THE RELIABILITY AND CONSTRUCT VALIDITY OF THE SELF-CONCEPT OF ACADEMIC ABILITY SCALE-FORM D FOR HEARING IMPAIRED STUDENTS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY
LEE M. JOINER

1966

This is to certify that the

thesis entitled
THE RELIABILITY AND CONSTRUCT VALIDITY OF THE SELF-CONCEPT OF ACADEMIC ABILITY SCALE-FORM D FOR HEARING IMPAIRED STUDENTS

presented by

LEE M. JOINER

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Special Education

W. Vernon Hicks
Major professor

Date May 20, 1966

O-169

ABSTRACT

THE RELIABILITY AND CONSTRUCT VALIDITY
OF THE SELF-CONCEPT OF ACADEMIC ABILITY SCALE-FORM D
FOR HEARING IMPAIRED STUDENTS

by Lee M. Joiner

This investigation was a pilot study to determine the reliability and construct validity of an eight-item scale (SCA-D) designed to assess the self-concept of academic ability of hearing impaired students. The SCA-D scale is a parallel form of the Michigan State University General Self-Concept of Ability Scale (GSCA). The SCA-D scale has a simple vocabulary and can be administered to the hearing impaired by using signs and finger spelling.

Research objectives included determining: (1) the equivalence of the SCA-D and the GSCA scales; (2) the stability of measurement of the SCA-D scale; (3) the reproducibility of the SCA-D scale; (4) the item to test homogeneity of each item in the scale; and (5) the construct validity of self-concept of academic ability with hearing impaired students.

Two groups of subjects were used: (1) a random sample (N = 97) of twelfth grade students in an urban public school system; and (2) all of the "academic" students in the Indiana School for the Deaf, grades 8-11 (N = 80).

To assess construct validity, five major variables besides self-concept of academic ability were examined: (1) grade point average; (2) IQ scores; (3) perceived parental evaluations; (4) perceived teacher evaluations; and (5) grade level.

Analysis of questionnaire and school record data involved the use of Z tests, t tests, F tests, skewness measures, kurtosis measures, Pearson Product Moment correlations, Phi coefficients, Loevinger's H_{it}, and Green's: (I), Rep, and Rep_{ind}.

Findings:

1. Coefficients of equivalence of the two forms (SCA-D, GSCA) ranged from .84 to .51 in three schools. The overall coefficient of equivalence was .75.

2. Responses to both the GSCA and SCA-D instruments form a normal distribution.

3. The coefficient of stability over six days for the SCA-D was .84.

4. Mean SCA-D scores on test and retest of the hearing impaired were equal.

5. The reproducibility of the SCA-D as measured by Green's summary statistic (I) was .90 for the hearing impaired and .93 for the non-impaired.

6. Each item of the scale was found to be homogeneous with the total test for both the hearing impaired and the non-impaired.

7. Six construct validity hypotheses derived from a social-psychological theory of learning were confirmed.

Major Conclusions:

Reliable and valid measures of self-concept of academic ability can be obtained with hearing impaired subjects under mass testing conditions. Also, instruments such as the SCA-D can be designed which

will permit comparisons to be made between the impaired and the non-impaired.

Construct validity findings which parallel those obtained with non-impaired students indicate that a general social-psychology of learning can be equally useful in understanding the learning of the deviant as it has been with "normals." No evidence was found in this research which would force professionals to seek a separate "social-psychology of learning" for the hearing impaired.

THE RELIABILITY AND CONSTRUCT VALIDITY OF THE SELF-CONCEPT OF ACADEMIC ABILITY SCALE-FORM D FOR HEARING IMPAIRED STUDENTS

By

Lee M. Joiner

A THESIS

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

DOCTOR OF PHILOSOPHY

College of Education
Department of Elementary and Special Education

1966

ACKNOWLEDGEMENTS

The author wishes to express his appreciation for the aid and assistance given during the past three years by Dr. William V. Hicks, Chairman of the Guidance Committee. Also, Drs. Donald Sellin, Douglas Gilmore, and Seymour Parker are due thanks for their efforts as members of the Guidance Committee.

Dr. Wilbur B. Brookover, Thesis Director, has been most generous in allowing the use, in this study, of data collected through U. S. Office of Education Cooperative Research Projects under his direction and in lending support to the development of social-psychological studies of impaired students. Association with Dr. Brookover has been a significant event in the author's doctoral studies.

Dr. Edsel L. Erickson, now associated with Western Michigan University, contributed greatly to this project by way of his ideas, friendship, and personal efforts. Special thanks are extended to him and to Mr. Jerry Crittenden who helped develop the major instrument used in this study and acted as translator and proctor in the testing of the hearing impaired students.

Dr. William McClure, Director of the Indiana School for the Deaf, was most cooperative in arranging for the testing of students in his school and offered many valuable comments regarding research procedures.

Many of the tedious difficulties which emerge during the course of a research project were avoided by enlisting the aid of Mr. Corwin Krugh and Mrs. Natalie Sproull. Those which could not be avoided were subdued by Miss Vicky Springer.

Finally, thanks is extended to the author's aunt, Doris M. Joiner, for her general encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

CHAPTER I

THE PROBLEM

Introduction

The development of reliable and valid instruments is a basic educational research activity. Without reliable and valid instruments, it is impossible to interpret research findings. Research projects organized around the use of untested instruments are unsatisfactory consumers of time, money, and effort. It is the purpose of this investigation to assess the reliability and construct validity of the Self-Concept of Academic Ability Form-D Scale for Hearing Impaired Students. General studies of the social-psychology of the hearing impaired which are now being planned will use this instrument.

Statement of the Problem

In a report to a national conference, Research Needs in the Vocational Rehabilitation of the Deaf (1960), sponsored by Gallaudet College and the U. S. Office of Vocational Rehabilitation, development of instruments for measuring social psychological factors was given high priority.[1] It was agreed that instruments are needed to study parental attitudes and educational attainment, attitudes toward

---

[1]Merril Rogers and Stephen Quigley (eds.), "Research Needs in the Vocational Rehabilitation of the Deaf," American Annals of the Deaf, CV (Sept., 1960), 335-370.

self, aspiration levels, vocational interests, the status of the family, etc.[2] Furthermore, members of the conference stressed the need for instruments which would yield comparable data from both impaired and non-impaired populations. "Existing tests could be used or modified, and tests specifically for the deaf could be constructed where necessary."[3]

Traditionally, comparisons of sociological, psychological and social-psychological data secured from impaired and non-impaired populations has not been possible because of differing instruments, populations, theories, designs, and definitions. The seriousness of this problem is noted by many scholars from the several areas of special education.[4] This problem is also seen as significant by scholars from the disciplines of sociology[5] and psychology.

Brookover and his associates have successfully conducted several developmental and cross-sectional studies of relationships between various

---

[2]Ibid.

[3]Ibid.

[4]Beatrice Wright, Physical Disability--A Psychological Approach (New York: Harper and Row, 1960); Lee Meyerson, "Physical Disability as a Social Psychological Problem," Journal of Social Issues, IV (1948), 1-10; Samuel Kirk, "Research in Education," Mental Retardation: A Review of Research, eds. Harvey Stevens and Rick Heber (Chicago: University of Chicago Press, 1964); Maynard C. Reynolds, "The Social Psychology of Exceptional Children: Part III in Terms of the Interaction of Exceptional Children with Other Persons," Exceptional Children, XXVI (1959), 243-247; William Geer and Evelyn Deno, "CEC and Legislation--Now and in the Future," Exceptional Children. XXXII (Nov., 1965), 187-194.

[5]Irving Goffman, Stigma: Notes on the Management of Spoiled Identity (Englewood Cliffs: Prentice Hall, 1963); Wilbur B. Brookover et al. Relationship of Self-Concept to Achievement in High School U.S. Office of Education, Cooperative Research Project No. 2831 (East Lansing: Michigan State University, to be published in September, 1966).

student conceptions of self and others, and their relationships to school performance and career development.[6] Other investigators have found further evidence for the validity of Brookover's models and instruments.[7,8,9] Brookover contends, however, that his statements concerning the social-psychology of learning and the accompanying instruments have only a modest validation because the research to date has largely been limited to students without known impairments; i.e., data must be obtained on a variety of impaired and non-impaired populations.[10]

The basic problem in developing instruments for use with the hearing impaired results from the language and communication deficiencies manifest by this group. If we design an instrument to measure a social-psychological construct, the language or concepts included in the item may be incomprehensible to the hearing impaired student. Error variance is therefore increased and our ability to reach conclusions is needlessly hindered. Conversely, if we design an instrument

---

[6]Brookover, loc. cit.

[7]Richard C. Towne and Lee M. Joiner, The Effects of Special Class Placement on the Self-Concept of Ability of the Educable Mentally Retarded, U.S. Office of Education, Cooperative Research Project No. 71-2073 (East Lansing: Michigan State University, to be published September, 1966).

[8]David L. Haarer, A Comparative Study of Self-Concept of Ability Between Institutionalized Delinquent Boys and Non-Delinquent Boys Enrolled in Public Schools, Ph.D. Thesis (East Lansing: Michigan State University, 1964).

[9]Edsel Erickson, Wilbur Brookover, Lee Joiner and Richard Towne, "A Social-Psychological Study of the Educable Mentally Retarded: An Educational Research Application of Symbolic Interactionism," A Report Prepared for the Council for Exceptional Children (Portland: By the Authors, 1965).

[10]Brookover, loc. cit.

strictly in terms of the limitations of the exceptional child, i.e., tamper with the wording and presentation of items, we have no assurance that this instrument will yield comparable and valid data. Therefore, assessment of the reliability of a modified instrument becomes the first necessary stage in social research with the hearing impaired.

In attempting to meet the need for instruments which will yield comparable data for impaired and non-impaired populations, Brookover, Erickson, and Joiner have modified the Self-Concept of Academic Ability Scale and other instruments. These modified instruments should make it possible to test a social-psychological theory of learning with impaired populations.

In summary, the problem of this research is based on the proposition that reliable and valid data from both impaired children and children not known to be impaired are essential for an adequate knowledge of the social-psychological context within which learning occurs. Restated, a methodological study focusing on the reliability and validity of an instrument designed to yield comparable data is necessary before a more definitive investigation contrasting social-psychological characteristics of impaired and non-impaired populations can be initiated.

Current Need for Social Psychological Studies of Hearing Impaired Children

Social-psychological perspectives on human behavior have received some exposition in the writings of scholars concerned with

exceptional children and special education.[11] Yet there is little evidence of either a significant trend toward social psychological theory testing of an empirical nature or the development of research procedures and instruments based on any one of several competing social psychologies. Despite the scarcity of empirical studies, it is quite likely that social psychological theories can be further tested with deviant populations and that the results of theory based research activity will lead to a better understanding of exceptional children.[12]

Empirical studies of the social psychological situation of the hearing impaired which stress self-definitions, perceptions of others, and role analysis are rarely encountered in the literature on hearing impairment. The situation is unfortunate because social psychological theories have been especially useful in suggesting ways for experimentally manipulating behavior. Recognition of the need for experimental research with the hearing impaired has been late in arriving but now appears in some writings in the field. Meyerson, for example, discusses

---

[11]See: Lee Meyerson, "Physical Disability as a Social Psychological Problem" Journal of Social Issues, IV (1948) 1-10. L. A. Dexter, "A Social Theory of Mental Deficiency" American Journal of Mental Deficiency, LXII (1958) 920-928. Merrill T. Hollingshead, "The Social Psychology of Exceptional Children: Part I in terms of the Characteristics of Exceptional Children" Exceptional Children, XXVI (1959) 137-140. Matthew J. Trippe, "The Social Psychology of Exceptional Children: Part II in terms of Factors in Society" Exceptional Children, XXVI (1959) 171-175. Maynard C. Reynolds, "The Social Psychology of Exceptional Children: Part III in terms of the Interaction of Exceptional Children with Other Persons" Exceptional Children, XXVI (1959) 243-247. Beatrice A. Wright, Physical Disability--A Psychological Approach, (New York: Harper & Row, 1960). Albert Cutler, "Place of Self Concept in the Education of the Physically Different Child" Exceptional Children, XXVIII (1962) 343-349. William C. Washburn, "Patterns of Self Concepts Related to Problems of Neuroses and Delinquency in Adolescents" Exceptional Children, XIX (1963) 341-347. Edward Frankell, "Effects of a Program of Advanced Summer Study on the Self Perception of Academically Talented High School Students" Exceptional Children, XXX (1964) 245-249.

[12]Brookover, loc. cit.

the type of research which has been conducted in attempting to establish a "psychology of impaired hearing."[13] He argued that although "testing-correlating-comparing strategies" yielding descriptions are fundamental in scientific investigation, an additional question ought to be asked: "Does the study contribute to the specification of the manipulable, environmental conditions under which the behavior investigated occurs, varies in strength or fails to occur . . ." "A major limitation hindering the further development of empirical documentation for a psychology of impaired hearing is the continued utilization in research of non-manipulable variables."[14]

It should be noted, in reference to Meyerson's comments, that in the relatively brief history of social-psychology, a rather formidible body of experimental research has accumulated. Most of these studies illustrate various experimental methods which have been devised to influence an individual's behavior.[15] Generally, the subject's behavior has been influenced or manipulated through the use of planned and controlled social interaction. One important part of Brookover's research, for example, determined the effectiveness of three treatment strategies directed toward "others" in the life of the student in bringing about behavioral changes on the part of the student himself.[16]

---

[13]Lee Meyerson, "A Psychology of Impaired Hearing," Psychology of Exceptional Children and Youth, ed. William Cruickshank (New York: Prentice-Hall, 1963), 118-191.

[14]Ibid., 175.

[15]Roger Brown, Social Psychology (New York: The Free Press, 1965).

[16]Wilbur Brookover, et al. Self-Concept of Ability and School Achievement, II, U.S. Office of Education, Cooperative Research Project No. 1636 (East Lansing: Michigan State University, 1965).

