3.0	2.2-3.0	
median:	2.7	2.6	
Mean Educ. (years)	1.15	1.10	1.12
range:	.5-2.0	.5-2.0	
median:	1.0	1.2	
Type of Loss			
Profound	7	8	15
Severe	3	2	5
Etiology of Loss			
Rubella	5	6	11
Unknown	5	4	9
Sex			
Male	5	5	10
Female	5	5	10
Race			
Caucasian	5	5	10
Negro	5	5	10

to establish a conditioned response (as defined by this study) to any of the stimuli employed. The rationale for replacement was based upon the fact that the primary purpose of the study was to evaluate the test phase of the CRA approach which is dependent upon successful conditioning. However, since it was believed that failure-to-condition could be of diagnostic significance, other evaluative procedures were scheduled for these three subjects after they were removed from the group and the results are reported in Chapter V.

Treatment Conditions

Three factors were of interest in this study: (a) an evaluation of test-retest reliability of threshold levels obtained by CRA procedures, (b) a comparison of hearing threshold levels obtained by the CRA approach with thresholds obtained by conventional play audiometry procedures, and (c) a determination of the validity of CRA threshold levels obtained from young children who have been demonstrated untestable by conventional hearing test procedures. However, since the second and last of these factors were incompatible by definition the use of two test groups and the inclusion of an objective test measure (ERA) were thought to provide an indirect approach to the assessment of these variables.

The three test measures utilized in this study included conventional play audiometry (CPA), the experimental CRA approach, and evoked response audiometry (ERA). Group I was tested by all three techniques, whereas Group II received only CRA and ERA.

Each test procedure was administered twice to each subject and was presented unilaterally with the test ear randomly determined. Threshold levels were obtained for the frequencies 500, 1000, and 2000 cps employing the descending method of stimulus presentation described in Chapter I, page 10. All sessions were carried out in the clinical facilities of the Speech Pathology and Audiology Department at the University of South Florida, Tampa, Florida. ERA testing was conducted in a commercially built sound-treated test room, whereas CPA and CRA conditioning procedures and tests were administered in untreated clinic rooms. Ambient noise levels in the clinic rooms were checked and handled employing the "Adequate Listener" technique previously detailed as part of the CRA procedures.

All scheduled sessions involved a one hour period. Conditioning sessions, for both groups, were carried out daily with individual assignments determined on the basis of subject availability.

Two hearing tests were administered during each testing session. These sessions were spaced from one day to one week apart with the interval determined by subject availability and scheduling needs. A minimum of three testing sessions was required for each subject in Group I. At the first scheduled session for each subject CPA and CRA were randomly assigned and administered to confound the effects of ordering. For the remaining two sessions CPA and CRA were randomly assigned for initial presentation with each to be followed by an ERA test. ERA was not considered to contribute to any ordering effect since it does not involve any active participation or learning on the part of the subject. In addition, it was considered most appropriate as a final procedure since it required a protracted period of relative inactivity. A minimum of two testing sessions was required for each subject in Group II. At each session CRA was presented first and followed by ERA.

The output of each earphone employed in this study was calibrated, according to ISO procedures,[2] prior to the initiation of any testing and was checked periodically.

Otoscopic examinations were administered by audiologists to each subject before each testing session to guard against any variations in hearing acuity due to temporary conductive

---

[2] Standard Reference Zero for Calibration of Pure Tone Audiometers, ISO/R 389-1964

impairments. Whenever such problems were detected the session was terminated, the subject referred out for an otologic examination, and the session was then rescheduled accordingly. During the course of this study two children were found to have impacted cerumen in the external auditory canals.

The specific techniques and procedures employed in each of the three test conditions are described in the following sections.

Conventional Play Audiometry

Conventional pure tone air conduction hearing test procedures were employed using intermittent social reinforcement (smiles and nods) and modifications of response modes which appeared to be appropriate to the interests and motor abilities of each subject. These response techniques included the picture/tone, paired-association method and a modification of the lantern slide approach described in Chapter II, pages 15 and 16.

Subjects were seated in the test room with earphones in place and response equipment within reach. The tester was seated beside the subject and the test equipment was placed on the opposite side beyond the subject's view. The total test procedure was modeled for the subject using a 1000 cps test signal. When it was apparent the subject understood the procedure, by completing a response unaided, testing was initiated at the same frequency and proceeded to 500 and 2000 cps. Test tones were attenuated as subjects responded appropriately to three consecutive tone presentations within a total of six.

Two audiologists were employed on a rotating schedule in the administration of all CPA tests so that no subject was tested twice by the same tester. Threshold levels for each subject were plotted on standard audiograms and placed in a file immediately following each test and were not made available to any tester until all tests had been completed on all subjects. (CPA test results for each subject in Group I are reported in Appendix A.)

Conditioned Response Audiometry

The sequential steps described in Chapter III for both phases of the CRA procedures were employed with all subjects in Groups I and II. However, in the administration of the conditioning phase to Group I it was possible to eliminate the second step which involved shaping small increments of the total response. In addition, since this phase did not involve testing and, with Group II, often involved prolonged periods of time, several subjects were scheduled for shaping at the same time. The conditioning was carried out by students enrolled in the Speech Pathology and Audiology

Program. None of these students had been previously exposed to any formal coursework in the area of behavior modification. Prior to their assignment to the administration of conditioning sessions each student was provided with a copy of the CRA-phase I procedures, outlined in the previous chapter.

The number of sessions required to achieve conditioning varied among the subjects. The total number of sessions involved for each subject, in each group, are reported in Appendix B.

Phase 2 was initiated, for each subject, employing the 500 cps tone used in the conditioning phase and proceeded to 1000 and 2000 cps. Test tones were attenuated as the subjects responded correctly to three consecutive tone presentations in a total of six. No-stimulus periods were randomly assigned by tester #2 prior to initiating a test. Figure 1 illustrates the form employed for this purpose. The N/S periods were only considered in respect to inappropriate responses and were not treated as a break in the continuity of tone presentations. For example, on the program depicted in Figure 1, under the 500 cps column at a stimulus level of 120 dB, if a subject responded appropriately to presentation numbers 1 and 2, the tone would not be attenuated unless he also responded correctly to numbers 3 and 4.

Subject #________ Testor #________ Date ________

FREQUENCY		500						1000						2000					
PRESENTATION		1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
120	Program	N	S	S	S	N	S	N	S	S	N	S	S	S	S	N	S	S	N
	Responses																		
115	Program	S	N	S	S	N	S	S	S	N	S	S	N	N	S	S	N	S	S
	Responses																		
110	Program	S	N	S	S	S	N	N	S	S	N	S	S	N	S	S	S	N	S
	Responses																		
105	Program	N	S	S	S	S	N	N	S	S	S	S	N	S	N	S	S	N	S
	Responses																		
100	Program	N	S	S	N	S	S	N	S	S	S	N	S	N	S	S	S	S	N
	Responses																		
95	Program	S	S	N	S	S	N	S	N	S	S	N	S	N	S	S	N	S	S
	Responses																		
90	Program	S	N	S	S	S	N	N	S	S	N	S	S	S	N	S	S	N	S
	Responses																		
85	Program	S	S	N	S	S	N	N	S	S	N	S	S	N	S	S	S	N	S
	Responses																		
80	Program	N	S	S	N	S	S	N	S	S	S	S	N	N	S	S	N	S	S
	Responses																		
75	Program	S	S	N	S	S	N	S	N	S	S	N	S	S	N	S	S	S	N
	Responses																		
70	Program	S	N	S	S	S	N	N	S	S	S	S	N	N	S	S	N	S	S
	Responses																		
65	Program	N	S	S	N	S	S	N	S	S	S	S	N	N	S	S	S	N	S
	Responses																		

N=No stimulus S=Stimulus

Figure 1---CRA test form with a stimulus-no-stimulus program inserted

Assignment to the role of Tester #2 was alternated between two audiologists so that no subject was tested twice by the same tester. CRA test forms were removed following each test, as previously described under CPA, and were not returned until all tests were completed. (CRA test results for each subject in Groups I and II are reported in Appendix A.)

Evoked Response Audiometry

All ERA tests were administered with subjects in a waking state and seated in a lounge chair. An attendant was present in the test room to maintain interest and to prevent excessive physical activity on the part of the subject. Pictures and small toys were provided and the subjects were given access to them between test runs.

Three scalp electrodes were attached to the subject's head. A recording electrode was positioned at the vertex, a referent electrode on the mastoid area, and a ground electrode fastened to the earlobe. Phones were positioned over the subject's ears.

Testing was initiated at a frequency of 500 cps and proceeded to 1000 and 2000 cps. At each frequency, the initial test was presented at a hearing level of 120 dB ISO and attenuated in 5 dB decrements. Two test runs, separated by a silent control run, were made at each hearing level. Test runs consisted of the presentation of 60 pure tone pulses with a rise/fall time of 20 milliseconds duration, and a pulse interval of 1 second. Total time for each test run was 60 seconds.

A signal averaging computer, with 100 averaging points, was employed to extract responses from ongoing EEG activity. Strip chart readouts of the averaged data were provided at the midpoint (30 presentations) and end (60 presentations) of each run. Silent control periods consisted of 60 second averaging periods during which there were no intentional auditory stimuli presented.

Online evaluations of responses were carried out at the end of each series of test-control-test runs. Response criteria included: (1) a temporal match of any negative or positive waveform occurring within the first 250 milliseconds on the averaged traces of the first and second test runs (amplitude variations were accepted since Teas[3] and other investigators have demonstrated they occur in some degree as a function of adaptation), and (2) an absence in the averaged trace of the silent control run of any negative or positive waveform, or any segment, which would coincide

---

[3]D. C. Teas, "Analysis of Evoked and Ongoing Electrical Activity at the Scalp of Human Subjects," _Journal of Speech and Hearing Research_, 8 (1965), pp. 371-387

temporally with the matched waveforms obtained in the test runs. An example of this wave-matching technique is presented in Figure 2.