Another argument for the development of social-psychological studies of the hearing impaired is based on the fact that the hearing impaired are "visible" in a population. They can be readily singled-out or identified by others. Meyerson has developed a basis for a psychology of physical disability in which "visibility" by others and the evaluations of others are assigned a central position. In partial summary of his somatopsychology he states:

> It is society, far more than the condition of the body, which determines what a person will be permitted to do and how he will behave. All cultures place values upon certain aspects of physique, although different aspects of physique may be differentiated as important in different cultures, and different values may be assigned to the same variations. Nevertheless, certain generalizations may be made:
>
> 1. Physique is a social stimulus.
> 2. It arouses expectations for behavior.
> 3. It is one of the criteria for assigning a person to a social role.
> 4. It influences the person's perception of himself both directly through comparison with others and indirectly through others' expectations of him.
> 5. Comprehension of the kind, extent, and degree of socially imposed handicaps on persons with atypical physiques is basic to an understanding of the somatopsychology of physical disability.[17]

On the basis of the above passage, it appears that studies of the person's perception of himself and the evaluations of his capacity and capabilities by others are needed.

The contention that social-psychological studies should be carried out which deal with the self-other interactions of the hearing

---

[17]Lee Meyerson, "Somatopsychology of Physical Disability," Psychology of Exceptional Children and Youth, ed. William Cruikshank (New York: Prentice-Hall, 1963), 15-16.

impaired is supported, again, by recent research findings. Evaluations and expectations which are communicated in interaction between the hearing impaired student and others attain special potency because the hearing impaired seem to rely heavily on others for evaluative judgments as well as behavioral decisions.[18] A study of the vocational status and adjustment of deaf women concluded with the suggestion that the interaction between parents, child, and school personnel be explored in research.[19] By starting at this point the field of "others" is narrowed from what might include all people in the general language community or participants in the common culture to those who repeatedly interact with the subject.

General Theoretical Orientation of the Study*

The major theoretical propositions which have been investigated in the Brookover studies at Michigan State University and upon which the present research is based are:[20]

---

[18]Richard L. Blanton and Jum C. Nunally, "Educational Language Processes in the Deaf," Psychological Reports, XV (1964), 891-894.

[19]Joseph Rosenstein and Alan Lerman, Vocational Status and Adjustment of Deaf Women (New York: Lexington School for the Deaf), 1963.

*This section is based on the theoretical orientation presented in: Edsel L. Erickson, Wilbur B. Brookover, Lee M. Joiner, and Richard C. Towne, "A Social-Psychological Study of the Educable Mentally Retarded: An Educational Research Application of Symbolic Interactionism," A Report Prepared for the Council for Exceptional Children (Portland: By the Authors, 1965).

[20]The development of these theoretical propositions are presented more fully in Brookover, LePere, Hamachek, Thomas and Erickson (1965), Op. Cit. and in Brookover, "A Social Psychological Conception of Classroom Learning." School and Society, LXXVII (1959) 84-87.

1. The self-concept of ability is a "functionally limiting" variable in school performance.[21]
2. Students, including the hearing impaired, develop their self-concept of ability largely through their perceptions of how their "significant others" evaluate their ability.[22]
3. The student must believe that engaging in a task is the appropriate thing to do.[23]
4. Whether or not a given task is viewed as appropriate by the student is dependent upon his "self-identity" in relation to others.[24]

Perceived evaluations of ability mentioned in proposition number 2 are not considered "prescriptive."[25] That is to say if a student perceives that a significant other evaluates him as being capable of performing a task at some set level, it is not always true that this significant other will expect the student to engage in that activity. For instance, a parent might evaluate his child as being

---

[21]Brookover et al. (1965) Op. Cit., Chapter I and VI.

[22]Ibid.

[23]George Herbert Mead, Mind, Self and Society (Chicago: University of Chicago Press, 1934), 150-151 and 364-369.

[24]Ibid.

[25]"Expectation" and "Prescription" are used synonymously and have reference to the particular behaviors an individual perceives others in his social system expect of him because he occupies a particular position (i.e. friends and child) in reference to them. Except as to the addition of the term "perceived" this use of "Expectations" is similar to that of Ralph Linton, The Study of Man (New York: Appleton-Century-Crofts, Inc., 1936) 113-114; Robert K. Merton, Social Theory and Social Structure, (Glencoe, Ill.: The Free Press, 1957) 369.

able to succeed in college on the basis of his knowledge of the child's performance but might not expect him to attend college because of the financial situation of the family or conflicting interests shown by the child.

Symbolic interaction theory serves as a basis for derivation of the above propositions and also posits that a person ". . . enters his own experience as a 'self' or individual not directly . . ." but by taking the standpoint of others.[26] Thus, in order for a person to intentionally act to achieve in a given task he must see the task as appropriate, its appropriateness being determined by his self-identity in relation to others. His self-identity as well as the expectation that he engage in the activity are social emergents or concepts which are established through communicative interaction between the student and others. Yet it is possible that even when an activity is considered appropriate by the student and he is expected to engage in the activity by his parents, he may still not act to achieve. He may have learned that the activity, whether he does it or not, will have little effect upon the social relationships he presently values or aspires to attain. If this is so, the likelihood of his pursuing the activity may be substantially decreased.

---

[26]George Herbert Mead, Mind, Self and Society. (Chicago: University of Chicago Press, 1934) 138 and 225. "The individual experiences himself as such, not directly, but only indirectly, from the particular standpoints of other individual members of the same social group, or from the generalized standpoint of the social group as a whole to which he belongs. For he enters his own experience as a self or individual not directly or immediately, not by becoming a subject to himself, but only insofar as he first becomes an object to himself just as other individuals are objects to him or in his experience; and he becomes an object to himself only by taking the attitudes of other individuals toward him within a social environment or context of experience and behavior in which both he and they are involved."

Three basic relationships have been explored in previous research with non-impaired children carried out by Brookover et al. These are the relationships of student-peer (friend), student-parent, and student-teacher. The rationale for examining these relationships is drawn from the early work of Ralph Linton and the recent work of Robert K. Merton on "Role Set."[27] Using the deaf as an illustration, the deaf child moves among relationships with his parents, teachers, and friends and in so doing may maintain a common set of role expectancies emerging from his being defined as "deaf." Although at any given moment the deaf student may be physically in proximity with any one of a number of others, he symbolically takes into account other relationships which he has experienced. In this fashion a self-identity as "deaf" or hearing impaired along with the previous evaluations and expectations relating to this condition are carried into new situations.

In examining the parent-student, friend-student, and teacher-student relationships a useful theoretical concept is that of reciprocal role relationships. A relationship is termed a "reciprocal role relationship" when it is based on a reciprocity of actions; when an individual "enacts a social role which is defined with reference to another role, as in the relationship between patient and doctor . . ."[28] or between

---

[27]Ralph Linton, The Study of Man, (New York: Appleton-Century-Crofts, Inc., 1936) 113-114. Op. Cit.; Robert K. Merton, Social Theory and Social Structure, (Glencoe, Ill.: The Free Press, 1957): The theoretical importance of stressing research upon "student role" is developed in Chapter XVI in Brookover and Gottlieb op. cit.

[28]Herbert C. Kelman, "Processes of Opinion Change," Public Opinion Quarterly, XXV (Spring, 1961), 64.

hearing impaired student and special class teacher. Kelman notes that ". . . if an individual finds a particular relationship satisfying, he will tend to behave in such a way as to meet the expectations of the other." He behaves in terms of his perception of the demands of that relationship. A hearing impaired student may learn that he is obliged to achieve at a high level in school if he is to maintain his present relationships with others or if he desires to establish a new relationship involving expectations for academic achievement.

In summary, the general social psychology behind the present study is symbolic interactional. Symbolic interaction theory is concerned with the genesis of "self" and its place in organized human behavior. Self arises through the internalization of symbolic gestures and involves the individual's perception of the expectations which others hold for his behavior. In this study, self-concept of academic ability represents the individual's view of himself as a student. Reciprocal-role relationships, normative influences, and the idea that performance in a task is in part compliant behavior are all factors in the arisal of self-concepts.

Development of the Self-Concept of Academic Ability-Form D Scale

In the fall of 1965 the author of the present study discussed the possibility of developing reliable and valid scales for the social-psychological study of the hearing impaired with Edsel Erickson and Vivian Stevenson. The scales were to be alternate forms of the instruments used in the Brookover studies, thus enabling comparisons to be made between impaired and non-impaired groups. The development of these scales would also permit future experimental research with

exceptional children to be conducted.

It was decided that an inter-university (Western Michigan University and Michigan State University) research project on the hearing impaired and the visually impaired would be initiated, involving students in schools in Michigan and Indiana. (See Appendix C for abstract). The relevance and appropriateness of the project was established on three grounds: (1) the need for social-psychological studies of exceptional children, (2) the existence of well formulated theory which had been tested with non-impaired subjects, (3) the availability of longitudinal, normative, data on non-impaired students.

The major instrument used in the Brookover studies has been the General Self-Concept of Academic Ability Scale (see Appendix D). This is an eight item Guttman-type scale for which the reliability and validity characteristics have already been determined.[29] Development of an alternate form for use with hearing impaired subjects involved modifying the language of the eight items in such a way that they could be directly translated into manual signs. An easy reading vocabulary level was also desired since some subjects were expected to need the double reinforcement of the printed word and the sign. By developing the scale with these two criteria in mind it was hoped that the likelihood of obtaining meaningful responses would be increased. Similar alterations were also performed on other scales. The characteristics of these scales, however, are not the subject of the present

---

[29]Brookover et al. op. cit.; Paterson, op. cit.

study. Future reports will be addressed to the reliability and validity of the other measures.

Mr. Jerry Critterden, an NDEA doctoral fellow at Michigan State University, played a major part in the development of these alternate forms. With his help, each item was examined and several simplified alternatives proposed, the decision as to selection being made on the basis of signing ease and "trueness" to the intended meaning of the original item. Response alternatives were reduced from five to three as it was observed in prior research by the present author that finely drawn comparative statements are not always fully understood by children with language problems.[30]

---

[30]Towne and Joiner, op. cit.

CHAPTER II

METHODS FOR DETERMINING RELIABILITY AND VALIDITY

Introduction

The meaning of reliability can be best understood by recognizing that both chance error and consistency enter into measurement. Chance error means that a series of measurements will never yield values which are exactly alike. Obtained scores or values for the measurement typically vary in large or small units depending upon the size of the unit of measurement. Variation from measurement to measurement, or chance error, has several possible sources. Included among the sources of chance errors:

I. Lasting and general characteristics of the individual
e.g. general ability to comprehend instruction
II. Lasting but specific characteristics of the individual
e.g. knowledge and skills specific to particular form of test items
III. Temporary but general characteristics of the individual
e.g. understanding of mechanics of testing
IV. Temporary and specific characteristics of the individual
e.g. unpredictable fluctuations in attention or accuracy
V. Systematic or chance factors affecting the administration of the test or the appraisal of test performance
e.g. conditions of testing--adherence to time limits, freedom from distractions, clarity of instruction, etc.
VI. Variance not otherwise accounted for (chance)
e.g. "luck" in selection of answers by "guessing"[1]

Despite the fact that error is involved in the scores when measures are repeated, there remains a tendency toward consistency.

---

[1]Robert L. Thorndike, "Reliability," Educational Measurement, ed. E. F. Lindquist, (Washington: American Council on Education, 1961), 568.

The objects or events measured tend to maintain their same relative position in reference to other objects or events.

High self-concepts of academic ability on the first test, for example, tend to remain among the highest on subsequent tests. "This tendency toward consistency from one set of measurements to another is the reverse of the fact of variation which we have just considered, and will be designated 'reliability.'"[2] Kerlinger lists the following synonyms for reliability: dependability, stability, consistency, predictability, accuracy.[3]

Construct validity is a concept which denotes a greater concern for the event being measured than the instrument which is used to measure it. As opposed to validation of the test, the researcher interested in construct validity attempts to validate a theory. The concept itself has been considered a "significant advance because it unites psychometric notions with scientific theoretical notions."[4]

Determining the Equivalence of Two Test Forms

Test constructors are often interested in designing a series of tests which assess the same skill, knowledge, or attribute but which do not exactly repeat the item content from one test to the next. Educational programs are usually concerned with developmental phenomena and periodical testing has been used as a means for determining development or change. The routine of yearly achievement tests is illustrative

---

[2] Ibid.

[3] Fred Kerlinger, Foundations of Behavioral Research, (New York: Holt, Rinehart and Winston, Inc., 1964), 429.

[4] Ibid. 448.

of the process. When repeated measures of achievement occur, however, it is obvious that if the same items are used time and time again, the student's performance would be likely to increase due to practice in the problems. To remedy this situation, several forms of the same test are devised by drawing samples of items from a pool of items assumed to be representative of a fixed content level. Try-outs are then undertaken to determine the equivalence of the forms.

The present study was concerned with a slightly different problem. Rather than sampling a pool of items representing some common content, an effort was made to maintain the integrity of items which were used with the non-impaired students. The vocabulary of the items in the original scale necessitated simplification before they could be used with hearing impaired students. A standard "signing" of the original items would have been impossible. The question which must be subsequently asked is, "do the two forms of the test result in a similar ordering of individuals on each?" Do the subjects who score high on the SCA remain high on the SCA-D and vice versa? According to Thorndike, "If we have two forms of a test, we may give each pupil first one form and then the other. They may follow each other immediately if we are not interested in stability over time, or may be separated by an interval if we are. The correlation between the two forms will provide an appropriate reliability coefficient."[5] "Coefficient of equivalence" is the term generally used for alternate form correlations for tests not separated by a time interval.

---

[5]Robert Thorndike and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education, (New York: John Wiley and Sons, Inc., 1961) 178.

Assessing the Stability of Measures by Test and Re-test

Stability of measurement, a more accurate and descriptive term than reliability for test-retest evaluation, is estimated by repeatedly administering the same test or test procedure and calculating the correlation between observed scores. Research reports in the literature of education frequently include test-retest coefficients. Often, these coefficients are the only "reliability" measures presented. A serious question can be raised regarding the sufficiency of test-retest correlations as a general index of test "reliability."

Thorndike states that "in most measures of intellect, temperment, or achievement . . . repetition of the same test form and correlation of the two sets of scores is less defensible as an operation for determining reliability (than test-retest with alternate forms). ". . . Repeating the same test form holds the sampling of items constant so that this factor is treated as systematic rather than error variance."[6]

Another confounding factor, when test-retest correlations are used as the only estimate of reliability, is that the subjects may recall their previous responses. Spuriously high reliability coefficients may be a consequence of holding the sampling of items constant and risking the subject's remembering his former response.

The point of this brief discussion is that "reliability" is always partially a result of the method or index used to determine and describe it and a stability index alone is not a sufficient reliability

---

[6]Robert Thorndike, "Reliability," Educational Measurement, ed. E. F. Lindquist, (Washington: Annual Council on Education, 1961), 577-578.

index. In this sense the test-retest correlation coefficient is only an observed fact; an indication of the extent to which the same ordering of subjects is obtained on two separate administrations of the same test. In interpreting this fact, consideration must be given to the conditions under which the coefficient was obtained.

In the present pilot project, several major changes were introduced in the second testing. A full description of the changes appears under "Details of Testing Procedures for the Hearing Impaired" in Chapter III. Changes in the testing situation were instituted in an attempt to minimize some of the sources of chance error listed in the introduction to this chapter. Among them were: (1) modification to increase the likelihood that all students would comprehend the instructions, (2) a greater understanding of the mechanics of testing on the part of the subjects due to practice, (3) freeing the subjects from unnecessary distractions in the testing situation.

Therefore, caution must be exercised in interpreting test-retest correlations as absolute indices of stability. The coefficients are likely to be spuriously high as a result of the subject possibly remembering his previous responses. Also the variance due to item sampling is treated as systematic rather than error variance. Counterbalancing this tendency toward a spuriously high stability estimate is the fact that testing conditions were not completely constant. Minimal estimates of stability of measurement are likely to be obtained under this condition.

One final point should be made in reference to stability estimates with exceptional children. When language difficulties are prominent

among a group, the researcher is forced to immediately doubt whether the subjects understood the questions which he asked them. One logical method for finding out whether or not the subjects understood the questions is to ask the same question at two separate times and to see whether or not the subject's responses agree. From this point of view the test-retest correlation of identical test forms is an index of the subject's consistency in response. "Understandings" of the item is then inferred from the consistency shown between responses. If individual subjects did not understand the item it would be reasonable to expect random responses to the item. Random responses on the part of a large segment of the subjects would reduce the test-retest correlation coefficient since it would be akin to correlating random events.[7]

Internal Consistency Reliability

At least three general approaches have been employed in determining the reliability of an instrument through analysis of its internal consistency. These methods are: (1) "Rational Equivalence" reliability (i.e. Hoyt's Analysis of Variance,[8] Kuder-Richardson Formula #20);[9]

---

[7]It is possible that "response set" might invalidate the idea that test-retest correlation between items which could not be understood would be near zero. One strategy for approaching this question would be to deliberately compose items containing vocabulary and concepts which are unfamiliar to the exceptional child. By calculating test-retest results, an estimate of "response set" could be obtained. For other approaches see G. C. Helmstadter, Principle of Psychological Measurement (New York: Appleton-Century-Crofts, Inc., 1964), 152-156.

[8]C. Hoyt, "Test Reliability Obtained by Analysis of Variance," Psychometrika, VI (1941), 153-160.

[9]G. F. Kuder and M. W. Richardson, "The Theory of Estimation of Test Reliability," Psychometrika, II (1937), 151-160.

(2)"Reproducibility" analysis (i.e. Guttman's,[10] Green's,[11] Loevinger's,[12] Jackson's,[13]); and (3) Empirical determination of sub-scales (i.e. McQuitty's Linkage Analysis,[14] Factor Analysis).

"Rational Equivalence" Techniques

Kuder-Richardson Formula #20 is probably the most well known and widely used technique for estimating the reliability of a test on the basis of a single administration.

$$r_{tt} = \frac{K}{K - 1} \cdot \frac{\sigma t^2 - K\overline{pq}}{\sigma t^2}$$

The Kuder-Richardson formula would be inappropriate for the present instrument analysis. Jackson and Ferguson contend that the derivation of this formula implicitly assumes that all items are of equal difficulty and furthermore that all items have equal standard deviations.[15] This is equivalent to saying that "there are at most

---

[10]L. Guttman, "The Cornell Technique for Scale and Intensity Analysis," Educ. Psychol. Msmt., VII (1947), 247-279.

[11]Reported in Benjamin White and Eli Saltz, "Measurement of Reproducibility," Psychological Bulletin, LIV (March, 1957), 81-99.

[12]Jane Loevinger, "A Systematic Approach to the Construction and Evaluation of Tests of Ability," Psychological Monographs, LXI (1947), No. 4.

[13]White and Saltz, Op. cit.

[14]Louis L. McQuitty, "Elementary Linkage Analysis for Isolating Orthogonal and Oblique Types and Typal Relevances," Educational and Psychological Measurement, XVII (Summer, 1957), 207-229.

[15]Robert Jackson and George Ferguson, "Studies on the Reliability of Tests," Bulletin No. 12, Dept. of Educational Research, University of Toronto, 1941.

two degrees of difficulty of items, that is, the number passing any item must equal either the number passing or failing any other item."[16] It can also be shown that perfect item inter-correlation is a necessary condition for obtaining perfect reliability using this method. "From the statement that reliability will equal one only if all the items are perfectly correlated and equal in difficulty, it is only one step to the statement that the reliability will equal one only if everyone has a score of zero or perfect. . . . Exactly as good results could be obtained by giving one item instead of the whole test."[17] A reproducible scale violates, by definition: (1) equal item difficulty, (2) inter-item correlations of unity, and (3) equal item variance. Therefore, Kuder-Richardson Formula #20 is ruled out.

Hoyt's Analysis of Variance, another "single test" reliability index is based on a definition of reliability as the ratio of true score variance to obtained score variance. Since this analysis has been used so often with the Self-Concept of Academic Ability Scale, Loevinger's cogent discussion of the limitations of the technique is included here. Loevinger writes as follows:

> His (Hoyt's) initial assumption is that the error component for each person on each item is normally distributed with the same variance as the error component in every other item. The error component is defined as the difference between the actual score and the true score of the person on the item. The true score is a constant based on the difficulty of the item and the ability of the person. Since the actual score on the item is either one or zero, and the true score is constant, the error component must equal

---

[16] Loevinger, op. cit. 11.

[17] Ibid. 12.

> either one minus the true score or simply minus the true score. The error component for any one person and any one item has only two possible values, which is a far departure from the normal curve. Moreover, the variance of the error component depends solely on the probability of the person passing the item; so the assumption of a constant variance for the error component is equivalent to the assumption that the probability of any person passing the item is a constant. Hoyt's assumptions are worse than Kuder and Richardson's; rather than simply restricting consideration to an unimportant special case, Hoyt has considered an impossible case, for his assumptions are mutually contradictory.[18]

Despite the serious limitations of this analysis when used with an instrument conceived of as a reproducible scale and scored dichotomously, the index will be calculated and the results consigned to Appendix B. By doing so, present findings can be compared with earlier reports.

Reproducibility

Cureton says that: "The most important requirement for a test whose scores are to be interpreted as measurements would seem to be that its items all draw upon the same sets of abilities or traits."[19] Tests which meet this requirement have been called: unified tests, uni-dimensional tests, univocal scores, reproducible scales and homogeneous tests. Not only have different names for tests which meet this requirement been coined, but several analytic devices are also available to determine if they do in fact meet the requirement.

---

[18]Loevinger, op. cit.

[19]Edward Cureton, "Quantitative Psychology as a Rational Science," Psychometrika, XI (1946), 191-196, cited in Jane Loevinger, "The Technic of Homogeneous Tests Compared with Some Aspects of 'Scale' Analysis and Factor Analysis" Psychological Bulletin, VL (1948), 507-529.

The necessity for developing uni-dimensional tests can be illustrated easily. Regardless how useful a test may be for predictive purposes or how "relationally fertile" it appears, it is of negligible value in theoretical research if it is composed of several obscure dimensions. When a theory is developed which includes self-concept of academic ability as its central construct any operationalization of the construct should deal with that dimension along. A scale could be devised which if untested for unidimensionality might include items tapping self-regard, internalized self-expectations, somatopsychological self-assessments, and social values. It would be termed a "multi-factor test" and conventional internal consistency analysis could result in a judgment of good reliability[20] test and retest correlations could be high. Yet when we discuss our results and further elaborate our theory it would be completely inaccurate and misleading to refer to self-concept of academic ability as a single variable. One would be hard pressed to define what is being discussed.

Apparently the question of whether or not the items in a test draw upon the same set of abilities or traits is usually approached indirectly through another question. Taking the Self-Concept of Academic Ability Scale-Form D as an example it can be shown that if responses are dichotomized (they originally appear trichotomously) there are 2^8 (256) different response patterns which may possibly

---

[20]See discussion and Multi-factor tests in G. C. Helmstadter, Principles of Psychological Measurement. (New York: Appleton-Century-Crofts, 1964), 101-103.

map a subject's item responses. A score of 2 can be obtained in 27 different ways. The secondary question is a direct outcome of the observation that there are many ways of obtaining any one total score. It asks: "To what extent can a pattern of responses to a psychological test be predicted on the basis of a subject's total score?"

Guttman's index of reproducibility (R) is the best known and most commonly used analysis procedure.[21] Basic to his procedure is the theoretical concept of "scalability." Scalability is not a characteristic of the test itself but is instead considered a "property of a universe of attributes" of which the items in a test are a sample.[22] In order to determine whether the universe of attributes represented by the sampled items are scalable it must be possible to order the items on the basis of subject response so that "persons who answer a given question favorably all have higher ranks than persons who answer the same question unfavorably."[23]

White and Saltz discuss several methods for determining the reproducibility of a psychological scale.[24] Each method attempts to provide a summary of the extent to which knowledge of a subject's total score enables reproduction of his response pattern. Guttman's

---

[21]Louis Guttman, "The Cornell Techniques for Scale and Intensity Analysis," Educ. Psychol. Msmt., VII (1947), 247-279.

[22]Loevinger, loc. cit. p. 511.

[23]L. Festinger and Katz, op. cit.

[24]Benjamin White and Eli Saltz, "Measurement of Reproducibility," Psychological Bulletin, LIV (March, 1957), 81-99.

(R) meets one important requirement for a good index in that it yields a theoretical maximum value which is the same for every test (1.00). It is deficient in the sense that it is subject to the effect of item difficulty or variations in percentage of persons passing each item. "The reproducibility figure can approach its absolute lower limit of 50 per cent only when all the items have a difficulty level of 50 per cent . . ." Any departure from this condition causes a rise in the lower limit of the reproducibility index.[25]

Guttman has suggested that an obtained coefficient (R) of .90 is necessary before a universe of attributes is considered scalable, but the difference between this coefficient and the lower limit of the coefficient may and does vary from scale to scale. On some occasions the (R) of .90 may be much higher than the minimum.[26]

Another great disadvantage of Guttman's technique, along with all other except Green's is that it does not permit evaluation of the null hypothesis that the obtained index is not significantly different from chance. There is no known method for obtaining an approximation of the standard error of (R).

Several techniques for scale analysis have been devised since the introduction of Guttman's (R) which are designed to correct the deficiencies in (R) which were noted. Jackson's Plus Percentage Ratio (PPR) and Loevinger's Index of Homogeneity (H) both yield indices which are uninfluenced by the distribution of item difficulties and permit specification of a minimum value. The advantage of these methods

---

[25] Ibid., 86.

[26] Ibid.

is incorporated into Green's Summary Statistics Method (I) which has the additional advantage of being amenable to the calculation of a standard error of measurement. Also, Green's (I) is the most straightforward and simple to calculate for large numbers of subjects or items. Therefore Green's method was selected for computing an index of reproducibility for the test as a whole.

The general formula for Green's Summary Statistics (I) is:

$$I = \frac{Rep - Rep_{ind}}{1.00 - Rep_{ind}}$$

In this formula Rep is the obtained reproducibility of the test for the sample or population which has responded to it.

$$Rep = 1 - \frac{1}{NK} \sum_{i=1}^{k-1} \overline{ni}\ ,\ i+1 - \frac{1}{NK} \sum_{i=2}^{k-2} \overline{\overline{n_i - 1},\ \overline{i},\ i+1,\ i+2},$$

Where N is the number of subjects, K the number of items. The quantity $n_{\overline{I}}$, i + 1 is the number of subjects who fail the i^{th} item and pass the next most difficult item (i + 1). The number of these pairs is k - 1. The quantity, $n\ \overline{\overline{i - 1},\ \overline{i},\ i+1,\ i+2}$ is the number of subjects having failed both items i - 1 and i and passed i + 1 and i + 2.

Green has developed the coefficient (Rep_{ind}) to describe the reproducibility that would be expected if the items were independent, i.e. exhibit zero covariance. If this were the case, knowledge of the total score and the order of difficulty of the items would lead to zero prediction of the subject's response pattern.

Rep_{ind}, the value of the reproducibility coefficient which would be obtained under the condition of zero covariance, is obtained

by the following formula:

$$Rep_{ind} = 1 - \frac{1}{N^2 K} \sum_{i=1}^{k-1} n_i \; ni + 1 - \frac{1}{N^4 K} \sum_{i=2}^{k-2} n_i \; n_i + 1n_i + 2n_{i+1}$$

According to Green, the value of I as calculated through the use of these three formulas should be .50 before the items of a test can be considered scalable. White and Saltz cite one study which showed that the average discrepancy between Green's index (I) and the exact reproducibility of ten scales was .002.

Green's formula for obtaining an approximation of the standard error of the obtained reproducibility of the scale (Rep) is:[27]

$$\sigma \; Rep \approx \sqrt{\frac{(1 - Rep)\ (Rep)}{N\ K}}$$

By utilizing the formula it is possible to establish the confidence limits within which the true value of Rep occurs. When values of a given statistic are calculated on the basis of a sample it is quite possible that the obtained value is different than the actual value for the population. A small standard error of measurement means that the sample statistics are clustering closer to the population parameter.[28]

---

[27]The discussion and formulas for Green's (I), Rep, Rep_{ind}, and Rep, are based on material covered in White and Saltz, loc. cit.

[28]Sanford H. Dornbusch and Calvin F. Schmid, A Primer of Social Statistics, (New York: McGraw-Hill, 1955), 120.

Item evaluation. The Self-Concept of Academic Ability-Form D scale is made up of eight items. Ordinarily, in scale construction, the investigator is faced with a set of items which he has composed and from which he proceeds to select the best ones. How to proceed in selecting the best ones has been a subject of much discussion and several methods have been proposed.[29]

In constructing and evaluating the alternate form of the Self-Concept of Academic Ability Scale, it was not necessary to commence with a pool of items and ask which set should be included in the final form. Instead, we ask: "Do the eight items which represent modifications of the original Brookover scale retain their former scale characteristics when revised so as to be directly "signable." The problem is not one of selecting the best items, but one of seeing whether the adapted items retain their goodness.