Threshold levels were defined as the lowest hearing level at which a response could be identified by the trace-matching technique described above.

Two audiologists were employed in the administration of all complete ERA tests so that no subject was tested twice by the same tester. No attempt was made to match waveforms between complete tests and testers were not given access to completed records until all subjects had been tested. (ERA threshold results for each subject in Groups I and II are reported in Appendix A.)

Instrumentation

The instrumentation employed in this study is described below under appropriate headings. As previously pointed out, although a clinical audiometer and a commercially built sound attenuated test room were available to this study, they were employed only for ERA testing. This approach appeared most expedient since CRA procedures were designed and developed for use with portable type instruments in field settings where commercial test rooms would not be widely available.

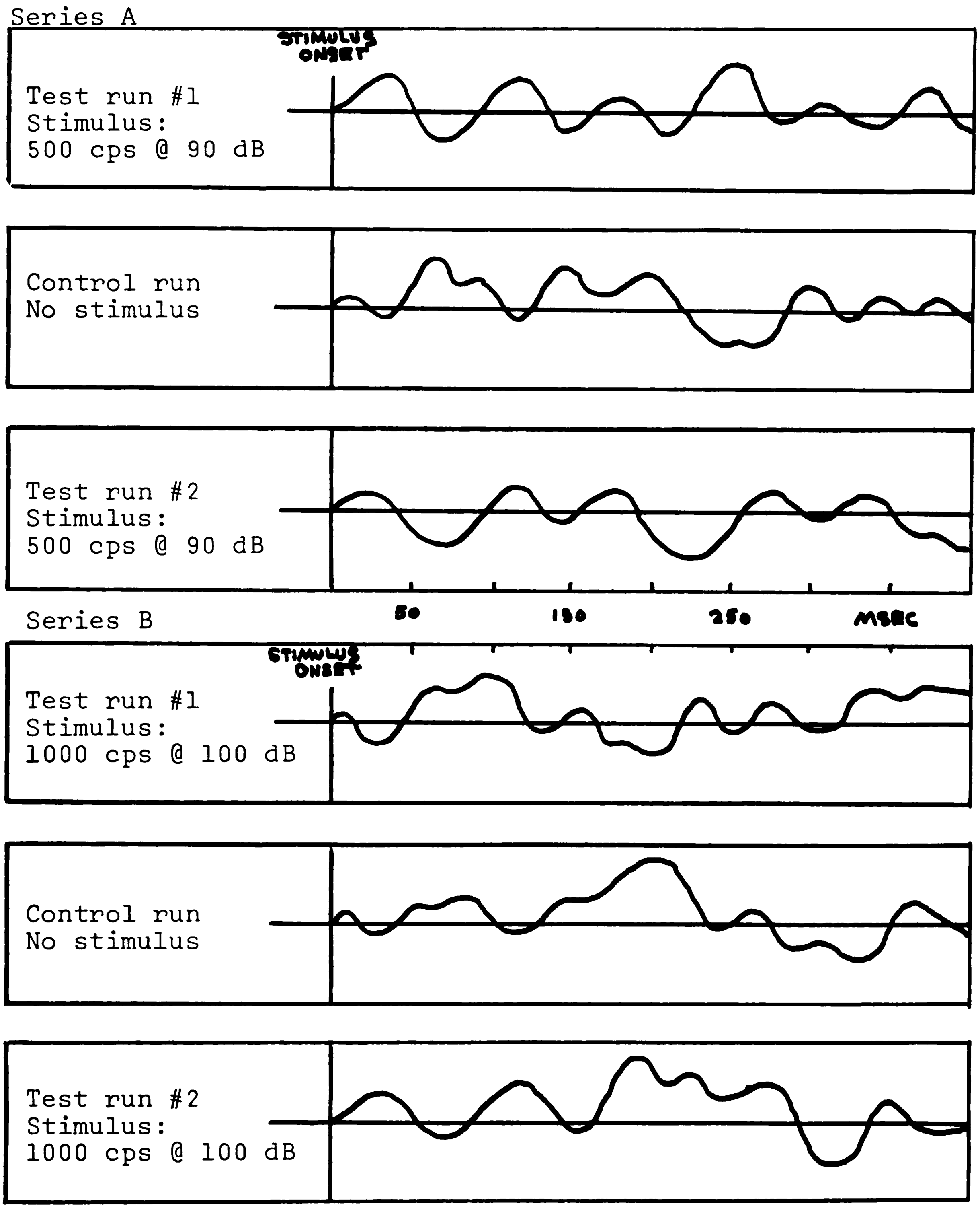

Figure 2---Two series of ERA test-control-test runs. "A" depicts an acceptable response on the basis of a match between wave forms on first and second run and no match with control run. "B" depicts no acceptable response on the basis of an absence of any match between stimulus runs and some degree of match between stimulus runs and control.

Conventional Play Audiometry

All CPA tests were administered using a Maico, MA-2 portable audiometer with a matched set of TDH-39 earphones.

The modified lantern slide response equipment consisted of a Kodak, Model 140, Carousel Projector on which the lamp circuit had been altered to provide subject response and tester control switches. These switches were wired in series so the lamp could be illuminated only when both were simultaneously activated. Slides were advanced by means of a remote control switch.

Conditioned Response Audiometry

The Maico portable audiometer, described above, was employed in both the conditioning and testing phases of this procedure. In addition, Phase 1 (conditioning) was carried out using the SHAPA device described in Chapter III. Six of these adaptors were constructed in the electronics shop of the Speech Pathology and Audiology Training Program.

Evoked Response Audiometry

ERA tests were carried out with the subjects seated in an Industrial Acoustics Company, Model 400-A, sound treated test room.

The tests were administered employing a Princeton Applied Research Corporation, Model 140, Evoked Response Audiometer. The system comprises a Tone Generator and

Headset calibrated to ISO, 1964 standards, and EEG Preamplifier, Averaging Computer, Strip Chart Recorder and a Programmer.

Input to the preamplifier was supplied through silver-disk electrodes, seated in an electrolyte gel and attached to the subject's scalp by adhesive collars.

Calibration Equipment

The initial calibration and periodic checks of earphones employed in this study were carried out through the use of an artificial ear (B & K, Model 4152, 6 cc coupler), a microphone (B & K, Model 4132), and a sound level meter (B & K, Precision Sound Level Meter, Model 2203).

Statistical Design

Threshold measurements for the subjects in both groups represented the raw data for entry into A x B x C analyses of variance, factorial design suggested by Winer.[4] These analyses were carried out independently for Groups I and II. Test-retest reliability coefficients and standard error of measurements relative to treatment conditions were also computed. Descriptive statistics were employed where necessary to specify particular aspects of subject performance. These statistical procedures provided an objective basis

---

[4] B. J. Winer, Statistical Principles in Experimental Design (New York: McGraw-Hill Co., 1962), pp. 337-349

for answering the questions posed by this study.

Summary

Twenty subjects, matched on the basis of age, education, and category of loss, were selected for this study from the population of a preschool, day care, program for the deaf.

These subjects were assigned to one of two test groups on the basis of their ability (Group I) or failure (Group II) to respond consistently and appropriately to conventional play audiometric techniques.

Group I received three test measures consisting of Conventional Play Audiometry (CPA), an experimental test technique called Conditioned Response Audiometry (CRA), and Evoked Response Audiometry (ERA) a relatively objective procedure.

Since Group II, by definition, could not be tested by Conventional Play Audiometry they received only CRA and ERA.

All tests were administered twice and threshold levels were obtained, unilaterally, for the frequencies 500, 1000, and 2000 cps.

Primary factors of interest included a comparison of threshold levels obtained by CPA and CRA, test-retest reliability of the CRA procedures, and an evaluation of the validity of the CRA approach with children demonstrated to be untestable by CPA.

CHAPTER V

RESULTS AND DISCUSSION

This chapter presents the major results obtained by the procedures described in Chapter IV. Twenty subjects were systematically assigned to one of two test groups on the basis of their ability to respond consistently and appropriately to conventional pure tone hearing test methods (Group I) or their failure to do so (Group II). Extraneous variables were matched as closely as possible among the groups. The results of these matching procedures are shown in Table 1, Chapter IV, page 101.

Conventional pure tone, conditioned response, and evoked response hearing tests were administered twice to each subject in Group I. These results were analyzed to determine test-retest reliability of the CRA and ERA procedures and to assess the validity of these techniques employing CPA as the criterion measure.

CRA and ERA were administered twice to each subject in Group II as a means of evaluating the reliability and the validity of the CRA procedures with a population untestable by conventional pure tone procedures.

Following the presentation of statistical results and associated conclusions, these results are interpreted and discussed in terms of the original questions posed in this investigation.

Obtained Test-retest Results

Tables 1 and 2 below present the distribution of the threshold levels obtained from all subjects in Groups I and II, for each of the three test frequencies, under each of the test methods.

Table 1. Frequency distribution of test-retest threshold levels obtained from Group I for three test frequencies under three test methods.

dB	CPA			CRA			ERA		
Level	500	1000	2000	500	1000	2000	500	1000	2000
55	2								
60	1			1					
65	12			2			1		
70	2	2		2					
75	112	1		112	12		112		
80	1112	12		122	1		112	2	
85	222	12		1122			112222	112	
90	112	112	12	1112	1122		1	11112222	2
95		11222	1112222	1	1112222	12	122	2	1
100	2	122	112			11122	2	1	111222
105		1	12		22	1122		12	122
110	1	12	122		1112	11222	2	1	
115									
120			112	12		1122		122	1122

Test = 1

Retest = 2

Table 2. Frequency distribution of test-retest threshold levels obtained from Group II for three test frequencies under two test methods.

dB Level	CRA			ERA		
	500	1000	2000	500	1000	2000
50	1					
55	2					
60						
65		1		1		
70				2		
75		2			12	
80						
85	1	12	1	2		
90	2		2	1	12	2
95	11222	12	112	2		1
100	11112222	1112	122	1		
105	12	1222	112	11222	1122	12
110	1	1122	1122	111111222	111222	1122
115						
120		12	11222		111222	111111222222

Test = 1

Retest = 2

As can be noted from an inspection of Table 1, there are some extreme high threshold levels for the low test frequencies under all test methods which results in a skewing of those distributions toward the lower intensity levels.