Loevinger suggests that one rational principle for item selection or item evaluation is that each item should be homogeneous with the total test.[30] A condition of perfect item to test homogeneity would be encountered when each item in the test is passed

---

[29]Much of this discussion falls under the rubric of "item validity," a confusing term which actually means the extent to which an item score correlates with the total test score. Biserial r, item by total chi-square, and point biserial are some commonly used methods. For a thorough treatment of conventional techniques see: Frederick B. Davis, "Item Selection Techniques," Educational Measurement, ed. E. F. Lindquist, (Washington: American Council on Education, 1961), 266-328.

[30]Loevinger, "A Systematic Approach to the Construction and Evaluation of Tests of Ability."

by individuals who have higher total scores than those who fail it. Perfect item to test heterogeneity appears when the total scores of individuals passing a certain item are randomly distributed within an ordering of total scores according to magnitude.

Two further points should be considered, according to Loevinger, if homogeneous scales are to be obtained. Whatever technique we apply in our assessment of homogeneity should remain unaffected by item difficulty since it is necessary to include items of varying difficulty or pass/fail ratio if we are to adequately describe varying levels of an attitude, knowledge, or other behavior. Secondly, the index should require no assumptions beyond ordinal measurement. The assumption of equal interval measurement is rarely justified in social and mental measurement. Therefore, the second condition sets realistic limits and emphasizes the use of widely applicable methodology.

The index H_{it} for item to test homogeneity meets the above stated requirements. Actually, it is a modification of an index proposed by Long.[31] Logically it is based on the ratio of wrong discriminations (i.e. subjects passing the item having lower total test scores than those failing the item or vice versa) to total discriminations made. It could be called the <u>percentage of wrong discriminations</u>.

---

[31]John Long, et al., "The Validation of Test Items," <u>Bulletin No. 3</u>. Department of Educational Research, University of Toronto, 1935.

Loevinger contends that Long's index generally meets the criteria for a useful item to test homogeneity index except for one major problem; what to do about the scores that are tied. Whenever the number of subjects is greater than the number of items in a dichotomously scored test, ties ensue. The important effect of ties can be illustrated as follows.

If two subjects have the same total score on a test but on the item under consideration one of them passes it and the other fails, elimination of that item from the test will result in the same score maintaining for the subject who failed the item, but a score lower by one point for the subject who passed it. "Thus, to eliminate a spurious correlation between item and total score, we must count all pairs of persons who are tied for total scores but who differ in item score as 'wrong discriminations' on the part of the item."[32] Another dilemma is encountered when subject A has a total score 1 point higher than subject B and subject A passed the item under question and subject B failed it. If we eliminate this item from the test, these two subjects will then have tied total scores.

Loevinger argues that the best solution to the tied scores problem is to eliminate tied pairs from the numerator and denominator of the "wrong discrimination/total discrimination" ratio. The number of instances of total scores, including the item in question, which are tied with one subject of the pair passing the item and the other failing the item are counted in the numerator. In the denominator,

---

[32]Loevinger, (1947), op. cit., p. 34.

the total number of subjects who passed the item and whose total score is one point above subjects who failed the item are counted. Her final formula becomes:

$$H_{it} = 1 - \frac{2\sum \text{"passes" below or tied with "fails"}}{PQ - \sum \text{"passes" one above fails}}$$

Where p is the # of passes on the item
q is the # of fails on the item

The expected value of this index given a condition of perfect item to test heterogeneity is zero.

White and Saltz advocate the use of the Phi Coefficient (ϕ_{it}) as a measure of item reproducibility. "This measure has the advantages of an absolute maximum of 1.00, an absolute minimum of 0.00, a known sampling distribution, and direct relationship to conventional test construction procedures."[33] Using this method, all subjects are ranked as to total score. For each item a cutting point along this ranked series of total scores is determined on the basis of the proportion of subjects passing and failing the item. An item score (pass-fail) by total score (high-low) table is made up for each item as follows:

		TOTAL SCORE		
		Low	High	Total
Item 1 Score	Pass Item J	A	B	A+B
	Fail Item J	C	D	C+D
		A+C	B+D	N

[33] White and Saltz, op. cit., 90.

Then the formula for Phi is:

$$\phi_{it} = \frac{BC-AD}{(A+B)\ (C+D)}$$

where A, B, C, and D are cell entries in the table.

A chi-square test is appropriate for the determining whether or not ϕ_{it} is statistically significant if $N < 30$. When $N > 30$, the significance test for the null hypothesis is $z = \phi\sqrt{n'}$

Empirical Determination of Sub-Scales

Factor analysis and cluster analysis for the determination of sub-scales and "purifying" a test are two methods commonly used in internal consistency studies. Paterson has used both McQuitty's method and centroid factor analysis with the eight items of the Self-Concept of Academic Scale administered to seventh grade non-impaired students.[34] The use of McQuitty's method led her to conclude that the items in the SCA Scale are best interpreted as representing a single cluster. A subsequent factor analysis again confirmed that "the SCA Scale is basically unidimensional with a time factor operating secondarily."[35]

Not only has the parent scale been shown to be basically unidimensional when factor or cluster analysis are applied, but there is also reason to believe that if a Phi-coefficient (ϕ_{it}) item analysis

---

[34]Ann Paterson, Evaluation of an Instrument Purporting to Measure the Construct Self-Concept of (Academic) Ability (unpublished Ph.D. dissertation, College of Education, Michigan State University, 1966, in process), p. 4-25.

[35]Ibid. 4-26.

is performed, a factor analysis becomes superfluous. Thurstone multiple factor analysis has been shown to result in sets of items which are nearly identical to those obtained when the Phi analysis is used.[36]

Construct Validity

Construct validity is one of four types of validity which have been mentioned in the literature on testing and measurement.[37] Other types of validity estimates are: content validity, concurrent validity and predictive validity. Rather than evaluating validity on the basis of how well an instrument predicts a specific outcome, its logical coherence, or its agreement with other measures of the same construct, construct validity is directly concerned with empirically testing hypotheses which involve the construct and thereby the theory which generates the hypotheses.

In the research on self-concept of academic ability, construct validity of the General Self-Concept of Academic Ability instrument was found when the correlations between general self-concept and theoretically relevant variables were examined."[38] This procedure

---

[36]White and Saltz, op. cit., 94.

[37]A comprehensive and basic reference is: American Educational Research Association, Committee on Test Standards, and National Councils on Measurement used in Education, Committee on Test Standards, "Techniques, "Psychological Bulletin, LI (March, 1954),Part 2.

[38]Wilbur B. Brookover, et al., "Improving Academic Achievement through Students' Self-Concept Enhancement," U.S. Office of Education Cooperative Research Project 2831, (East Lansing: Michigan State University, Bureau of Educational Research Services, 1965), 55.

is in agreement with Cronbach's three part elaboration of construct validation:

1. Suggesting what constructs might account for test performance. This is an act of imagination based on observation or logical study of the test.

2. Deriving testable hypotheses from the theory surrounding the construct. This is a purely logical operation.

3. Carrying out an empirical study to test this hypothesis.[38]

Kerlinger stresses the point that "it is not enough to predict, for instance, that the measure presumably reflecting the target property should be positively correlated with one theoretically relevant variable. One should, through deduction from the theory, predict more than one such positive relation. In addition, one should predict zero relations between the principle variable and variables 'irrelevant' to the theory."[39]

Helmstadter states that ". . . no single, numerical estimate of the degree of construct validity will be found. Rather, a wide variety of approaches and evidences may be used to support claims about what a test measures and consequently the ways in which it can be used."[40] Five general types of evidences for construct validity have been cited:

---

[38]Lee Cronbach, Essentials of Psychological Testing. (New York: Harper and Bros., 1960), 121.

[39]Fred Kerlinger, Foundations of Behavioral Research, (New York: Holt, Rinehart and Winston, Inc. 1964), 451.

[40]G. C. Helmstadter, Principles of Psychological Measurement, (New York: Appleton-Century-Crofts, 1964), 139-144.

1. Group Differences - i.e. "many traits are postulated in such a way that persons in different groups are conceived to possess different amounts of the characteristics involved."
2. Changes in Performance - i.e. "performance on an aptitude measure should remain the same in spite of the introduction of a training program, while performance on an achievement test in the same situation would increase."
3. Correlations - i.e. measures of a construct should correlate with measures of a logically related variable. Conversely, measures of a construct should not correlate with measures of a logically irrelevant variable.
4. Internal Consistency - High internal consistency is an argument for construct validity except in those cases where a complex measure is involved.*
5. Study of the Test-Taking Process - This involves performing the tasks which a subject is expected to do and determining if the actions required are relevant. It is especially useful, as a method, in developing tests which call for the manipulation of materials.

The construct validity hypotheses which follow are correlational and comparisons of group differences. Internal consistency, although sometimes viewed as evidence for construct validity, is dealt with in the section devoted to reliability analysis. In the present project internal consistency shall be viewed as primarily a measure of reliability. Changes in performance cannot be considered unless longitudinal data is available.

All of the following hypotheses except numbers 5 and 6 are replications of hypotheses tested with non-impaired subjects. The theoretical rationale for each has been treated extensively elsewhere

---

*A complex measure does not assess a single behavioral dimension, universe or trait. The notion of complex measures is somewhat inconsistent with the idea of a construct unless it can be assumed that drawing upon measures of several different constructs simultaneously results in a new hyper-construct.

and therefore will not be repeated here.[41]

Construct Validity Hypotheses:

1. The Self-Concept of Academic Ability of hearing impaired students in associated with their school performance.
2. The Self-Concept of Academic Ability of hearing impaired students is associated with perceived parental evaluations of academic ability.
3. The Self-Concept of Academic Ability of hearing impaired students is associated with perceived teacher evaluations of academic ability.
4. The Self-Concept of Academic Ability of hearing impaired students is associated with their measured intelligence.
5. The Self-Concept of Academic Ability of hearing impaired students is not associated with grade level.
6. The mean Self-Concept of Academic Ability of the hearing impaired equals the mean Self-Concept of Ability of the non-impaired.

Confirmation of the above hypotheses would give an indication of at least three things:

1. Self-Concept of academic ability as measured by this scale has construct validity with respect to certain central variables in the theory.

---

[41]Some basic references are:
Brookover, Wilbur B., Lepere, Jean M., Erickson, Edsel L., Thomas, Shailer. "Definitions of others, Self-Concept, and Academic Achievement: A Longitudinal Study," paper presented at the American Sociological Association at Chicago, Illinois, 1965.; Brookover, Wilbur B., Paterson, Ann and Thomas, Shailer. "Self-Concept of Ability and School Achievement," U.S. Office of Education Cooperative Research Project #845, (East Lansing: Office of Research and Publication, Michigan State University, 1962).; Erickson, Edsel L., and Thomas, Shailer "The Normative Influence of Parents and Friends Upon School Achievement," paper presented at meetings of American Educational Research Association, Chicago, Illinois, 1965.; Erickson, Edsel L., Brookover, Wilbur B., Joiner, Lee M., Towne, Richard C. "A Social-Psychological Study of the Educable Mentally Retarded: An Educational Application of Symbolic Interactionism," paper presented at 1965 national meeting of the Council for Exceptional Children, Portland, Oregon.; LePere, Jean M. "A Study of the Impact of the Parent on the Child's Concept of Self as it relates to Academic Achievement," paper presented at meetings of American Educational Research Association at Chicago, Illinois, 1964.; LePere, Jean M. "A Longitudinal Analysis of the Relationship of Self-Concept and Academic Achievement of Students from Seventh through Tenth Grades," paper presented at meetings of American Educational Research Association at Chicago, Illinois, 1965.

2. A social-psychology of learning which has been developed primarily for non-impaired students may have usefulness for understanding the hearing impaired.
3. Rather than being conceived of a totally "different," needing a special psychology or social psychology, the hearing impaired may be approached from the point of view of major similarities which they show with the non-impaired.

Cronbach and Meehl have pointed out that if construct validity hypotheses are not confirmed, three possible conclusions might be drawn:

1. The test does not measure the construct variable.
2. The theoretical network which generated the hypothesis is incorrect.
3. The experiment failed to test the hypothesis correctly.[42]

---

[42]L. Cronbach and P. Meehl, "Construct Validity in Psychological Tests," Psychological Bulletin, LII (1955), 295.

CHAPTER III

PROCEDURES

Non-Impaired Subjects Used in the Study

In November 1965 the first group of subjects was randomly selected from among the senior class of all high schools (3) in a midwestern city where mass testing was being undertaken as part of U.S. Office of Education Cooperative Research Project #2831. The testing took place in each of the three high schools and was completed within a period of three weeks. In supplement to the major questionnaire, the Self-Concept of Academic Ability Scale-Form D was randomly distributed to approximately 30 Caucasian students (15 male, 15 female) in each of the three schools. Out of a total group of approximately 1,500 students tested under the major project design, 97 were selected as subjects in the present study. Table I shows the mean self-concept of ability score, IQ score, and grade-point average for this group along with standard deviations and the distribution characteristics of each variable. Hereafter, the above sample will be referred to as the "Non-Impaired High School Seniors."

Table 1 IQ, Self-Concept of Academic Ability, and GPA: characteristics of the distributions for hearing impaired and non-impaired

Statistic	(Test #2) Hearing Impaired (N = 80)			Non-impaired (N = 97)		
	IQ	SCA-D	GPA	IQ	SCA-D	GPA
Mean	106.6	18.34	2.25	103.7	18.99	2.22
SD	11.5	3.34	.77	10.3	2.96	.72
Skewness	- 0.02	.24	-.03	.10	.08	.43
Kurtosis	.04	.17	.04	.08	- .15	- .04

Hearing Impaired Subjects Used in the Study

In December, 1965, testing was undertaken at the Indiana School for Deaf at Indianapolis, Indiana. Due to the fact that serious questions were raised as to the feasibility of mass testing some of the low-achieving hearing impaired students, it was decided that random sampling of the resident population would not be undertaken. Instead, all of the residential students in grades 8 (N = 23), 9 (N = 23), 10 (N = 22), and 11 (N = 12) who were in academic sections (programs) were tested en masse.

A total of 81 subjects were tested in the first session. In the second session 85 subjects were tested of whom 80 had previously been tested. Test-retest analyses, therefore, show an N of 80 subjects, while scale analyses show an N of either 81 or 85 subjects. Of the 80 subjects used for the correlational analyses 50 were male and 30 female. Only three of the subjects were Negro and therefore were not excluded from the analysis. If a large number of Negro subjects were present

it would have been necessary to treat them separately in the analyses since Morse found major differences between Negro and Caucasion students on self-concept of ability.[1]

Instruments were administered twice to this group with a test-retest interval of six days. Hereafter, the above population shall be referred to as the "Hearing Impaired Students-Test #1" or "Hearing Impaired Students-Test #2." The mean self-concept of academic ability scores, IQ and GPA are presented in Table I along with standard deviation and distribution characteristics of the variable.

Instruments

Although several scales and questions were used in the complete pilot project (see Appendix A for entire questionnaire) only three scales are relevant to the present report:

1. The Self-Concept of Academic Ability Scale-Form D for Hearing Impaired Students.
2. Perceived Parental Evaluation of Ability Scale-Form D for Hearing Impaired Students.*
3. Perceived Teacher Evaluation of Ability Scale-Form D for Hearing Impaired Students.

---

[1]Richard J. Morse, "Self-Concept of Ability, Significant Others and School Achievement of Eighth Grade Students: A Comparative Investigation of Negro and Caucasian Students," unpublished M.A. thesis, Michigan State University, 1963.

*Analysis of the perceived parental and teacher evaluation of ability scale is currently being conducted under the direction of Dr. Edsel L. Erickson at Western Michigan University.

The exact format of the three scales as presented to the subjects in the study is as follows:

Self-Concept of Academic Ability-Form D**

Pick one. Circle the letter for your answer.

1. Think of your friends. Do you think you can do school work better, the same, or poorer than your friends?
   a. better
   b. the same
   c. poorer

2. Think of the students in your class. Do you think you can do school work better, the same, or poorer than the students in your class?
   a. better
   b. the same
   c. poorer

3. When you graduate from high school, do you think you will be with the best students, average students, or below average students?
   a. the best
   b. average
   c. below average

4. Do you think you could graduate from college?
   a. yes
   b. maybe
   c. no

5. If you went to college, do you think you would be one of the best, average, or poorest students?
   a. the best
   b. average
   c. poorest

6. If you want to be a doctor or a teacher, you need more than 4 years of college. Do you think you could do that?
   a. yes
   b. maybe
   c. no

7. Forget how your teachers mark your work. How good do you think your own work is?
   a. excellent
   b. average
   c. below average

**Scale titles were not used in the questionnaire.

8. What marks do you think you <u>really</u> <u>can</u> get if you try?
   a. A's and B's
   b. B's and C's
   c. D's and E's

Perceived Parental Evaluations of Ability-Form D

Pretend you are your mother or father. Answer like they would. Pick one. Circle their answer.

1. Think of your mother and father. Do your mother and father say you can do school work better, the same, or poorer than your friends?
   a. better
   b. the same
   c. poorer

2. Would your mother and father say you would be with the best, average, or below average students when you graduate from high school?
   a. the best
   b. average
   c. below average

3. Do they think you could graduate from college?
   a. yes
   b. maybe
   c. no

4. Remember, you need more than four years of college to be a doctor or teacher. Do your mother and father think you could do that?
   a. yes
   b. maybe
   c. no

5. What grades do your mother and father think you <u>can</u> get?
   a. A's and B's
   b. B's and C's
   c. D's and E's

STOP
DO NOT TURN THE PAGE YET

Perceived Teacher Evaluation of Ability-Form D

Pretend you are your teacher. The one you like the best. Answer like he or she would. Pick one. Circle their answer.

1. Think of your teacher. Would your teacher say you can do school work better, the same or poorer than other people your age?
   a. better
   b. the same
   c. poorer

2. Would your teacher say you would be with the best, average or below average students when you graduate from high school?
   a. the best
   b. average
   c. below average

3. Does your teacher think you could graduate from college?
   a. yes
   b. maybe
   c. no

4. Remember you need more than four years of college to be a teacher or doctor. Does your teacher think you could do that?
   a. yes
   b. maybe
   c. no

5. What grades does your teacher think you can get?
   a. A's and B's
   b. B's and C's
   c. D's and E's

STOP
DO NOT TURN THE PAGE YET

Two additional variables were dealt with in the validity analysis; grade point average and intelligence. Grade point average represented an average of the subject's grades in social studies, English, science, and math for spring, 1965, and fall, 1965. Letter grades were utilized in reporting academic level at the Indiana School

for the Deaf so a numerical code was assinged to each letter grade: E = 0, D = 1, C = 2, B = 3, A = 4.

In some cases grades were available in all four subjects for both spring and fall. Frequently, however, grades were available for less than four subjects or grades were available for social studies and English in the spring, and English and math in the fall. A necessary, but limiting, compromise was to average the numerical values for all reported grades over these two report periods. In other words, grade point average might be based on eight marks for some students and three marks for others.

Intelligence was measured in most cases by W.I.S.C. or Chicago test scores. In all cases the tests had been administered by a diagnostician or psychologist associated with the Indiana School for the Deaf. Test scores were less than two years old in nearly all cases. The vast majority of tests had been given in 1965.

Details of Testing Procedures for the Hearing Impaired

In the first testing session 81 subjects were gathered in the recreation room of the Indiana School for the Deaf. The subjects were seated in groups of two and three at tables facing a forward podium where the main test administrator stood. Four proctors were provided who were positioned on each side of the room and moved about helping the students follow the direction and interpreting or repeating statements and directions given by the main test administrator. Two of the proctors were members of the staff of the Indiana School for the Deaf and were able to communicate with the students with signs and finger spelling when necessary. The other two proctors were not familiar with signs and finger spelling.

Testing began with a brief explanation of what was to take place and who the strangers were. Stress was placed on the idea that there were no right or wrong answers to the items and that everyone might have different ideas as to the best answer to them. These were questions about how the students saw themselves and others in their life and answers were expected to differ from person to person. Attention was directed to the appropriate page in the questionnaire by flicking the lights and then holding up the questionnaires, these having been turned to the correct page. The main test administrator then indicated the item number and signed the content of the item. The students read along with the manual presentation if they so desired.

It should be noted that these students were inexperienced when it came to mass psychological tests. Complete testing took approximately one hour.

Discussion of the first testing session led to the introduction of some important changes in the second testing session. Innovations were introduced on the basis of observed problems. This was a defensible strategy since part of the purpose of the pilot study was to determine the best ways in which to go about mass psychological testing with hearing impaired students.

The first change in the testing situation related to the behavior of the main test administrator. Although a few practice "run-throughs" had taken place prior to the first test, there were no subjects present. On the first testing session there was an element of "newness" for both the subjects and the main test administrator.

One important alteration which was suggested by the proctors from the Indiana School was the simultaneous use of "saying" and "signing" by the main test administrator. On the first session concentration had been directed toward the "signing" without much attention being paid to the simultaneous use of speech. The second testing integrated both communication channels.

The most troublesome items in the questionnaire were the "open-ended" significant-others and academic significant-others items. For some unknown reason the nature of these items forced the administrator and proctors to spend a great deal of time communicating direction to the students. Questions such as "how many names," "are three enough," "what is important," occurred frequently. It was felt that the confusion regarding these items might have been in part due to the fact that the students were unacquainted with the general testing situation and had not developed a knowledge of the respondent role.

On these grounds it was decided to reorganize the questionnaire so that the open-ended items appeared last. The general smoothness and efficiency the second testing may have been partly as a result of this change. It is possible, however, that practice in the first session was the factor which led to later testing ease with the open-ended questions.

The use of page numbers in the questionnaire had been overlooked prior to the first administration. When testing non-impaired subjects, page numbers were not necessary because the subject simply read through and answered the items independently. With the hearing impaired subjects, however, it was necessary to constantly refer to the present

location in the questionnaire. The group progressed through the questionnaire together, page by page.

Another distracting element which was eliminated during the second testing was the motion of the proctors. It became obvious during the first session that the movement of the proctors up and down the room visually distracted the subjects. Many "got lost" as a result of these movements. So for the second session, the proctors were directed to station themselves at each side of the group in such a way that eye contact could be obtained by all the subjects with one or more of the proctors. Minimum proctor movement was suggested.

Otherwise the second testing was like the first. The total testing took 35 minutes as opposed to the hour needed for the first session. It should be noted that changes in the testing situations would be expected to adversely affect estimates of stability (test-retest correlations). The tendency would be for the coefficient of stability to become a minimal estimate when important changes in the testing situation are instituted.

Scale Length

One strategy in constructing tests for cognitive processes is to obtain the most information with the least number of questions. Particularly with deviant populations we must attempt to work according to this principle. Among some groups of exceptional children we find attention span difficulties, distractibility, communication problems, and, as with all children, a great potential for just plain boredom with the testing enterprise. With deaf students, the extensive presentation of the scale through signing the content, lip-reading, and

directing attention to the written form of the questions serves to slow up the administration of the test. A somewhat delicate balance must be maintained between elaborating the content enough to insure understanding and progressing rapidly enough to hold the interest and attention of the students. Although lengthening the scale has been said to be one method of increasing its reliability, it is doubtful that doing so would be the best approach to use in instrument development with exceptional children.

Dichotomization of Responses

The assessment of a test's reproducibility or unidimensionality is usually carried out on an instrument whose items can be scored dichotomously. This necessity has arisen for two reasons: (1) a greater interest in ability testing wherein answers are usually scored on a pass-fail basis; and (2) the problem of developing means for treating multiple-response alternatives in some measures of unidimensionality is complex.

Loevinger notes: "The writer personally doubts whether multiple choice items have any advantage over dichotomous ones to offset the methodological difficulties in most contexts."[2] At any rate, of the more commonly employed methods for determining unidimensionality, Guttman's is the only one which provides means for analyzing multiple response level items. Even in the case of this method it has been noted that we will seldom find an item whose multiple response categories

---

[2]Loevinger, "The Technic of Homogeneous Tests Compared With Some Aspects of 'Scale Analysis' and Factor Analysis."

can be scored as discreet in the final stage of the analysis and still yield an adequate index of reproducibility. Instead, an attempt is made to combine response categories so as to minimize the error of reproducibility.

Note that the effect of combining response categories is to decrease the number of "scale types" scale patterns or total scores. In the case of the Self-Concept of Academic Ability Scale-Form D, dichotomization of responses results in nine scale types or total scores of 0-8. The strategy employed in combining response categories within an item is to establish a cutting point which will minimize the "error" when the reproducibility index is calculated.

CHAPTER IV

FINDINGS*

Reliability

Reliability analysis attempts to estimate the extent to which both chance error and consistency enter into measurement. The analyses presented in this chapter are addressed to: (1) the determination of the equivalence of two forms of a scale; (2) the determination of the stability of measures by test and retest with the same instrument; (3) the determination of internal consistency of the instrument and; (4) the evaluation of individual items in the scale. The instrument whose reliability was investigated is the Self-Concept of Academic Ability Scale-Form D for hearing impaired students.

The Equivalence of the SCA and the SCA-D Scales

Determination of the equivalence of two forms of a test involves examination of the distribution of scores on both instruments and correlation of the two sets of scores. Mean scores cannot be compared since the range of scores differs for each scale. The SCA scale yields a range of scores from 8-40, while the SCA-D scale yields a range of scores from 8-24. The skewness and kurtosis of both score

*All Pearson Product Moment Correlations, Means, Standard Deviation, Standard Errors of Measurement, Skewness values, and Kurtosis values are calculated by the CDC 3600 computer using a program entitled Bastat.

distributions can, however, be compared and a correlation coefficient for the two sets of scores calculated.

H_1: A high, positive correlation will be observed between the scores of the non-impaired subjects on the SCA and SCA-Form D scales.

H_R: $P > 0$

H_Δ: $p \leq 0$

Statistic: one-tailed, t test

The Pearson Product Moment Correlation between scores for the non-impaired subjects on the Self-Concept of Academic Ability and Self-Concept of Academic Ability-Form D scales was .75. This coefficient was obtained using the total (N = 97) group. A separate analysis of males (N = 49) and females (N = 48) disclosed no difference in correlations for these groups. The coefficient of equivalence was .762 for males and .756 for females. Table 2 shows, however, that high school #1 results functioned to lower the overall coefficient of equivalence.

Table 2 **Z** deviate correlations* between responses of non-impaired seniors to SCA and SCA-Form D scales (three high schools)

	High School #1 (EV) N = 22	High School #2 (S) N = 39	High School Combined #3 (E) N = 36
r	.513	.842	.805
r^2	.260	.709	.648

*The **Z** deviate correlation formula is: $XzYz = 1 - \frac{\sum n(Z_X - Z_Y)^2}{2n}$

In high school #1 a correlation of .513 appeared. Squaring this coefficient yields an index of determination of .26, meaning that 26 percent of the variance in either SCA or SCA-D scores is associated with, or predictable from measures of either variable. In high schools #2 and #3 between 65 and 71 percent of the variance of scores on either of the two scales is associated with or predictable for measures of either one. Although post facto explanations of observed phenomena have little scientific merit, it should be parenthetically noted that mass testing problems were encountered in high school #1. Only one class period was available for testing purposes and time was beginning to "run out" when the students commenced the SCA-D items. By having to hurry through the scale a larger error factor may have been introduced.

The delta hypothesis that the correlation between SCA and SCA-D scores is less than or equal to zero was rejected at the .05 level of confidence.

Table 3 Means, standard deviations, skewness, and kurtosis of the Self-Concept of Academic Ability Scale and the Self-Concept of Academic Ability Scale-Form D

Statistic	Scale	N = 97
	SCA	SCA-D
Possible Raw Score Range	8-40	8-24
Mean	27.330	18.610
S D	6.910	2.970
*S E_m	.3455	.1188
Skewness	.149	.001
Kurtosis	.080	.030

* $S E_m = S D \sqrt{1 - r_{12}}$, Where r_{12} = test-retest correlation

Another consideration in establishing the equivalence of the two forms is the shape of their respective distribution of scores. Table 3 shows the means, standard deviations, standard error of measurement, skewness and kurtosis for the SCA and SCA-D scales. Mean comparisons were not possible due to differing scoring procedures on the two instruments.

A skewness index of .149 for the SCA scale and of .001 for the SCA-D scale indicates that the score distributions were nearly symmetrical, with the SCA score showing a slight positive skewness, i.e., measures piling up slightly toward the lower values and the upper tail extended.