Table 2 reveals an opposite trend with some extreme low threshold levels for the low and middle frequencies, under both test methods which results in a skewing of those distributions toward the higher intensity levels.

These departures from statistical normality in the threshold-level distributions of both Groups appeared to be a function of the selection of subjects for this study and to

result from range restrictions imposed by the "ceiling effect" of the severe to profound hearing losses exhibited by the subjects.

One of the assumptions underlying the use of parametric statistical procedures is that each population has a normal distribution of dependent scores. However, according to Guilford, nonnormal distributions resulting from the truncating effects inherent in the measurement of certain psychological or physiological variables can be appropriately transformed into values on a new scale in which the distribution is normal.[1] Accordingly, prior to the analysis of the data in this study, all obtained threshold level measurements were converted to logarithmic equivalents.

Test-retest Reliability

A number of studies are to be found in the literature which have been directed toward establishing the reliability of conventional pure tone audiometric procedures. These studies have been carried out under both clinical and laboratory conditions and with various types of populations. Unfortunately there has been little uniformity among such studies in the types of statistical procedures employed to assess and report reliability data. As a consequence, opportunities for evaluating and equating results have been

---

[1]J. P. Guilford. Fundamental Statistics in Psychology and Education. (New York: McGraw-Hill Book Company, 1965) pp. 252-253

limited.

High, Glorig, and Nixon[2] and Jerger[3] pointed up these procedural incongruities and suggested the use of correlation coefficients as measures of "relative consistency"; i.e., the stability of measurement within uniform populations, such as might result from the categorization of hearing loss and the standard error of measurement as an index of "absolute consistency" i.e., the stability of individual scores.

Following this proposed approach, the test-retest reliability data resulting from this study are reported in terms of coefficients of correlation and standard error of measurement for each test group, under the appropriate headings below.

Test Group I

Conventional play audiometry (CPA), Conditioned response audiometry (CRA) and Evoked response audiometry (ERA) reliability coefficients (r) and standard error of measurement (Se) relative to test-retest administration are shown in the following table.

---

[2]W. S. High, A. Glorig, and J. Nixon. "Estimating the Reliability of Auditory Threshold Measurements." Journal of Auditory Research, 1 (1961), pp. 247-262

[3]J. Jerger. "Comments on Estimating the Reliability of Auditory Threshold Measurements." Journal of Auditory Research, 2 (1961), pp. 138-142

Table 3. Test-retest reliability coefficients and standard error of measurement for test methods by frequency for Group I

Test Method	Frequency (cps) 500	1000	2000
CPA			
r	.96	.91	.89
Se	2.23	3.09	3.33
CRA			
r	.96	.94	.89
Se	2.93	2.85	2.67
ERA			
r	.92	.97	.97
Se	3.33	1.94	1.41

As can be noted, the obtained r's for the three test frequencies under each of the three test methods were all of sufficient magnitude to indicate a relatively high degree of test-retest reliability. This assumption was substantiated, in part, by reference to tables presented by Guilford[4] which indicate that for a sample as small as ten an obtained coefficient of .76 or greater is significantly different from zero at the .01 level.

The lowest obtained Se's were for ERA at 1000 and 2000 cps. However, doubling Se values for all test methods at all test frequencies to obtain 95% confidence zones for expected scores indicated these variances would be close to the clinically accepted standard of ± 5 dB.

[4]Guilford, op. cit., pp. 538-539

Test Group II

Reliability coefficients and standard error of measurement for the two test methods administered to this group are presented in the following table.

Table 4. Test-retest reliability coefficients and standard error of measurement for test methods by frequency for Group II

Test Method	Frequency (cps) 500	1000	2000
CRA	.98	.96	.95
r	2.25	2.89	2.37
Se			
ERA			
r	.98	.96	.99
Se	1.95	2.72	1.83

With a single exception (ERA at 1000 cps) all of the r's obtained for this group were slightly larger than those reported for the corresponding frequency for Group I and were significantly different from zero beyond the 0.01-level of confidence.

The Se's of each of the test frequencies under both test methods approximated those obtained for the same tests administered to Group I. Relative to the confidence zones described above, the dispersion of the expected scores for these obtained Se's would be close to the ±5 dB clinical standard.

Predictive Validity

Product-moment correlations were computed with the threshold scores for the three test frequencies between the various test methods employed with each of the test groups. The obtained coefficients are presented under appropriate headings below.

Test Group I

The results of the inter-method analyses are indicated in the following table:

Table 5. Correlation coefficients between test methods, by frequency, for Group I

Tests	Frequency	r
CPA/CRA	500	.91
	1000	.94
	2000	.92
CPA/ERA	500	.89
	1000	.79
	2000	.76
CRA/ERA	500	.92
	1000	.81
	2000	.61

All of the obtained correlations attained or exceeded significance at the 0.01-level with the exception of the CRA/ERA comparison at the 2000 cps frequency which, with an r of .61, did not reach the 0.05-level.

As can be noted there is a trend toward a progression in coefficient strength from the high to low frequencies.

Test Group II

The following table presents the results of the inter-method analyses carried out for this group.

Table 6. Correlation coefficients between test methods by frequency for Group II.

Tests	Frequency	r
CRA/ERA	500	.98
	1000	.94
	2000	.89

The obtained coefficients in the above table are all significantly different from zero beyond the 0.01-level. A frequency-coefficient magnitude progression is apparent similar to that reported for Group I.

This tendency toward lower test frequencies exhibiting higher inter-method coefficients, a trend also apparent in the obtained test-retest reliability data (Tables 3 and 4), is at variance with results obtained from comparable studies with hard of hearing adults and mentally retarded populations.

An early study by Witting and Hughson reported the smallest variance on repeated audiograms, with hard of hearing adults, at 1000 cps.[5] These results were confirmed in a study by Gardner[6] and have resulted in the widespread clinical practice of employing the 1000 cps test tone as an intra-test reliability check.

However, a number of investigators, such as Fulton[7] and Lloyd, Reid and McManis,[8] have reported a test-retest progression in reliability strength from low to high test frequencies using standard audiometric procedures with groups of mentally retarded children. This trend was minimal, and it was concluded there was insufficient evidence to indicate any type-stimuli was substantially more reliable than any other.

The trend toward increasing coefficient strength with decreasing test frequencies, which is apparent in this study,

---

[5]E. Witting, and W. Hughson. "Inherent Accuracy of a Series of Repeated Clinical Audiograms." Laryngoscope, 50 (1940), pp. 259-269

[6]M. Gardner. "A Pulse-tone Clinical Audiometer." Journal of Acoustical Society of America, 19 (1947), pp. 178-190

[7]R. Fulton. "Standard Pure Tone and Bekesy Audiometric Measures with the Mentally Retarded." American Journal of Mental Deficiency, 71 (1967), pp. 60-73

[8]L. Lloyd, M. Reid, and D. McManis. "Pure Tone Reliability of Institutionalized MR Children.: Report #51, (1967) Parsons State Hospital, Parsons, Kansas.

might be attributed to "practice effects" resulting from the use of a 500 cps tone in the conditioning phase of the CRA procedure prior to the initiation of testing. This explanation is supported, in part, by the absence of such a trend in the reported reliability data from repeated measures of ERA for both test groups. Since ERA is a measure of involuntary physiologic change, it seems reasonable to expect some attenuation on the effects of learning relative to the other voluntary measures.

Analysis of Variance

Of primary interest in this study were comparisons across levels of each of the main factors or independent variables for each of the test groups. Therefore, test methods, test frequencies and test sessions were comparisons of interest. A factorial design with associated analysis of variance procedures provided for the above comparisons.

Homogeneity of Variance

A basic assumption underlying the use of analysis of variance is that the error variance among treatment populations is the same. A statistical test of this assumption, comparisons among error variances (estimated from sample values) was made. The estimated error variances for each of the test groups are shown in Table 7.

Table 7. Estimates of error variance within treatment populations.

Test Tone	Group I CPA	Group I CRA	Group I ERA	Group II CRA	Group II ERA
500	127.24	215.00	139.00	254.00	191.00
1000	106.00	135.02	125.21	209.09	185.23
2000	101.00	65.29	65.93	112.15	68.89

Hartleys' F-max test[9] was used to evaluate the homogeneity of variance hypothesis for each test group. The computation for Group I gave a value of 2.06 with a critical value of 3.89 required for rejection. Group II variances yielded a value of 3.69 with a critical value of 3.76 required for rejection. Thus for a 0.01-level test the hypothesis of homogeneity of variance was not rejected for either group.

F-Tests and Individual Comparisons

The present analysis of variance procedures were carried out independently for the data obtained from each test group.

A three dimensional (3X2X3) repeated measures analysis of variance[10] design was employed with Group I. The main

---

[9] B. J. Winer, Statistical Principles in Experimental Design (New York: McGraw-Hill Book Co., 1962), pp. 92-96

[10] Winer, op. cit., pp. 337-349

factors were test frequencies (500-1000-2000 cps), test sessions (test, retest), and test methods (CPA, CRA, ERA). A summary of this analysis is given in the following table:

Table 8. Summary of analysis of variance comparing differences between test frequencies, sessions and methods for Group I.

Source of Variance	df	Mean Square	F
A (Frequency)	2	7602.6389	58.7889**
B (Session)	1	2.2223	.0172
C (Method)	2	397.6389	3.0748*
AB	2	.9719	.0075
AC	4	9.5139	.0735
BC	2	5.1385	.0397
ABC	4	3.2640	.0252
S/ABC	162	129.3209	

** Significant beyond the 0.01-level
 * Significant beyond the 0.05-level
$P[F\ (df=2/162 > 3.04] = .05$

As can be noted, two factors (frequency and method) showed statistical significance. No significant interactions or other significant main effects were obtained.

<u>Frequency</u>. Summing across sessions and methods the mean hearing levels obtained by frequency were 82.75 dB at 500 cps, 94.66 dB at 1000 cps, and 105.25 at 2000 cps. The differences between each pair of frequencies are shown in

Table 9. The use of a priori individual comparison procedures (k = 3, df = 59)[11] showed critical differences to be 4.15 (0.05-level) and 5.52 (0.01-level).

Table 9. Mean differences between pairs of frequencies across sessions and methods for Group I.

cps	500	1000	2000
500	--	12.0*	22.5*
1000	--	--	10.6*

* Significant beyond the 0.01-level

These results indicate significant differences between all frequency means. As the frequency increased so did the obtained hearing levels associated with them which is typical of the pattern of loss associated with a sensorineural impairment.

Method. Mean hearing levels obtained by methods, across sessions and frequencies, were 91.25 dB for CPA, 95.66 dB for CRA, and 95.75 for ERA. Differences between means are shown in Table 10.

Test of the significance of the difference in means for the pairs of treatments showed that the mean differences between CPA and both CRA and ERA exceeded the critical

[11]Winer, *ibid*, p. 85

difference of 4.15 for significance at the 0.05-level. The mean difference of .04 between CRA and ERA procedures was non-significant.

Table 10. Mean differences between pairs of methods across sessions and frequencies for Group I.

Method	CPA	CRA	ERA
CPA	--	4.41*	4.50*
CRA	--	--	.04

*Significant beyond the 0.05-level.

On the basis of a strict statistical interpretation these results indicate a significant difference between thresholds obtained by conventional play audiometry and those resulting from the experimental CRA procedure and ERA. The clinical implications of these findings is discussed later in this section.

Group II data were analyzed using a three dimensional (2X2X3) repeated measures analysis of variance design.[12] The results of this analysis are presented in Table 11. Two main factors, frequency and methods, showed statistical significance. No other significant main effects or

[12]Winer, ibid, p. 337-349

significant interactions were obtained.

Table 11. Summary of analysis of variance comparing differences between test frequencies, sessions and methods for Group II.

Source of Variance	df	Mean Square	F
A (Frequency)	2	1563.3333	8.9404*
B (Session)	1	.8333	.0047
C (Method)	1	1540.8333	8.8117*
AB	2	20.8333	.1191
AC	2	5.8333	.0333
BC	1	40.8333	.2335
ABC	2	1.6666	.0095
S/ABC	108	174.8611	

* Significant beyond the 0.01-level
$P[F\ (df\ 2/108) > 4.82) = .01$
$P[F\ (df\ 1/108) > 6.90) = .01$

Frequency. The mean hearing levels for frequencies, obtained by summing across levels of the other factors were 97.25 dB at 500 cps, 103.25 dB at 1000 cps, and 109.75 at 2000 cps. The differences between each pair of means are shown in Table 12. A priori test procedures (k = 3, df = 39) showed critical differences for significance between the means to be 5.98 (0.05-level) and 7.99 (0.01-level). As can be noted, all three frequency comparisons reached or exceeded the 0.05-level of significance.

Table 12. Mean differences between pairs of frequencies across sessions and methods for Group II.

cps	500	1000	2000
500	--	6.0**	12.5*
1000	--	--	6.5**

**Significant at or beyond the 0.05-level
*Significant beyond the 0.01-level

Method. Summing across frequencies and sessions the mean hearing levels obtained by methods were 99.83 dB for CRA and 107 dB for ERA, with a mean difference of 7.17 dB. This exceeded the critical difference of 6.42 for significance at the 0.01-level, (k = 2, df = 59).

Descriptive Statistics

In addition to the inferential statistics utilized in the above comparisons, a number of descriptive statistics were calculated. This was done in an attempt to gain additional information regarding factors that may have influenced present results and to facilitate comparisons with other studies.

Mean threshold levels, standard deviations and percent of thresholds falling within the traditional audiometric error of ± 5 dB are reported under the appropriate test-group headings below.

Test Group I

Mean threshold levels and the standard deviation of threshold scores were computed for each test method employed with Group I and are shown in the following table:

Table 13. Mean threshold levels and standard deviations of threshold scores, within and across test sessions, obtained under CPA, CRA, and ERA procedures for Group I.

	CPA		CRA		ERA	
	M	SD	M	SD	M	SD
Test						
500	79.5	11.9	85.0	15.4	84.0	12.4
1000	92.5	10.9	95.0	12.2	96.5	11.8
2000	103.0	10.6	106.5	8.5	107.0	8.6
Retest						
500	79.0	13.1	84.5	15.3	84.5	11.4
1000	92.0	11.3	95.0	9.8	96.5	14.0
2000	101.5	9.4	107.5	8.2	106.0	9.4
Test / Retest						
500	79.25	12.2	84.75	15.0	84.25	11.6
1000	92.25	10.8	95.25	10.9	96.5	12.6
2000	102.25	9.8	107.0	8.1	106.5	8.7

An examination of the dispersion of the threshold levels obtained under the various test methods revealed that CPA and CRA methods were in closest agreement at the 1000 cps test tone. Across test sessions, at that frequency the mean threshold difference was 3 dB and the standard deviations were closely approximated. A similar trend was apparent, although less pronounced, for the comparison of CPA and ERA results. In addition, there was a noticeable trend toward decreasing variation with increasing test-tone frequency. There was no apparent difference between the means and variances associated with test occasions.

Test-retest and inter-method threshold levels were inspected to determine the percentage of agreement with the accepted clinical error criterion of ±5 dB. In an attempt to further delineate the variation occurring between methods, attention was also directed to the percent of scores falling within a ±10 dB range. These data are presented in the table below.

Table 14. Percent of repeated measure and inter-method threshold differences falling within ± 5 dB and ± 10 dB variance ranges for Group I.

	CPA		CRA		ERA	
	± 5	± 10	± 5	± 10	± 5	± 10
Test/retest						
500	100%	---	100%	---	90%	10%
1000	90%	10%	90%	10%	90%	10%
2000	90%	10%	100%	---	100%	---
	CPA/CRA		CPA/ERA		CRA/ERA	
	± 5	± 10	± 5	± 10	± 5	± 10
Inter-method						
500	75%	95%	85%	90%	80%	95%
1000	90%	100%	70%	90%	75%	95%
2000	70%	95%	80%	90%	80%	100%

It was of interest to note that the above percentages indicated that stability on repeated measures was approximately the same for all methods. However, by frequencies, 10% of the CPA scores exceeded the clinical range at 2000 cps, whereas 10% of the ERA scores did so at 500 cps.

Examination of individual-subject thresholds between methods revealed that thresholds obtained by CPA and CRA, with the 1000 cps test tone, did not vary beyond a range of $\pm$ 10 dB.

Test Group II

The mean threshold levels and standard deviations of threshold scores for each of the test methods employed with Group II are shown in Table 15.

As can be noted from a comparison of Tables 13 and 14, the patterns of the variations for CRA and ERA, across frequencies and test occasions, are similar to those reported for Group I. That is, variance tends to decrease with increasing test tone frequency and there are no apparent differences between means associated with test-retest performance. However, it is apparent that mean threshold level differences, between methods for this group, are significantly larger than those reported under CRA and ERA for Group I. As previously reported, analysis of variance procedures indicated these differences to be significant beyond the 0.01-level of confidence with the mean difference, across sessions and frequencies equal to 7 dB. It is also of interest to note that these between-method differences are somewhat larger on initial test administrations.

Table 15. Mean threshold levels and standard deviations of threshold scores, within and across test sessions, obtained under CRA and ERA procedures for Group II.

	CRA		ERA	
	M	SD	M	SD
Test				
500	94.0	16.7	102.0	14.5
1000	99.0	15.2	106.5	14.3
2000	104.5	11.2	114.0	8.7
Retest				
500	93.75	14.1	99.5	13.0
1000	101.0	13.0	106.5	14.0
2000	107.0	10.8	113.5	10.0
Test / Retest				
500	93.75	16.8	100.75	13.5
1000	100.0	13.9	106.5	14.0
2000	105.75	10.8	113.75	9.1

The percentages of threshold scores falling within ± 5 and 10 dB clinical variance ranges were calculated for this group and are presented in Table 16.

As was the case with Group I none of the test-retest difference scores were greater than 10 dB in either direction. However, unlike Group I, in the test-retest comparisons ERA appeared more consistent with CRA for the high and low frequency test tones.

Inter-method comparisons of CRA/ERA thresholds revealed a greater degree of variation for this group than was previously reported for Group I. As can be noted, 20% of these subjects' threshold scores exceeded the ± 10 dB range of clinical variance, whereas only 5% did so in Group I. It would appear, on the basis that ERA demonstrated slightly less test-retest variation, that the major part of this inter-method variability was associated with the CRA method. However, inspection of the raw scores revealed that all such differences were in the direction of higher relative thresholds for ERA, at all test tone frequencies.

Table 16. Percent of repeated measures and inter-method threshold differences falling within ± 5 dB and ± 10 dB variance ranges for Group II.

	CRA		ERA			CRA/ERA	
Test-Retest	± 5	± 10	± 5	± 10	Inter-Method	± 5	± 10
500	90%	10%	100%	---	500	70%	80%
1000	80%	20%	80%	20%	1000	75%	80%
2000	90%	10%	100%	---	2000	60%	80%

Discussion

Several questions were posed in the previous chapter regarding the efficiency and effectiveness of the experimental CRA procedures in obtaining threshold measurement with young and non-communicative children. These questions are restated in the following sections with an effort made to answer them in terms of the present results.