Kurtosis indices for the SCA and SCA-D were .08 and .03 respectively. Values of this order describe a mesokurtic distribution, i.e., the "normal" or Leplace-Gaussian probability curve. Thus, it is concluded that self-concept of academic ability, as measured by both instruments, is a normally distributed variable.

Stability of Measurement

Test-retest correlations (Pearson Product Moment) were calculated in order to determine the stability of the SCA-D responses. These test-retest correlations are likely to be minimal estimates since changes were introduced in the format and presentation of the second testing. A more specific purpose of assessing the test-retest stability is to obtain some indication as to whether or not the questions were understood by the hearing impaired subjects. It is assumed that if the items were not understood, correlations between scores obtained at separate times would be low. High correlations

would be indicative of response consistency and therefore "understanding" of the items.

H_2: A high, positive correlation will be observed between the hearing impaired subjects' SCA-D scores on test #1 and test #2.

H_R: $P > 0$

H_Δ: $p \leq 0$

Statistic: one-tailed, t test

The findings presented in Table 4 illustrate that the correlation between scores on test #1 and test #2, over a six day interval, for the hearing impaired is .84. Squaring, 71 percent of the variance of scores on test #2 is predictable from test #1.

Table 4 Test-retest correlations for SCA and SCA-Form D scales: hearing impaired and non-impaired subjects

Scale	Hearing-Impaired (N = 80)	Non-Impaired (N = 466)		
	Test-Retest Interval	Test-Retest Interval		
	6 days	7th-8th Grade	8th-9th Grade	9th-10th Grade
SCA		.65	.74	.72
SCA-D	.84			

All r's significant beyond the .05 level.

Table 4 also shows the test-retest correlations over one year periods for the non-impaired (7th - 8th grade, 8th-9th grade, 9th-10th grade). These test-retest correlations were used as stability measures

of the parent (SCA) instrument. The correlations (.65, .74, and .72) are lower than the correlation obtained for the hearing impaired. This is as expected since some change in self-concept of academic ability is likely to occur during a one year period.

The delta hypothesis that the test-retest correlation is equal to or less than zero was rejected at the .05 level of significance.

H_3: The mean SCA-D scores for test #1 will equal the mean SCA-D scores for test #2.

$$H_R: \mu 1 - \mu 2 \neq 0$$

$$H_\Delta: \mu 1 - \mu 2 \neq 0$$

Statistic: two-tailed, "t" test for related samples

A t value of 1.00 calculated for the means in Table 5 was not statistically significant at the .05 level of confidence. This led to the inference that the means on the first and second tests are equal.

Table 5 Means, standard deviations, and characteristics of distributions for the Self-Concept of Academic Ability-Form D scale: test #1 and test #2

Statistic	Test	
(N = 80)	#1	#2
Mean	18.000	18.340
S D	3.140	3.340
*S E_m	1.260	1.340
Skewness	.311	.330
Kurtosis	.250	.170
t	1.00	

*S E_m = S D $\sqrt{1 - r_{12}}$

Analysis of the distributions (Table 5) gives an impression of positive skewness of scores on both tests. Scores piled up a little below the mean. Also, there is a tending toward leptokurotsis. These skewness and kurtosis values do not represent radical departures from normalcy but give sufficient grounds to question the appropriateness of using statistics which assume a normally distributed variable.

Reproducibility of the SCA-D

Reproducibility concerns the pattern of item scores obtained on a test and the total score. A test is said to be reprodubible if knowledge of a subject's total score allows us to predict, within certain limits of error, the way the subject responded to each item in the test. Usually, a test's high reproducibility is evidence of its unidimensionality. A test which is unidimensional measures a single factor, attribute, or skill.

H_4: The reproducibility of the SCA-D scale, when responded to by the non-impaired will be greater than .50, the minimum value for considering a set of items a scale.

H_R: $I > .50$

H_Δ: $I \leq .50$

Table 6 shows that the reproducibility value which would be obtained with the present data if the items of the scale were independent (i.e., exhibited zero covariance) would be .121. This value, referred to as Rep_{ind}, in the table, was calculated on the basis of the non-impaired students' responses to the SCA-D scale. The reproducibility value, .941, greatly exceeds the minimum reproducibility for zero item covariance. The standard error of Rep is .0085, meaning that chances are 99 out of 100 that repeated sampling will show a value

Table 6 Green's Rep_{ind}, Rep, (I), and standard error of Rep for the SCA-D Scale: hearing impaired and non-impaired subjects

Population or Sample	Green's Rep_{ind}	Green's Rep	σ Rep***	Green's (I)	N
Non-impaired high school seniors	.121	.941	.0085	.932	97
Hearing impaired test #2	.154	.915	.011	.900	85

*** $\sigma_{Rep} = \sqrt{\frac{(1-Rep)(Rep)}{NK}}$

for the population of between .9283 and .9537.

The (I) value, whose formula takes into account both the obtained reproducibility coefficient (Rep) and the reproducibility coefficient which would be obtained if the items were independent (Rep_{ind}), was .932. This value is greater than .50 and therefore permits rejection of the delta hypothesis.

> H_5: The reproducibility of the SCA-D scale, when responded to by the hearing impaired will be greater than .50, the minimal value for considering a set of items a scale.
>
> H_R: $I > .50$
>
> H_Δ: $I \leq .50$

The results with the hearing impaired contained in Table 6 parallel the results obtained with the non-impaired. Rep_{ind} equalled .154 with a Rep of .915. Both of these values are slightly lower than those obtained with the non-impaired. The standard error of Rep was .011, meaning that chances are 99 out of 100 that repeated sampling will yield values of from .883 to .917. The (I), or summary statistic,

showed a value of .900. This value is substantially greater than .500 and permits rejection of the delta hypothesis.

Item Evaluation

One rational principle for item evaluation is that each item should be homogeneous with the total test. A condition of item to test homogeneity exists when subjects who pass a given item have higher total scores than those who fail it. Complete absence of homogeneity of items to test has been called "heterogeneity" and occurs when the total scores of subjects passing an item are randomly distributed along the ordered (high to low) total scores. Two methods of item evaluation were used: Loevinger's H_{it} and the Phi (ϕ_{it}) coefficient.

H_6: Each item (8) in the SCA-D scale, when administered to the non-impaired, will be homogeneous with the total scale.

H_R: $\phi_{it} > 0$ (for each item)

H : $\phi_{it} = 0$ (for each item)

Statistic: **Z** deviate

Table 7 shows the Phi (ϕ_{it}) coefficients which were obtained in the item analysis. **Z** values were calculated for each ϕ_{it} and ranged from 5.59 to 7.45. Inasmuch as **Z** deviate values of this magnitude are quite unlikely ($p < .001$) to result from random variation from a ϕ_{it} of zero, the delta hypothesis is rejected for each item.

Table 7 Phi (ϕ_{it}) coefficients and **Z** deviates for each item in the Self-Concept of Academic Ability Scale-Form D: hearing impaired and non-impaired subjects

Subjects	Scale	ϕ_{it} Item Coefficient 1	2	3	4	5	6	7	8
Non-Impaired	SCA								
High School Seniors	Form D	.599	.568	.765	.759	.756	.617	.556	.703
N = 97	**Z**	5.90	5.59	7.54	7.48	7.45	6.08	5.48	6.93
Hearing Impaired	SCA	.573	.639	.570	.558	.523	.773	.436	.401
Test #2	Form D								
N = 80	**Z**	5.13	5.72	5.10	4.99	4.77	6.92	3.90	3.59

$$Z = \phi_{it} \sqrt{n}$$

H_7: Each item (8) in the SCA-D scale, when administered to the hearing impaired, will be homogeneous with the total scale.

$$H_R: \phi_{it} > 0$$

$$H_\Delta: \phi_{it} = 0$$

Statistic: **Z** deviate

The phi coefficients (ϕ_{it}) for the SCA-D item responses of the hearing impaired resulted in **Z** deviate values ranging from 3.59 to 6.92. The probability of obtaining values of this magnitude through random vairation from a ϕ_{it} of zero is less than .001. Therefore, the delta hypothesis is rejected for each item.

A supplementary item analysis using Loevinger's (H_{it}) index is summarized in Table 8. The expected value of the statistic, given

Table 8 Loevinger's (H_{it}) Values for each item in the Self-Concept of Academic Ability-Form D Scale: hearing impaired and non-impaired subjects

Population or Sample	Scale	(H_{it}) Item Value 1	2	3	4	5	6	7	8
Non-Impaired High School Seniors	SCA Form D	.987	.981	.997	.984	.988	.940	.989	.970
Hearing Impaired Students Test #2	SCA Form D	.963	.978	.977	.984	.989	.985	.977	.934

complete item to test heterogeneity is zero. Sampling properties of the coefficient are unknown and formal statistical hypothesis testing is impossible. In every instance the H_{it} coefficients for both non-impaired and hearing impaired were greater than .93. Since the expected value of H_{it} is zero given complete item to test heterogeneity and 1.00 given item to test homogeneity, it seems reasonable to assert that each item in the SCA-D scale is homogeneous with the total test for both sets of responses.

Construct Validity

Construct validity, as one type among several, is primarily addressed to the theory in which a construct is embedded. Rather than assessing how well an instrument predicts some specific outcome, the establishment of construct validity involves the empirical testing of hypotheses involving the construct. These hypotheses are derived from the theory which posits the construct. Positive, zero, and

negative relationships should be hypothesized in a complete construct validity analysis. It should be remembered that correlation analysis is but one of five means for assessing construct validity. Some other procedures were described in Chapter II.

Construct Validity Hypotheses

H_8: The Self-Concept of Academic Ability of the hearing-impaired students will be associated with their school performance (past GPA)

H_R: $P > 0$

H_Δ: $p \leq 0$

Statistic: "Student's "t," one-tailed

The correlation between self-concept of academic ability and past school performance was .32 for the hearing impaired. A "t" value of 2.97 (78 d.f.) was computed for this correlation. The probability of a t value of that magnitude is $<.05$. These findings permit rejection of the delta hypothesis and acceptance of the research hypothesis. The Self-Concept of Academic Ability of the hearing impaired is associated with past school performance.

H_9: The Self-Concept of Academic Ability of the hearing impaired will be associated with perceived parental evaluations of academic ability.

H_R: $P > 0$

H_Δ: $p \leq 0$

Statistic: "Student's t," one-tailed

A correlation coefficient of .50 was obtained for the SCA-D - PPEv comparison. This coefficient produces a "t" value of 5.16 (78 d.f.) which is statistically significant at less than the .05 level. The delta hypothesis that the correlation between self-concept

of academic ability is associated with perceived parental evaluation among hearing impaired students.

H_{10}: The Self-Concept of Academic Ability of the hearing impaired will be associated with perceived teacher evaluations of academic ability.

H_R: $P > 0$

H_Δ: $P \leq 0$

Statistic: "Student's t," one-tailed

Perceived teacher evaluations was correlated with Self-Concept of Academic Ability. An observed correlation of .48 yielded a "t" value of 4.89. The probability of obtaining such a value by chance is .05. The delta hypothesis is rejected. Hypothesis 10, that the self-concept of academic ability of the hearing impaired will be associated with perceived teacher evaluation, is accepted.

H_{11}: The Self-Concept of Academic Ability of the hearing impaired will be associated with measured intelligence.

H_R: $P > 0$

H_Δ: $P \leq 0$

Statistic: "Student's t," one-tailed

A correlation of .29 was found between self-concept of academic ability and IQ for this population. A "t" of 2.68 calculated on the basis of this coefficient is statistically significant at the .05 level. The delta hypothesis is thereby rejected and hypothesis 11, that the Self-Concept of Academic Ability of the hearing impaired will be associated with intelligence, is accepted.

H_{12}: The Self-Concept of Academic Ability of hearing impaired students is <u>not</u> associated with grade level.

H_R: $P = 0$

H_Δ: $P \neq 0$

Statistic: "Student's t," two-tailed

The correlation between Self-Concept of Academic Ability and grade level was -.04. On the basis of a "t" value of .035 calculated with the correlation coefficient, the delta hypothesis was rejected. The Self-Concept of Academic Ability of the hearing impaired students was not associated with grade level.

H_{13}: The mean Self-Concept of Academic Ability of the hearing impaired equals the mean Self-Concept of Academic Ability of the non-impaired.

$$H_R: \mu_1 - \mu_2 = 0$$

$$H_\Delta: \mu_1 - \mu_2 \neq 0$$

Statistic: t test for independent samples, two tailed

The SCA-D mean for the hearing impaired (test #2) was 18.34 as compared to an SCA-D mean for the non-impaired of 18.61. Before a "t" test of differences between means can be used it is necessary to determine whether or not the parent populations can be considered to have equal variances. An F test for homogeneity of variances was undertaken, the null hypothesis being:

$$\frac{\sigma_1^2}{\sigma_2^2} = 1.0$$

An F value of 1.37 was obtained. This is not statistically significant at the .05 level, the null hypothesis being therefore accepted. SCA-D responses for the hearing impaired and non-impaired are homogeneous in variance.

A "t" value of 1.02 was calculated for the two means. This was not statistically significant, leading to rejection of the delta hypothesis. The mean Self-Concept of Academic Ability of the hearing impaired is equal to the mean Self-Concept of Academic Ability of the non-impaired.

Other Findings

Tables 9 and 10 present the responses to the items in the SCA-D scale which characterize each scale type (total score). Table 9 represents the responses of the non-impaired and Table 10 the hearing impaired. Since the items of the scale are ordered according to difficulty for both groups a comparison can be drawn.

The item difficulty order is exactly the same for both groups of subjects except for item #4. Only those hearing impaired students who score the maximum of 8 believe that they can graduate from college. Table 11 shows that this is only four percent of the hearing impaired subjects. For the non-impaired, the item concerning graduating from college is fifth in difficulty. As is illustrated in Table 11, 59 percent of the non-impaired believe that they can graduate from college.