One of the primary questions asked was whether CRA test procedures would produce stable or reliable results over repeated administrations to the population of interest. An attempt was made to answer this question employing techniques of analysis that would facilitate a comparison of the data with results obtained from similar studies. Reliability coefficients (r) and standard errors of measurement (Se) were computed relative to test methods and test tone frequencies. Correlations ranging from .89 to .99 were obtained for all methods, at all frequency levels, among both test groups. These were all of sufficient magnitude to indicate an adequate degree of test-retest reliability for all methods, with both groups. Standard errors for the two test groups, across methods and frequencies, ranged from a low of 1.41 for ERA to 2000 cps to a high of 3.33 which was also obtained for ERA but at the 500 frequency. This range indicated a relatively small amount of measurement error for all test methods. For example, doubling the largest standard error (3.33) would permit predicting true scores to be within $\pm$ 6.66 dB of obtained scores, which would be close to the accepted clinical error range. It was also relevant to this question to note that the standard errors for CRA, at all frequencies, remained relatively stable over both test groups. Similarly, standard deviations were computed for CRA scores, between test occasions, and were found to vary by no more than 2.5 dB for either group. These positive results were further substantiated by the analysis of variance procedures which indicated there were no significant differences between CRA scores related to test sessions.

On the basis of these findings it appears that CRA test procedures can be employed to obtain reliable threshold measurements with young and non-communicative children, as defined by this study. It is recognized that there are some restrictions to be observed in extending this assumption to other populations. Of primary import is the fact that the test phase of CRA is contingent upon success in establishing stimulus control during the conditioning phase. Failure to do so, of course, prohibits the use of the test. As previously pointed out, this situation occurred with the initial selection of some subjects for Group II. Three were unable to demonstrate an adequate degree of stimulus control after a comparatively large number of conditioning trials and were replaced under the rationale that the purpose of the study was to assess the CRA test procedures rather than the conditioning phase. However, it should be noted that these three children were subsequently referred for medical and psychological testing and were diagnosed as brain-injured. Thus there appears to be some potential for the initial phase of CRA to provide important preliminary diagnostic information to the speech and hearing clinician by pointing up those children whose primary problem may not be deafness.

The second question asked was whether threshold levels obtained by CRA procedures were comparable to those obtained by conventional play audiometry techniques. This question could not be assessed with the primary population of interest who, by definition, were untestable by conventional methods. Consequently the problem was approached indirectly by employing a matched group of subjects (Group I) who could respond to conventional test procedures.

The results of analysis of variance procedures employed with the test-retest data for Group I indicated significant differences for the two main factors; frequency and methods. Frequency differences were found to be significant at the 0.01-level which undoubtedly reflects the effects of the pronounced sloping pattern characteristic of sensorineural hearing loss. Method differences attained significance at the 0.05-level. Tests of the differences of individual pairs of treatments indicated CPA/CRA and CPA/ERA comparisons to be significant at the 0.05-level but CRA/ERA to be nonsignificant.

As previously pointed out, a strict statistical interpretation of these results indicates that thresholds obtained by CRA and ERA procedures are not directly comparable to those obtained by conventional methods.

However, it was considered important to equate and evaluate these findings relative to critical procedural differences existing between the methods. This involved the procedures and criterion employed to establish threshold levels. Under CPA "Threshold" was defined as a point of uncertainty (50% responses), and this intermittent kind of discrimination was socially reinforced. Conversely, under the operant procedures employed in CRA, reinforcement was contingent upon 100% stimulus discrimination. It can be assumed, from this, that the subjects under each method would develop a different "preparatory set" or "attentiveness" toward the stimulus. On the basis of these facts it was postulated that the threshold-level differences between these methods may represent the difference between the thresholds of detectibility (something different than nothing) and the threshold of perceptibility (recognition of tonality). Pollack studied this factor with normal hearing subjects and found that the variation between such thresholds could be as large as 6.5 dB.[13] Consequentially it was hypothesized

---

[13]I. Pollack. "The Atonal Interval." Journal of the Acoustical Society of America, 20 (1948), pp. 146-149

that the obtained statistical difference between methods reflects this "response-criterion" difference. Since this represents a relatively constant factor, rather than random variation, it would not invalidate the clinical usefulness of the CRA procedure. While it was recognized that this study did not provide a sufficient number of repeated measures to readily identify and define such a factor, all relevant data were carefully inspected and the results of other analyses were considered in an attempt to further assess the validity and clinical usefulness of this procedure.

Inter-method coefficients between CPA and CRA for the frequencies 500, 1000 and 2000 cps were .91, .94 and .92 respectively. These were significant from zero beyond the 0.01-level and indicate linear prediction is possible between these measures. Inspection of mean threshold levels across sessions indicated corresponding mean differences for CPA/CRA and CPA/ERA of approximately 4.5 dB. It was also of some interest to note that in a CPA/CRA comparison of inter-method threshold levels 75% of the obtained differences at 500 cps were within a variance range of ± 5 dB; 90% at 1000 cps and 70% at 2000. It was noted with equal interest that at 1000 cps all threshold differences were within a ± 10 dB range and at 500 and 2000 cps only 5% of the total difference scores exceeded this range. The relative import of these results can best be emphasized through comparison with studies of similar type but involving populations that are generally considered to be more easily and accurately tested. For example, Witting and Hughson obtained a minimum of ten audiograms of 7 normal and 17 hard of hearing adults, of assumed normal intelligence, and reported 23.5% and 29% of these subjects respectively demonstrated threshold variations in excess of ± 5 dB between the initial test and the average of the retests.[14] It should be remembered that these comparisons were between repeated measures of the same test, whereas the 10% variation referred to in the present study was between scores on different measures.

In consideration of these facts it would appear appropriate to assume that this second research question can only be answered, by this study, in a restricted sense. That is, that the experimental CRA procedures produce thresholds that are statistically different from those obtained by conventional methods. However, there is evidence to suggest that thresholds are clinically comparable and the differences represent a systematic increment with thresholds tending to be elevated under CRA methods by approximately 5 dB. On this basis it can also be postulated that in consideration of the accepted clinical error, CRA thresholds relative to conventional scores, could be expected to range from equality to a plus 10 dB. There is also evidence to

---

[14]Witting and Hughson, op. cit., pp. 259-269

suggest that maximum intra- and inter-method agreement occurs with the use of 1000 cps test tone.

The third and last primary question asked in this study concerned the validity of CRA thresholds obtained from young children who have been demonstrated untestable by conventional methods. As previously pointed out, it was necessary to use an indirect approach to questions concerning the test Group II population. As a consequence, ERA procedures were included with both groups in an attempt to establish the efficacy of this method through Group I and thereby provide an objective hearing measurement technique that could substantiate the results of the CRA analysis with Group II. Since ERA results with Group I did not differ significantly from those reported for CRA, its effectiveness, relative to the following analysis can be assumed with similar qualifications.

Group II mean threshold differences for CRA/ERA methods were significant at the 0.01-level as tested by analysis of variance. Since inter-method coefficients were of sufficient magnitude to indicate a linear relationship between these methods, these significant differences were interpreted in the same manner as previously described for Group I results. In other words, it was assumed that the CRA/ERA thresholds obtained for this population differed by a constant amount.

It was also hypothesized that since significant CRA/ERA differences were not obtained for Group I the factors influencing this variation for Group II may involve some subjective variables reflecting the primary difference between the groups. Up to this point that group difference has been generally classified as the "ability to respond to conventional hearing test methods." However, as previously pointed out, the type of test-response behavior elicited is in some part contingent upon the subjects' "set" or "attentiveness". Thus, rather than differing in "ability to respond", the groups can be considered as differing in some qualitative degree of set or attentiveness to the stimuli. Research was cited in Chapter III which indicates that the detectibility of auditory evoked responses is enhanced when subjects are "attending" to the stimulus (identifying some component). On the basis of these facts it can be postulated that since Group II could not achieve an appropriate or consistent degree of set or attentiveness to respond to conventional hearing test methods, this condition would also prevail with the passive, unreinforced ERA procedures. The consequence would be failure to detect evoked responses until the stimulus intensity reached some level above that at which thresholds of detectibility or perceptibility might be obtained by other measures. It is also relevant to this point to note that the previous review of the literature indicates ERA audiograms most often approximate those obtained by conventional means to a level about 15 dB above threshold. Interestingly this is close to the 12 dB difference Hirsh[15] suggests separates the threshold of detectibility and that of intelligibility (point at which 50% of speech test items can be repeated).

An inspection of mean threshold differences revealed that, across frequencies, these ranged from 6.5 to 8 dB. Thus, proceeding on the basis of assumptions derived in the evaluation of Group I data, it can be postulated that with this specific population, thresholds obtained by ERA differ from those obtained by CRA by an average of 7 dB. However, there is some additional evidence to indicate that this estimate might be slightly low. The data on percent of scores falling within ± 5 and 10 dB ranges indicate that, over all frequencies, 20% of the inter-method difference scores exceeded the top of this range. Raw score comparisons reveal maximum variations to be 20 dB in the direction of elevated thresholds for ERA. Thus there is evidence to suggest that the average difference of ERA over CRA scores may be closer to 10 dB. If, again, consideration is given to the added effects of accepted clinical error, it can then

---

[15] I. J. Hirsh, The Measurement of Hearing. (New York: McGraw-Hill Book Company, 1952), p. 170

be postulated that with this population, ERA thresholds relative to CRA, may be expected to range from plus 5 to plus 15 dB higher.

In summary, it appears that reliable and valid hearing thresholds can be obtained by CRA procedures from young and non-communicable children whose primary problem may be deafness. Such thresholds can be considered valid in the sense that they represent a consistent approximation of hearing levels which might be obtained by conventional hearing test methods. There is evidence to suggest that the threshold difference between CRA and CPA is a reflection of the difference in criterion measures employed to behaviorally define threshold under the two methods.

It is also apparent that there are some young and non-communicative children whose hearing ability cannot be adequately assessed by CRA methods. These are children, such as the three subjects initially replaced in Group II, for whom reinforcement contingencies and stimulus control cannot be readily established or consistently maintained. However, such subjects may represent a group for whom hearing loss is not, educationally, their primary problem. That is, children with CNS dysfunctions and concomitant learning problems whose speech and language development might not be appropriately enhanced through auditory amplification and speech reading techniques. If this is the case, then it appears that the initial phase of CRA can be instrumental in avoiding inappropriate placement and training by pointing up those children who cannot be conditioned to respond consistently to any of the three stimuli employed and in expediting their referral for other kinds of diagnostic services.

CHAPTER VI

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The basic purpose of this research was to develop and evaluate hearing test procedures, incorporating principles of operant conditioning, which would have immediate and widespread clinical application for use with young and non-communicative children.

The resulting procedures were labeled Conditioned Response Audiometry, and questions were posed regarding the efficacy of this approach with the population of interest.

SUMMARY

The experimental CRA method utilized a two-phase structure. The first phase was limited to conditioning procedures directed toward establishing stimulus control. The second phase was concerned with the measurement of hearing thresholds. The conditioning procedures carried out in the initial phase involved a multisensory approach using visual, tactile and auditory stimuli. The technique consisted of the simultaneous presentation of the three signals, modeling the response and administration of a tangible reinforcer. The modeling was faded out as stimulus control was established. At that point the visual stimulus was removed. If the response was maintained the tactile stimulus was eliminated. When conditioning was maintained to the auditory stimulus, the second or testing phase of CRA was initiated.

To facilitate the widespread use of the CRA approach, the techniques used in both phases were structured to be administered employing a portable audiometer. However, since the portable type instrument is not generally equipped to permit the simultaneous presentation of visual, tactile and auditory stimuli an adaptor was developed to accomplish this function. It was called a Shaping Adaptor for Portable Audiometers (SHAPA).

The second or testing phase of the CRA approach was structured to utilize the conditioned response developed during the initial phase. Two testers were employed to check and maintain the validity of subject responses. One tester was positioned in front of the subject and functioned to maintain subject attention, signal for stimulus presentations and administer reinforcement. The second tester, with the audiometer, was positioned behind the subject. His function was to present a schedule of test tones randomly interspersed with no-stimulus presentations. Any response occurring during a no-stimulus period indicated the subject was responding to something other than the test tone and signaled the termination of testing. Remedial procedures were then instituted to establish appropriate stimulus control.

Twenty subjects were selected from the total population of a community pre-school deaf education program. These subjects were assigned to test groups on the basis of their ability (Group I), or failure (Group II), to respond consistently and appropriately to conventional pure tone hearing test procedures. The groups were matched on age, educational experience and category of loss.

Three factors were of interest in this study: (a) an evaluation of test-retest reliability of threshold levels obtained by CRA procedures; (b) a comparison of hearing threshold levels obtained by CRA methods with thresholds obtained by conventional play audiometry procedures; and, (c) the determination of the validity of CRA threshold levels obtained from young children who had been demonstrated to be untestable by conventional hearing test procedures. Since the later two factors were incompatible by definition, the use of two test groups and the inclusion of an objective test measure (ERA) was thought to provide an indirect approach to the assessment of these variables.

Three test measures were employed with Group I. These included conventional play audiometry (CPA), evoked response audiometry (ERA) and conditioned response audiometry (CRA). The results of these tests were employed as criterion measures for the results of ERA and CRA measurements obtained from Group II. All tests were administered twice to each subject to assess the reliability of the procedures within the two groups.

The results of correlation coefficients computed for threshold levels obtained from test/retest sessions for both groups indicated a relatively high degree of reliability for CRA procedures. Standard errors of measurement computed for the same data revealed the dispersion of expected scores for both groups were close to the standard clinical error range of $\pm$ 5 dB.

Inter-method correlations obtained between CPA, CRA and ERA indicated a high degree of linear prediction was possible between thresholds obtained by all three methods at any of the three frequency levels.

Analysis of variance procedures were carried out independently for the data obtained from each test group. The three main factors of interest were methods, test frequencies and test sessions.

The results for Group I indicated there were significant differences between threshold levels obtained by methods and frequency. No significant interactions or other significant main effects were obtained. A priori individual comparison procedures indicated significant differences between thresholds obtained at all frequency levels. Similar individual comparisons revealed significant differences between threshold levels obtained by CPA and both CRA and ERA. However, no significant differences were found between threshold levels obtained by CRA and ERA.

Analysis of variance results for Group II also indicated significant differences between threshold levels for frequency and methods.

Descriptive statistics computed for Group I revealed that mean threshold level differences, over all frequencies, between CPA and CRA ranged from 3 to 6 dB for both test and retest occasions. Standard deviations for both methods, over all frequencies, were within 3 dB.

Percentage of repeated measures and inter-method threshold differences falling within a ± 5 dB and a ± 10 dB variance range indicated that the test/retest stability for all methods was approximately the same and that thresholds obtained by CPA and CRA with the 1000 cps test tone did not vary beyond the ± 10 dB range.

Similar comparisons made for Group II revealed that the mean threshold level differences for CRA and ERA ranged from approximately 5 to 10 dB. Standard deviations between the methods were within 3 dB.

Inter-method comparisons of the percentage of CRA/ERA threshold differences falling within a $\pm$ 5 or 10 dB range revealed a greater degree of variation for this group than was found for Group I. Twenty percent of the scores for Group II exceeded the $\pm$ 10 dB range, whereas only five percent did so in Group I. Inspection of the raw scores revealed that this variation was all in the direction of higher relative thresholds for ERA, at all test frequencies.

CONCLUSIONS

Within the limits imposed by the test methods and procedures employed and the specific populations involved the following conclusions seem warranted:

1. CRA test procedures can be employed to obtain reliable threshold measurements over test occasions separated in time.

2. Some young and non-communicative children, as defined by the present study, cannot be conditioned to respond consistently to any one of the three stimuli employed in the initial phase of CRA. As a consequence, the testing phase of CRA cannot be employed with such children. However, there is some evidence to suggest that such "failure to achieve conditioning" may be indicative of a group of children for whom the symptoms associated with deafness are a reflection of a more pervasive CNS disorder. Thus it appears that the initial

phase of CRA might be instrumental in pointing up children for whom hearing loss is not, educationally, their primary problem, in avoiding the possibility of inappropriate placement and training and consequently in expediting referral for other kinds of diagnostic services.

3. Threshold levels obtained by CRA test procedures are linearly related to threshold levels obtained by conventional hearing test methods. The results of the present study indicate that the difference factor may represent a "response-criterion variable" resulting from differences inherent in the measurement techniques employed. In other words, the threshold difference between CPA and CRA appears to reflect the difference in the kinds of psychophysical thresholds measured by each of the methods.

RECOMMENDATIONS FOR FURTHER RESEARCH

To expand and facilitate the use of CRA test procedures, further research should be directed toward quantifying the difference in obtained threshold levels under CPA and CRA methods. It would also be of interest to provide further evidence regarding the hypothesis that inter-method differences represent a response-criterion variable. This might best be approached by retesting the population of this study, with CPA and CRA procedures, and recording all thresholds as the lowest stimulus level at which 100% stimulus-discrimination was last achieved. This should equate the measurement criteria and result in a close approximation of the thresholds obtained between methods.

The CRA, Phase I, procedures should be studied to identify and delineate differences in establishing and maintaining stimulus control which might be of differential-diagnostic significance. Some of the specific questions which might be asked are as follows: is failure to achieve or maintain conditioning indicative of brain damage; is there a significant difference between the mean number of trials necessary to achieve conditioning for deaf children and children classified as mentally retarded.

Since CRA test procedures were designed for use in field settings where limited numbers of professionally trained personnel might be available, it would be of some advantage to redesign the test procedures and to develop instrumentation that would permit application by a single tester. This could be accomplished through the use of an adaptor for the test phase similar to the one employed in the conditioning phase. The adaptor would function as an electronic switch, with a no-stimulus system built in, to interrupt the test signal from the audiometer to the ear phones. The no-stimulus system would involve the use of an electronic flip-flop circuit that would establish a random pattern of stimulus presentations. Procedures would involve the tester operating a #1 button to present a tone and, if the subject initiated a response, subsequently depressing a #2 button that would activate a visual signal if the no-stimulus circuit was closed and the tone was being presented to the subject. In other words, as a result of the electronic no-stimulus circuit, the tester would not know when he depressed the stimulus-presentation button (#1) whether a tone was being presented to the subject until he checked it by depressing the #2 button. It would be necessary for the tester to wear ear muffs and plugs.

BIBLIOGRAPHY

American Academy of Ophthalmology and Otolaryngology, American Speech and Hearing Association, and American Academy of Pediatrics, "Joint Committee Statement On Infant Hearing Screening," ASHA, 13 (1971), p. 79.

Barr, B. "Pure Tone Audiometry for Pre-School Children," Acta-Otolaryngologica, Supplementum 121 (1955), pp. 1-84.

Bloomer, Harlan. "A Simple Method of Testing the Hearing of Small Children." Journal of Speech Disorders, 58 (1949), pp. 751-759.

Bordley, J. E. and Hardy, W. G. "A Study in Objective Audiometry With the Use of Psychogalvanometric Response." Annals of Otology, Rhinology, and Laryngology, 58 (1959), pp. 751-760.

Bricker, W. and Bricker, D. "A Programmed Approach to Operant Audiometry for Low-Functioning Children." Journal of Speech and Hearing Disorders, 34 (1969), pp. 312-319.

Bricker, W. and Bricker, D. "Four Operant Procedures for Establishing Auditory Stimulus Control With Low-Functioning Children." American Journal of Mental Deficiency, 73 (1969), pp. 981-987.

Cody, D. and Bickford, R. G. "Cortical Audiometry: An Objective Method of Evaluating Auditory Acuity in Man." Mayo Clinic Proceedings, 40 (1965), pp. 273-287.