Table 9 Responses of the non-impaired high school seniors to the Self-Concept of Ability items according to scale types*

Item #	Essence of Item	Scale Type Response Patterns 8	7	6	5	4	3	2	1	0
7	Disregarding teacher's evaluation, thinks own work is	excellent	average/ below average	→	→	→	→	→	→	→
3	On graduation from h.s., thinks he will be with	the best students	the best students	average/ below av.	→	→	→	→	→	→
2	In reference to other students in class, thinks his work is	better	better	better	the same poorer	→	→	→	→	→
1	In reference to friends, thinks he can do school work	better	better	better	better	the same poorer	→	→	→	→
4	Thinks he can graduate from college	yes	yes	yes	yes	yes	maybe/no	→	→	→
8	If he really tries, he can get	A's & B's	A's & B's	A's & B's	A's & B's	A's & B's	A's & B's	B's & C's D's & E's	→	→
6	Thinks he can go to grad. school to become a teacher or a doctor	yes/maybe the best/ average students	yes/maybe the best/ average students	yes/maybe the best/ average students	yes/maybe the best/ average students	yes/maybe the best/ average students	yes/maybe the best/ average students	yes/maybe the best/ average students	no	→
5	If he went to college, he thinks he would be one of								the best/ average students	poor

*Read down the column for the typical responses of a subject who obtains any one of the nine (0-8) scale scores. Some errors occur and therefore not all scale types of 6, for example, showed the above pattern.

Table 10 Responses of the hearing impaired students to the Self-Concept of Academic Ability items according to scale types* Test #2

Item #	Essence of Item	Scale Type Response Patterns 8	7	6	5	4	3	2	1	0
4	Thinks he can graduate from college	yes	maybe/no	→	→	→	→	→	→	→
7	Disregarding teacher's evaluation, thinks own work is	excellent	excellent	average/ below average	→	→	→	→	→	→
3	On graduation from h.s., thinks he will be with	the best students	the best students	the best students	average/ below average	→	→	→	→	→
2	In reference to other students in class, thinks his work is	better	better	better	better	the same/ poorer	→	→	→	→
1	In reference to friends, thinks he can do school work	better	better	better	better	better	the same/ poorer	→	→	→
8	If he really tries, he can get	A's & B's	A's & B's	A's & B's	A's & B's	A's & B's	A's & B's	B's & C's D's & E's	→	→
6	Thinks he can go to grad. school to become a teacher or doctor	yes/maybe	yes/maybe	yes/maybe	yes/maybe	yes/maybe	yes/maybe	yes/maybe	no	→
5	If he went to college he thinks he would be one of	the best/ average students	the best/ average students	the best/ average students	the best/ average students	the best/ average students	the best/ average students	the best/ average students	the best/ average students	poor

*Read down the column for the typical responses of a subject who obtains any one of the nine (0-8) scale scores. Some errors occur and therefore not all scale types of 6, for example, showed the above pattern.

Table 11 Percentage distribution of scale types: hearing impaired and non-impaired subjects

Population/Sample	Scale Type								
	8	7	6	5	4	3	2	1	0
Non-Impaired High School Seniors N = 97	.05	.07	.14	.14	.19	.14	.12	.08	.05
Hearing Impaired Students (Test #2) N = 85	.04	.06	.15	.09	.16	.21	.18	.07	.04

The correlations between Self-Concept of Academic Ability are four other variables (previous GPA, Perceived Parental Evaluation, Perceived Teacher Evaluations, and Perceived Peer Evaluations) were the same for the hearing impaired and non-impaired. These results are shown in Table 12. The test of significance for differences between r's utilized Fisher's **Z** transformation. The null hypothesis of zero difference between correlations was accepted in every case.

Table 12 Pearson Product Moment Correlations between Self-Concept of Academic Ability* and major variables: hearing impaired and non-impaired subjects

Sample/Population	Previous GPA	Perceived Parental Evaluations	Perceived Teacher Evaluations	Perceived Peer Evaluations
Hearing Impaired (Test #2) N = 80	.32	.50	.48	.53
	↑ NS ↓	↑ NS ↓	↑ NS ↓	↑ NS ↓
Non-Impaired N = 73	.39	.43	.40	.56

NS = Not a statistically significant difference between correlations, **Z** for independent samples.

*Raw scores - SCA-Form D

CHAPTER V

SUMMARY

This has been a pilot investigation of the reliability and construct validity of a scale designed to measure the self-concept of academic ability of the hearing impaired. The Self-Concept of Academic Ability Scale-Form D (SCA-D) is a modification of an instrument used extensively in social-psychological research with non-impaired students. The modified scale is simpler in vocabulary and can be easily administered through the use of signs and finger-spelling.

Summary of Research Objectives

The research objectives involved assesing: (1) the equivalence of the two forms of the scale, (SCA and SCA-D); (2) the stability of measurement of the SCA-D scale; (3) the reproducibility of the SCA-D scale; (4) the item to test homogeneity of each item in the scale; and (5) the construct validity of self-concept of academic ability for hearing impaired students.

Summary of Methodology

The subjects of the study were: (1) a random sample of twelfth grade students in a regular public high school; and (2) all of the "academic" students in the Indiana School for the Deaf grades 8-11). All subjects were Caucasian except for three of the hearing impaired students. There were 97 non-impaired subjects, (50 male and 47 female) and 80 hearing impaired subjects, (50 male and 30 female).

To determine the equivalence of the two forms of the scale, both versions were administered, en masse, to the twelfth grades. Stability of measurement for the hearing impaired was determined by analyzing data obtained in test and retest separated by six days. Reproducibility and item analyses were conducted using data from the second testing of the hearing impaired. Comparable reproducibility and item analyses were conducted utilizing the twelfth grade subjects' responses.

To assess construct validity, four major variables besides self-concept of academic ability were examined: (1) grade point average; (2) IQ scores; (3) Perceived Parental Evaluation; and (4) Perceived Teacher Evaluations.

The data were collected through the administration of questionnaires and from the students' school records. Analysis involved the use of Z tests, t test, F tests, skewness measures, kurtosis measures, Pearson Product Moment Correlation, Phi (ϕ_{it}) Coefficients, and Green's (I), Rep. and Repind. All correlations, means, standard deviations, and distribution summaries were calculated by the CDC 3600 at the Michigan State University Computer Center.

Summary of Research Findings

Reliability

a. Coefficients of equivalence ranged from .842 to .513 for the SCA-D scale in three schools. Coefficients of equivalence were not affected by the sex of the respondent, but were low in one of the three high schools. The overall coefficient of equivalence was .75 (Table 2).

b. Responses to both the SCA and SCA-D instruments form a normal distribution (i.e. mesokurtic with zero skewness) (Table 3).

c. The test-retest correlation for the hearing impaired on the SCA-D was .84 over a six-day interval. Test-retest correlations over a one year interval with non-impaired respondents were lower, .65 to .72. This is as expected (Table 4).

d. The mean SCA-D scores on test and retest were equal (Table 5).

e. The reproducibility of the SCA-D as estimated by Green's (Rep) was .915 for the hearing impaired and .941 for the non-impaired. The respective σ Rep's were .011 and .008. Green's summary statistic (I) was calculated and a value of .90 for the hearing impaired and .93 for the non-impaired obtained. These figures exceed the suggested minimum of .50 for a reproducible scale (Table 6).

f. Each item of the scale was found to be homogeneous with the total test for both the hearing impaired and the non-impaired. Phi (ϕ_{it}) coefficient yielded Z deviates ranging from 5.48 to 7.54 for the non-impaired and from 3.59 to 6.92 for the hearing impaired. All were statistically significant at the .05 level or better (Table 7).

A general conclusion is that all findings support the contention that the Self-Concept of Academic Ability Scale-Form D is a reasonable reliable instrument for research with the hearing impaired. Despite the fact suggestions will be given for its improvement in a later section, interpretable results can be obtained using the instrument as its stands. Some basis comparisons can be made between the Self-Concepts of Academic Ability of hearing impaired and non-impaired students using this instrument.

Construct Validity

a. A correlation of .32, significant at the .05 level, confirmed the hypothesis that Self-Concept of Academic Ability is associated with school performance (past GPA) among the hearing impaired.

b. Perceived Parental Evaluations of Academic Ability were found to be associated with Self-Concept of Academic Ability (R=.50) for this population.

c. Perceived Teacher Evaluations of Academic Ability were associated with the hearing impaired students' Self-Concept of Academic Ability (R=.48).

d. Measured intelligence was associated with (R=.29) Self-Concept of Academic Ability.

e. A zero association hypothesis, that Self-Concept of Academic Ability is not associated with grade level, was accepted.

f. Equality of mean Self-Concepts of Academic Ability for the hearing impaired and non-impaired was hypothesized and confirmed.

Other Findings

a. An analysis of the respective items difficulties orders disclosed that the hearing impaired respondents differed from the non-impaired in only one respect. Whether or not a hearing impaired student defined himself as being able to graduate from college determined whether he had a perfect or less than perfect score. Response to this item made the difference between a score of 8 and 7. Whereas for the non-impaired, the item on college graduation was 5th of 8 in difficulty.

b. Another finding of interest was that the correlations between Self-Concept of Academic Ability and other variables (previous GPA, Perceived Parental Evaluations of Academic Ability, Perceived Peer Evaluations of Academic Ability, and Perceived Teacher Evaluations of Academic Ability) were the same for both the hearing impaired and the non-impaired.

Accepting the construct validity hypotheses leads to the conclusion that the same central variable relationships exist for the hearing impaired as exist for the non-impaired. Since these relationships have been used as an argument for the construct validity of Self-Concept of Academic Ability with non-impaired individuals, it is legitimate to conclude the same for the impaired. In essence, the parallel construct validity findings mean that a general social-psychology of learning can be equally useful in understanding the learning of the deviant as it has been with "normal" populations. Rather than being seen as totally different, the hearing impaired may be profitably studied in the same framework as the non-impaired. No evidence was found in this research which would force professionals to seek a separate "social- psychology of learning for the hearing impaired."

Limitations of the Study and Implications for Future Research

It should be remembered that the present research was a pilot study. The major purposes being to evaluate instruments and testing procedures. Few systematic attempts have been heretofore made to develop instruments which will yield data comparable to that obtained with impaired subjects. And little attention has been given to questionnaire administration with hearing impaired students. Therefore, the limitations of the research are inseparably connected with the implication for future research. In fact, the discovery of difficulties and deficiencies in instrumentation and testing procedures is the main objective of a pilot study. It is not designed to be the final word regarding the instrument. An improved study has already been proposed which takes into account the limitations of this one.

The first limitation of the present research is the hearing impaired sample. While the non-impaired group reported here was a random sample of all twelfth graders in an urban school system, the hearing impaired student sample was of known bias. These students were the very best in one residential school for the deaf. A better sample would include all levels of academic performance and several residential schools. Probably hearing impaired students in public school programs should also be studied. Of course it would be necessary to initially analyze the two groups of students separately. Study of the "non-academic" students might show problems in instrumentation and procedures which are not yet apparent. Perhaps it is impossible to test, <u>en masse</u>, hearing impaired students who show very low academic performance. An attempt should be made to determine whether or not reliable and valid responses can be

obtained with the lower achievers.

A footnote in the text alluded to the possibility of "response set" spuriously increasing the test-retest correlations. No control was provided for this eventuality in the present research. However, by designing items which are so ambiguous or difficult that understanding is precluded, an index of minimal test-retest correlation could be obtained. Although the expected (r) would be zero, response set might easily increase the expected correlation to some unknown values.

Not only is it possible that the test-retest correlations are spuriously high due to an unknown "response-set" factor, the reproducibility estimate may be spuriously high due to the dichotomization of a three point response scale. Minimization of error is the rule which is followed when combining categories on an item. No explicit of rigorous decision rule has been developed. On these grounds it could be argued that the reproducibility of the scale is still questionable. Reproducibility obtained when a two choice item is used should be compared with reproducibility obtained when a three or five choice response is dichotomized.

Because of the problem of treating item sampling error as systematic variance when the same form of a test is administered twice, an alternate form should be designed to be used in checking test-retest reliability. Also, the procedural innovations which were introduced in the second testing distinctly limited the usefulness of test-retest data in the present study. A better approach would have been to treat the first test as a teaching session in which students learned to respond. Then, a second and third testing with alternate forms should be conducted and the data from these used in the analysis.

Since a major variable in the theoretical network is future school achievement, the self-concept of ability-achievement relationship is still untested. Time and financial limitation made it impossible to obtain school grades subsequent to the determination of students' self-concepts of ability. Were it possible to obtain these grades, three important things could have been accomplished: (1) the correlation between self-concept of academic ability and future GPA could have been determined, (2) intelligence and other variables could have been controled in the SCA_GPA relationship, and (3) a multiple prediction formula using GPA and SCA could have been developed for cross-validation on another sample. The latter would have been one approach to establishing the predictive validity of the scale.

At this point, certain suggestions can also be made regarding the use of factor analytic techniques. A worthwhile research activity would be to do a factor analysis of all items used in the entire questionnaire. Self-concept of academic ability, perceived evaluations, expectation of others, etc., should form separate factors if they are unique. Evidence of the empirical validity of self-concept of academic ability would be gained through the use of factor analysis.

The concurrent validity of the self-concept of academic ability scale also merits attention. The best strategy for going about this would be to conduct extensive interviews with hearing impaired students. Specifically, the interviewer should attempt to bring forth self-statements of academic ability with some probing being made in the direction of self-other comparisons. Recorded protocals should be evaluated and scored, then compared

with later information on the same subjects obtained through use of the self-concept of academic ability scale.

Although the self-concept of academic ability has definitely not been conceived of as a "diagnostic" instrument, and incidentally should not, another strategy for elaborating the general theory with the hearing impaired might take the following form. A subject who obtains a particular scale type score could be subjected to intensive evaluative increase on those facets of self-concepts of academic ability on which he scores low. Rather than a non-focused or general enhancement of evaluation, an experimenter could work in the areas of observed, low, self-definition. By so doing, it would be anticipated that both self-concept of academic ability and school grades would be more likely to increase than when non-specific methods are used.

Finally, standardization of the testing procedure should be developed. At the moment it seems that a filmed "signing and saying" of the total questionnaire should be developed. A filmed presentation, along with an improved test booklet, would permit an extension of the research into a nationwide sampling of students in schools for the deaf. A film, booklet, and standard instructional format would enable the testing to be carried out by untrained personnel in many schools. Costs would thereby become minimal and good data could be obtained.

BIBLIOGRAPHY

American Educational Research Association, Committee on Test Standards, and National Councils on Measurement Used in Education, Committee on Test Standards. "Technical Recommendations for Psychological Tests and Diagnostic Techniques," Psychological Bulletin, LI (1954), Supplement, 201-238.

Blanton, Richard L. and Nunnally, Jum C. "Educational Language Processes in the Deaf," Psychological Reports, XV (1964), 891-894.

Brookover, Wilbur B. "A Social Psychological Conception of Classroom Learning," School and Society, LXXVII (1959), 84-87.