Davis, H. "Enhancement of Evoked Cortical Potentials in Humans Related to a Task Requiring a Decision." Science, 145 (1964), pp. 182-183.

Davis, H., et. al. "The Slow Response of the Human Cortex to Auditory Stimuli: Recovery Process." Journal of Electroencephalography and Clinical Neurophysiology, 21 (1966), pp. 105-113.

Davis, H. "Slow Cortical Responses Evoked by Acoustic Stimuli." Acta Otolaryngologica, 59 (1965), pp. 179-185.

Davis, H. and Goldstein, R., Chapter 8, Hearing and Deafness, (eds.) Davis, H. and Silverman, S. R. (New York: Holt, Rinehart and Winston, 1970), p. 238.

Davis, H. and Yoshie, N. "Human Evoked Cortical Responses to Auditory Stimuli," _Physiologist_, 6 (1963), p. 164.

Davis, H. and Zerlin, S. "Acoustic Relations of the Human Vertex Potential." _Journal of the Acoustic Society of America_, 39 (1966), pp. 109-116.

Davis, P. A. "Effects of Acoustic Stimulation on the Waking Human Brain." _Journal of Neurophysiology_, 2 (1939), pp. 494-499.

Dawson, G. D. "A Summation Technique for the Detection of Small Evoked Potentials." _Journal of Electroencephalography and Clinical Neurophysiology_, 6 (1954), pp. 65-84.

Derbyshire, A. J., _et. al_. "Audiometric Measurements by Electroencephalography." _Journal of Electroencephalography and Clinical Neurophysiology_, 8 (1956), pp. 467-478.

Derbyshire, A. J., _et. al_. "EEG Responses to Words." _ASHA_ 6 (1964), p. 392.

Derbyshire, A. J. and McDermott, M. "Further Contributions to the EEG Method of Evaluating Auditory Functions." _Laryngoscope_, 68 (1958), pp. 558-570.

Dix, M. R. and Hallpike, C. S. "The Peep Show: A New Technique for Pure Tone Audiometry in Young Children." _British Medical Journal_, 2 (1947), pp. 719-723.

DiCarlo, L. and Bradley, W. "A Simplified Auditory Test for Infants and Young Children." _Laryngoscope_, 71 (1961), pp. 628-646.

Downs, M. and Sterritt, G. "A Guide to Newborn and Infant Hearing Screening Programs." _Archives of Otolaryngology_, 85 (1967), pp. 15-22.

Education of the Deaf, The Challenge and the Charge, A Report of the National Conference on Education of the Deaf (Washington, D. C.: U. S. Government Printing Office, 1967), p. 72.

Ewertsen, H. "Teddy-bear Screening Audiometry for Babies." _Acta Otolaryngologica_, 61 (1966), pp. 279-280.

Ewing, A. W. G. Aphasia in Children. (London: Oxford Medical Publication, 1930).

Ewing, I. and Ewing, A. "The Ascertainment of Deafness in Infancy and Early Childhood." Journal of Laryngology and Otology, 59 (1944), pp. 309-333.

Frisina, D. R., Chapter 4, Modern Developments in Audiology, (ed.) Jerger, James (New York: Academic Press, 1963), p. 128.

Froding, C. "Acoustic Investigation of Newborn Infants." Acta Otolaryngologica, 52 (1960), pp. 31-40.

Froeschels, E. and Beebe, H. "Testing the Hearing of Newborn Infants," Archives of Otolaryngology, (1946), pp. 710-714.

Fulton, R. "Standard Pure Tone and Bekesy Audiometric Measures With the Mentally Retarded." American Journal of Mental Deficiency, 71 (1967), pp. 60-73.

Fulton, R. F. and Lloyd, L. L. Audiometry for the Retarded: With Implications for the Difficult-To-Test (Baltimore: The Williams and Wilkins Company, 1969), p. 13.

Fulton, R. T. and Spradlin, J. E. "Operant Audiology With Severely Retarded Children." Parsons Demonstration Project Report No. 90. (Kansas: Parsons State Hospital and Training Center, 1968).

Galambox, R., Rosenberg, P. E., and Glorig, A. "The Eyeblink Response as a Test of Hearing." Journal of Speech and Hearing Disorders, 18 (1953), pp. 373-378.

Gardner, M. "A Pulse-tone Clinical Audiometer." Journal of Acoustical Society of America, 19 (1947), pp. 178-190.

Glorig A., Audiometry: Principles and Practices. (Baltimore: The Williams and Wilkins Company, 1965, p. 89.

Goldstein, R. "Electrophysiologic Audiometry," in Modern Developments in Audiology, (ed.) J. Jerger, (New York: Academic Press, 1963).

Goldstein, R. and Derbyshire, A. J. "Suggestions for Terms Applied to Electrophysiologic Tests of Hearing." Journal of Speech and Hearing Disorders, 22 (1957), pp. 696-697.

Green, D. "The Pup-Show: A Simple, Inexpensive Modification of the Peep-Show." Journal of Speech and Hearing Disorders, 23 (1958), pp. 118-120.

Gross, M., et. al. "Auditory Evoked Response Comparison During Counting Clicks, and Reading." Journal of Electroencephalography and Clinical Neurophysiology, 18 (1965), pp. 451-454.

Guilford, J. P. Fundamental Statistics in Psychology and Education. (New York: McGraw-Hill Book Company, 1965) pp. 252-253.

Guilford, F. and Haug, C. "Diagnosis of Deafness in the Very Young Child." Archives of Otolaryngology, 55 (1952), pp. 101-106.

Guilder, R. P. and Hopkins, L. A. "Program for the Testing and Training of Auditory Function in the Small Deaf Child During Pre-School Years," (Part I) Volta Review, 37 (1935), pp. 5-11.

Hardy, J., Dougherty, A., and Hardy, W. "Auditory Screening of Infants." Annals of Otology, Rhinology, and Laryngology, 71 (1962), pp. 759-766.

Hardy, J., Dougherty, A. and Hardy, W. "Hearing Responses and Audiologic Screening in Infants." Journal of Pediatrics, 55 (1959), pp. 382-390.

Hardy, W. and Bordley, J. "Special Techniques in Testing the Hearing of Children." Journal of Speech and Hearing Disorders, 16 (1951), pp. 122-313.

Hardy, W. C. "Early Detection and Assessment," Proceedings of International Conference on Oral Education of the Deaf, Volume I (Washington, D. C.: Alexander Graham Bell Association for the Deaf, 1967), p. 8.

Haug, C. and Guilford, F. "Hearing Testing on the Very Young Child: Follow-Up Report on Testing the Hearing of 968 Pre-school Patients With the Pediacoumeter," Transactions of the American Academy of Ophthalmology and Otolaryngology, 64 (1960), pp. 269-271.

High, W. S., Glorig, A. and Nixon, J. "Estimating the Reliability of Auditory Threshold Measurements," Journal of Auditory Research, 1 (1961), pp. 247-262.

Hirsh, I. J. The Measurement of Hearing. (New York: McGraw-Hill Book Company, 1952).

Human Communications and Its Disorders - An Overview, A report prepared and published by the Subcommittee on Human Communication and Its Disorders (Bethesda, Maryland: U. S. Department of Health, Education and Welfare, 1969), p. 63.

Jerger, J. "Comment on Estimating the Reliability of Auditory Threshold Measurements." Journal of Auditory Research, 2 (1961), pp. 138-142.

LaCrosse, E. L. and Bidlake, H. "A Method to Test the Hearing of Mentally Retarded Children." Volta Review, 66 (1964), pp. 27-30.

Lentz, W. E. and McCandless, G. A. "Averaged Electro-encephalic Audiometry in Infants." Journal of Speech and Hearing Disorders, 36 (1971), pp. 17-28.

Lloyd, L. "Behavioral Audiometry Viewed As An Operant Procedure." Journal of Speech and Hearing Disorders, 31 (1966), pp. 128-136.

Lloyd, L. and Frisina, D. (eds.). The Audiologic Assessment of the Mentally Retarded (Parsons State Hospital and Training Center, 1965).

Lloyd, L., Reid, M. and McManis, D. "Pure Tone Reliability of Institutionalized MR Children." Report #51, (1967) Parsons State Hospital, Parsons, Kansas.

Lloyd, L., Spradlin, J., and Reid, M. "An Operant Audiometric Procedure for Difficult-to-Test Patients." Journal of Speech and Hearing Disorders, 33 (1968), pp. 236-245.

Marcus, R. "Hearing and Speech Problems in Children: Observations and Use of Electroencephalography." Archives of Otolaryngology, 53 (1951), pp. 131-146.

Marcus, R., Gibbs, E. and Gibbs, F. "Electroencephalography in the Diagnosis of Hearing Loss in the Very Young Child." Diseases of the Nervous System, 10 (1949), 170-173.

McCandless, G. A. and Best, L. "Evoked Responses to Auditory Stimuli in Man Using A Summing Computer." Journal of Speech and Hearing Research, 7 (1964), pp. 193-202.

McConnell, Freeman (ed.). Proceedings of the Conference on Current Practices in the Management of Deaf Infants (Nashville, Tennessee: The Bill Wilkerson Hearing and Speech Center, Vanderbilt University, 1968).

Meyerson, L. and Michael, J. "Assessment of Hearing by Operant Conditioning Techniques." In Proceedings International Congress on Education of the Deaf. (Washington, D. C.: Gallaudet College, 1963).

Mulholland, A. M. and Fellendorf, G. W. Final Report of the National Research Conference on Day Programs for Hearing Impaired Children (Washington, D. C.: Alexander Graham Bell Association for the Deaf, 1968), p. 23.

Murphy, K. "Ascertainment of Deafness in Children." Audecibel, 14 (1966), pp. 89-93.

Myklebust, H. Auditory Disorders in Children. (New York: Grune and Stratton, 1954).

Newby, H. Audiology, (New York: Appleton-Century-Crofts, 1964), p. 155.

O'Neill, J., Oyer, H., and Hillis, J. "Audiometric Procedures Used With Children." Journal of Speech and Hearing Disorders, 26 (1961), pp. 61-66.

Perin, C. T. "A Quantitative Investigation of the Delay-of-Reinforcement Gradient." Journal Experimental Psychology, 32 (1943), pp. 37-51.

Premack, D. "Prediction of the Comparative Reinforcement Values of Running and Drinking." Science, 139 (1963), pp. 1062-1063.

Pollack, I. "The Atonal Interval." Journal of the Acoustic Society of America, 20 (1948).

Price, L. "Evoked Response Audiometry: Some Considerations." Journal of Speech and Hearing Disorders, 34 (1969), pp. 137-141.

Price, L., et. al. "The Averaged Evoked Response to Auditory Stimulation." Journal of Speech and Hearing Research, 9 (1966), pp. 361-370.

Price, L. L. "Cortical-Evoked Response Audiometry." In, Audiometry for the Retarded, (eds.) R. Fulton and L. Lloyd. (Baltimore: The Williams & Wilkins Co., 1969), pp. 211-212.

Rapin, I., et. al. "Evoked Responses to Clicks and Tones of Varying Intensity in Waking Adults." Journal of Electroencephalography and Clinical Neurophysiology, 21 (1966), pp. 335-344.

Rapin, I. and Graziana, I. "Auditory-Evoked Responses in Normal, Brain Damaged and Deaf Infants." Neurology, 17 (1967), pp. 881-894.

Reddell, R. and Calvert, D. "Conditioned Audio-Visual Response Audiometry." The Voice, 16 (1967), pp. 52-57.

Richmond, J., Grossman, H. and Lustman, S. "Hearing Test for Newborn Infants." Pediatrics, 11 (1958), pp. 634-638.

Roeser, R. and Price, L. "Effects of Habituation on the Auditory Evoked Response." ASHA, 10 (1968), p. 396.

Rose, D. and Rittmanic, P. "Evoked Response Tests With Mentally Retarded." Archives of Otolaryngology, 88 (1968), pp. 495-498.

Rose, D. E. and Ruhn, H. B. "Some Characteristics of the Peak Latency and Amplitude of the Acoustically Evoked Response." Journal of Speech and Hearing Research, 9 (1966), pp. 412-422.

Rosenthal, R. "Self-Fulfilling Prophecy," Psychology Today, 2 (1968), pp. 44-51.

Silverman, S. R. and Lane, H. S., Chapter 16, Hearing and Deafness, (eds.) Davis, H. and Silverman, S. R. (New York: Holt, Rinehart and Winston, 1970), p. 393.

Skinner, B. F. "A Case History in Scientific Method." In Koch, S. (ed.), Psychology: A Study of a Science, Vol. 2 (New York: McGraw-Hill, 1959).

Smith, W. I. and Moore, J. W. Conditioning and Instrumental Learning. (New York: McGraw-Hill Book Co., 1966).

Spradlin, J. E., Locke, B. J. and Fulton, R. T. "Conditioning and Audiological Assessment." In Fulton, R. T. and Lloyd, L. (eds.), Audiometry for the Retarded (Baltimore: The Williams & Wilkins Co., 1969).

Standard Reference Zero for Calibration of Pure Tone Audiometers, ISO/R 389-1964.

Statten, P. and Wishart, D. "Pure-Tone Audiometry in Children: Psychogalvanic-Skin-Resistance and Peep Show." Annals of Otology, Rhinology, and Laryngology, 65 (1956), pp. 511-534.

Suzuki, T. and Asawa, I. "Evoked Potential of Waking Human Brain to Acoustic Stimuli." Acta Otolaryngologica, 24 (1962), pp. 217-224.

Suzuki, T. and Asawa, I. "Evoked Potential of Waking Human Brain to Acoustic Stimuli." Acta Otolaryngologica, 48 (1957), pp. 508-515.

Suzuki, T. and Ogiba, Y. "Conditioned Reflex Audiometry: A New Technique for Pure Tone Audiometry in Children Under Three Years of Age." Archives of Otolaryngology, 74 (1961), pp. 192-198.

Suzuki, T. and Sato, I. "Free Field Startle Response Audiometry." Annals of Otology, Rhinology and Laryngology, 70 (1961), pp. 997-1007.

Suzuki, T. and Taguchi, K. "Cerebral Evoked Response to Auditory Stimuli in Young Children During Sleep." Annals of Otology, Rhinology and Laryngology, 77 (1968), pp. 103-110.

Teas, D. C. "Analysis of Evoked and Ongoing Electrical Activity at the Scalp of Human Subjects." Journal of Speech and Hearing Research, 8 (1965), pp. 371-387.

Thorne, B. "Conditioning Children for Pure Tone Testing." Journal of Speech and Hearing Disorders, 27 (1962), pp. 84-85.

Type and Size of Educational Programs Attended by Hearing Impaired Students in the United States: 1968-69," Annual Survey of Hearing Impaired Children and Youth, Series D, Number 4 (Washington, D. C.: Office of Demographic Studies, Gallaudet College).

Waldrop, W. F. "A Puppet Show Hearing Test." Volta Review, 55 (1953), pp. 488-489.

Walter, W. G. "The Contingent Negative Variation: An Electrocortical Sign of Significant Association in the Human Brain." Science, 146 (1964), p. 434.

Walter, W. G., et. al. "Responses to Semantic Stimuli in the Human Brain." Journal of Electroencephalography and Clinical Neurophysiology, 19 (1965), p. 314.

Weaver, R. M. "The Use of Filmstrip Stories in Slide Show Audiometry." In, The Audiologic Assessment of the Mentally Retarded: Proceedings of a National Conference, (eds.) Lloyd, L. L. and Frisina, D. R. (Kansas: Parsons State Hospital and Training Center, 1965), pp. 71-88.

Webb, C., et. al. "Procedures for Evaluating the Hearing of the Mentally Retarded." Cooperative Research Project No. 1731, U. S. Office of Education, 1964.

Wedenberg, E. "Auditory Tests on New-born Infants." Acta Otolaryngologica, 46 (1956), pp. 446-461.

Williams, W. G. and Graham, J. T. "EEG Responses to Auditory Stimuli in Waking Children." Journal of Speech and Hearing Research, 6 (1963), pp. 57-62.

Winer, B. J. Statistical Principles in Experimental Design (New York: McGraw-Hill Co., 1962), pp. 337-349.

Withrow, F. B. and Goldstein, R. "An Electrophysiologic Procedure for Determination of Auditory Thresholds in Children." Laryngoscope, 68 (1958), pp. 1674-1699.

Witting, E. and Hughson, W. "Inherent Accuracy of a Series of Repeated Audiograms." Laryngoscope, 50 (1940), pp. 259-269.

APPENDIX A

RESULTS OF

CPA-CRA-ERA

TESTING FOR

GROUPS I & II

GROUP I

Subject	CPA* 500	1000	2000	CRA* 500	1000	2000	ERA* 500	1000	2000
#1									
Test	80	95	95	85	95	105	80	90	95
Retest	85	95	100	85	95	100	85	90	90
#2									
Test	90	95	105	90	90	110	85	90	105
Retest	85	90	105	90	95	110	85	90	105
#3									
Test	75	90	100	85	95	105	75	85	100
Retest	75	95	95	80	95	110	80	85	100
#4									
Test	90	105	110	90	110	110	90	105	110
Retest	90	110	110	95	105	110	95	105	110
#5									
Test	100	110	120	120	110	120	110	120	120
Retest	100	100	110	120	110	120	100	120	120
#6									
Test	75	80	90	75	80	100	80	90	110
Retest	70	80	95	75	90	105	85	90	105
#7									
Test	65	90	95	75	95	100	75	100	100
Retest	65	95	95	70	95	95	75	95	100
#8									
Test	80	100	120	90	110	120	95	110	120
Retest	85	100	120	85	105	120	95	120	120
#9									
Test	60	75	100	60	75	100	65	85	110
Retest	55	70	95	65	75	105	60	80	110
#10									
Test	80	85	95	80	90	95	85	90	100
Retest	80	85	90	80	90	100	85	90	100

*Hearing threshold levels in decibels (ISO-1964)

GROUP II

Subject	CRA* 500	CRA* 1000	CRA* 2000	ERA* 500	ERA* 1000	ERA* 2000
#1						
Test	100	100	105	110	110	120
Retest	105	100	105	110	110	120
#2						
Test	95	110	120	105	120	120
Retest	95	110	120	100	120	120
#3						
Test	100	110	110	110	110	120
Retest	100	105	120	110	120	120
#4						
Test	50	65	85	65	75	95
Retest	55	75	90	70	75	90
#5						
Test	85	85	100	90	90	110
Retest	90	85	100	85	90	110
#6						
Test	100	100	105	110	120	120
Retest	100	110	110	105	110	120
#7						
Test	100	95	95	110	105	110
Retest	95	95	95	105	105	110
#8						
Test	95	100	95	100	105	105
Retest	95	105	100	95	105	105
#9						
Test	110	120	120	110	120	120
Retest	100	120	120	110	120	120
#10						
Test	105	105	110	110	110	120
Retest	100	105	110	105	110	120

*Hearing Threshold Levels in Decibels (ISO-1964)

APPENDIX B

TOTAL CONDITIONING TIME

FOR GROUPS I & II

GROUP I

Subjects	Number of 30 minute conditioning sessions	Total hours
1	8	4
2	7	3.5
3	11	5.5
4	6	3
5	8	4
6	9	4.5
7	7	3.5
8	10	5
9	12	6
10	9	4.5

GROUP II

Subject	Number of 30 minute conditioning sessions	Total hours
1	9	4.5
2	10	5
3	9	4.5
4	7	3.5
5	6	3
6	11	5.5
7	9	4.5
8	14	7
9	12	6
10	10	5