_____, and Gottlieb, David. Sociology of Education. New York: American Book Company, 1964.

_____, LePere, Jean M., Erickson, Edsel L., Thomas, Shailer. "Definitions of Others, Self-Concept, and Academic Achievement: A Longitudinal Study," a paper presented at the American Sociological Association at Chicago, Ill., 1965.

_____, LePere, Jean M., Hamachek, Don E., Thomas, Shailer, and Erickson, Edsel. "Improving Academic Achievement through Students' Self-Concept Enhancement, II," U.S. Office of Education, Cooperative Research Project No. 2831. East Lansing: Bureau of Educational Research Services, Michigan State University, 1965.

_____, Paterson, Ann, and Thomas, Shailer. "Self-Concept of Ability and School Achievement," U.S. Office of Education Cooperative Research Project No. 845. East Lansing: Office of Research and Publications, Michigan State University, 1962.

Cronbach, Lee. Essentials of Psychological Testing. New York: Harper and Bros., 1960.

_____, and Meehl, P. "Construct Validity in Psychological Tests," Psychological Bulletin, LII (1955), 281-302.

Cruickshank, William (ed.). Psychology of Exceptional Children and Youth. New York: Prentice-Hall, 1963.

Cureton, Edward, "Quantitative Psychology as a Rational Science," Psychometrika, XI (1946), 191-196.

Cutler, Albert. "Place of Self Concept in the Education of the Physically Different Child," Exceptional Children, XXVIII (1962) 343-349.

Dexter, Lewis A. "A Social Theory of Mental Deficiency," American Journal of Mental Deficiency, LXII (1958), 920-928.

Dornbusch, Sanford H. and Schmid, Calvin, F. A Primer of Social Statistics. New York: McGraw-Hill, 1955.

Erickson, Edsel, Brookover, Wilbur, Joiner, Lee, and Towne, Richard. "A Social-Psychological Study of the Educable Mentally Retarded: An Educational Research Application of Symbolic Interactionism," a paper presented at the Council for Exceptional Children, Portland, Oregon, 1965.

_____, and Thomas, Shailer. "The Normative Influence of Parents and Friends Upon School Achievement," a paper presented at the meetings of the American Educational Research Association, Chicago, Ill., 1965.

Frankell, Edward. "Effects of a Program of Advanced Summer Study on the Self Perception of Academically Talented High School Students," Exceptional Children, XXX (1964), 245-249.

Geer, William and Deno, Evelyn. "CEC and Legislation - Now and in the Future," Exceptional Children, XXXII (Nov., 1965), 187-194.

Goffman, Irving. Stigma: Notes on the Management of Spoiled Identity. Englewood Cliffs: Prentice Hall, 1963.

Guttman, Louis. "The Cornell Technique for Scale and Intensity Analysis," Educational and Psychological Measurement, LIV (March, 1957), 81-99.

Haarer, David L. A Comparative Study of the Self-Concept of Ability Between Institutionalized Delinquent Boys and Non-Delinquent Boys Enrolled in Public Schools. unpublished Ph.D. Thesis, East Lansing: Michigan State University, 1964.

Helmstadter, G. C. Principles of Psychological Measurement. New York: Appleton-Century-Crofts, Inc., 1964.

Hollingshead, Merrill T. "The Social Psychology of Exceptional Children: Part I in Terms of the Characteristics of Exceptional Children," Exceptional Children, XXVI (1959), 137-140.

Hoyt, C. "Test Reliability Obtained by Analysis of Variance," Psychometrika, VI (1941), 153-160.

Jackson, Robert and Ferguson, George. "Studies on the Reliability of Tests," Bulletin No. 12, Department of Educational Research, University of Toronto, 1941.

Kelman, Herbert C. "Process of Opinion Change," Public Opinion Quarterly, XXV (Spring, 1961), 57-78.

Kerlinger, Fred. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc., 1964.

Kuder, G. F. and Richardson, M.W. "The Theory of Estimation of Test Reliability," Psychometrika, II (1937), 151-160.

LePere, Jean M. "A Longitudinal Analysis of the Relationship of Self-Concept and Academic Achievement of Students from Seventh through Tenth Grades," a paper presented at the meetings of the American Educational Research Association, Chicago, Ill., 1965.

_____, "A Study of the Impact of the Parent on the Child's Concept of Self as it Relates to Academic Achievement," a paper presented at the meetings of the American Educational Research Association, Chicago, Ill., 1964.

Lindquist, E. F. Educational Measurement. Washington: American Council on Education, 1961.

Linton, Ralph. The Study of Man. New York: Appleton-Century-Crofts, Inc., 1936.

Loevinger, Jane. "A Systematic Approach to the Construction and Evaluation of Tests of Ability," Psychological Monographs, LXI (1947), No. 4.

_____, "The Technic of Homogeneous Tests Compared With Some Aspects of 'Scale Analysis' and Factor Analysis." Psychological Bulletin, XLV (1948), 507-529.

McQuitty, Louis "Elementary Linkage Analysis for Isolating Orthogonal and Oblique Types and Typal Relevances," Educational and Psychological Measurement, XVII (Summer, 1957), 207-229.

Mead, George Herbert. Mind, Self and Society. Chicago: University of Chicago Press, 1934

Merton, Robert K. Social Theory and Social Structure. Glencoe, Ill.: The Free Press, 1957.

Meyerson, Lee. "Physical Disability as a Social Psychological Problem," Journal of Social Issues, IV (1948), 1-10.

Morse, Richard J. Self-Concept of Ability, Significant Others and School Achievement of Eighth-Grade Students: A Comparative Investigation of Negro and Caucasian Students, unpublished M.A. Thesis, Michigan State University, 1963.

Paterson, Ann. Evaluation of an Instrument Purporting to Measure the Construct Self-Concept of (Academic) Ability. unpublished Ph.D. Thesis, East Lansing: Michigan State University, 1966.

Reynolds, Maynard C. "The Social Psychology of Exceptional Children: Part III in Terms of the Interaction of Exceptional Children with Other Persons," Exceptional Children, XXVI (1959), 243-247.

Rogers, Merril and Quigley, Stephen. "Research Needs in the Vocational Rehabilitation of the Deaf," American Annals of the Deaf, CV (Sept., 1960), 335-370.

Rosenstein, Joseph and Lerman, Alan. Vocational Status and Admustment of Deaf Women. New York: Lexington School for the Deaf, 1963.

Stevens, Harvey and Heber, Rick (eds.). Mental Retardation: A Review of Research. Chicago: University of Chicago Press, 1964.

Thorndike, Robert and Hagen, Elizabeth. Measurement and Evaluation in Psychology and Education. New York: John Wiley and Sons, Inc. 1961.

Towne, Richard C. and Joiner, Lee M. "The Effects of Special Class Placement on the Self-Concept of Ability of the Educable Mentally Retarded," U.S. Office of Education, Cooperative Research Project No. 71-2073. East Lansing: Bureau of Educational Research Services, Michigan State University, to be published September, 1966.

Trippe, Matthew J. "The Social Psychology of Exceptional Children: Part II in Terms of Factors in Society," Exceptional Children, XXVI (1959), 171-175.

Washburn, William C. "Patterns of Self Concepts Related to Problems of Neuroses and Delinquency in Adolescents," Exceptional Children, XIX, (1963), 341-347.

White, Benjamin and Saltz, Eli. "Measurement of Reproducibility," Psychological Bulletin, LIV (March, 1957), 81-99.

Wright, Beatrice. Physical Disability - A Psychological Approach. New York: Harper and Row, 1960.

APPENDIX A

COMPLETE QUESTIONNAIRE ADMINISTERED AT INDIANA SCHOOL FOR THE DEAF

DS/65

WESTERN MICHIGAN UNIVERSITY

High School Study

INTRODUCTION: (to be read aloud)

Some people at Western Michigan University are trying to find out what students think about some questions. This is not a test. Just tell me what you think or feel. Everyone has different feelings. This is O.K. If you do not understand, ask me to explain. No one will see your answers but me.

Answer these questions. Do not turn the page until I tell you.

PLEASE PRINT

Name: _____ _____ _____
(Last name) (First name) (Middle name)

Birthdate: _____ _____ _____ Sex: M____ F____

Name of present school: _____

What school did you attend last year? _____

STOP

DO NOT TURN THE PAGE YET

DS/65

There are many people who are important in our lives. List the names of people who you feel are important in YOUR life. Tell me who each person is.

NAMES	WHO IS THIS PERSON?

STOP

DO NOT TURN THE PAGE YET

There are many people who wonder about children's school work. List the NAMES of the people you feel are wondering about your school work. Please tell me who each person is.

NAMES	WHO IS THIS PERSON?

STOP

DO NOT TURN THE PAGE YET

Pick one. Circle the letter for your answer.

1. Think of your friends. Do you think you can do school work better, the same, or poorer than your friends?

   a. better
   b. the same
   c. poorer

2. Think of the students in your class. Do you think you can do school work better, the same, or poorer than the students in your class?

   a. better
   b. the same
   c. poorer

3. When you graduate from high school, do you think you will be with the best students, average students, or below average students?

   a. the best
   b. average
   c. below average

4. Do you think you could graduate from college?

   a. yes
   b. maybe
   c. no

5. If you went to college, do you think you would be one of the best, average, or poorest students?

   a. the best
   b. average
   c. poorest

6. If you want to be a doctor or a teacher, you need more than 4 years of college. Do you think you could do that?

   a. yes
   b. maybe
   c. no

7. Forget how your teachers mark your work. How good do <u>you</u> think your own work is?

   a. excellent
   b. average
   c. below average

8. What marks do you think you <u>really</u> <u>can</u> get if you try?

a. A's and B's
b. B's and C's
c. D's and E's

STOP

DO NOT TURN THE PAGE YET

Pretend you are your mother or father. Answer like they would. Pick one. Circle their answer.

1. Think of your mother and father. Do your mother and father say you can do school work better, the same, or poorer than your friends?

   a. better
   b. the same
   c. poorer

2. Would your mother and father say you would be with the best, average, or below average students when you graduate from high school?

   a. the best
   b. average
   c. below average

3. Do they think you could graduate from college?

   a. yes
   b. maybe
   c. no

4. Remember, you need more than four years of college to be a teacher or doctor. Do your mother and father think you could do that?

   a. yes
   b. maybe
   c. no

5. What grades do your mother and father think you <u>can</u> get?

   a. A's and B's
   b. B's and C's
   c. D's and E's

STOP

DO NOT TURN THE PAGE YET

Pretend you are your best friend. Answer like he or she would. Pick one. Circle their answer.

1. Think of your best friend. Would your best friend say you can do school work better, the same, or poorer than other people your age?

   a. better
   b. the same
   c. poorer

2. Would your best friend say you would be with the best, average, or below average students when you graduate from high school?

   a. the best
   b. average
   c. below average

3. Does your best friend think you could graduate from college?

   a. yes
   b. maybe
   c. no

4. Remember you need more than four years of college to be a teacher or doctor. Does your best friend think you could do that?

   a. yes
   b. maybe
   c. no

5. What grades does your best friend think you can get?

   a. A's and B's
   b. B's and C's
   c. D's and E's

STOP

DO NOT TURN THE PAGE YET

Pretend you are your teacher, the one you like the best. Answer like he or she would. Pick one. Circle their answer.

1. Think of your teacher. Would your teacher say you can do school work better, the same, or poorer than other people your age?

   a. better
   b. the same
   c. poorer

2. Would your teacher say you would be with the best, average, or below average students when you graduate from high school?

   a. the best
   b. average
   c. below average

3. Does your teacher think you could graduate from college?

   a. yes
   b. maybe
   c. no

4. Remember you need more than four years of college to be a teacher or doctor. Does your teacher think you could do that?

   a. yes
   b. maybe
   c. no

5. What grades does your teacher think you can get?

   a. A's and B's
   b. B's and C's
   c. D's and E's

STOP

DO NOT TURN THE PAGE YET

Pick one. Circle the letter for your answer.

3. How far would you like to go in school?

a. quit now
b. go to high school for a while
c. graduate from high school
d. go to a school to be a secretary or learn a trade
e. go to college for a little while
f. graduate from college
g. more than 4 years of college

4. Sometimes what you expect to do isn't the same as what you'd like to do. How far in school will you really go?

a. quit now
b. go to high school for a while
c. graduate from high school
d. go to a school to be a secretary or learn a trade
e. go to college for a little while
f. graduate from college
g. more than 4 years of college

STOP

DO NOT TURN THE PAGE YET

Pick one. Circle the letter for your answer.

1. Which marks would you have to get to make your mother and father happy?

   a. A's
   b. B's
   c. C's
   d. D's
   e. E's
   f. my mother and father don't care what marks I get in school

2. Do your parents know how you are doing in school? Pick one.

   a. They know everything I do in school.
   b. They know almost everything about my school work.
   c. They know some things about my school work.
   d. They only know a little bit about my school work.
   e. They know nothing about my school work.

3. What would be the lowest marks you could get in school and still satisfy your mother and father?

   a. A's
   b. B's
   c. C's
   d. D's
   e. E's
   f. My mother and father don't care what marks I get in school.

4. How far in school do your mother and father think you will go?

   a. quit now
   b. go to high school for a while
   c. graduate from high school
   d. go to a school to be a secretary or learn a trade
   e. go to college for a little while
   f. graduate from college
   g. more than 4 years of college

STOP

DO NOT TURN THE PAGE YET

Pick one. Circle the letter for your answer.

1. Which marks would you have to get to make your best friend happy?

   a. A's
   b. B's
   c. C's
   d. D's
   e. E's
   f. My best friend doesn't care what marks I get in school

2. Does your best friend know how you are doing in school? Pick one.

   a. He knows everything I do in school.
   b. He knows almost everything about my school work.
   c. He knows some things about my school work.
   d. He only knows a little bit about my school work.
   e. He knows nothing about my school work.

3. What would be the lowest marks you could get in school and still satisfy your best friend?

   a. A's
   b. B's
   c. C's
   d. D's
   e. E's
   f. My best friend doesn't care what marks I get in school.

4. How far in school does your best friend think you will go?

   a. quit now
   b. go to high school for a while
   c. graduate from high school
   d. go to a school to be a secretary or learn a trade
   e. go to college for a little while
   f. graduate from college
   g. more than 4 years of college

STOP

DO NOT TURN THE PAGE YET

Pick one. Circle the letter for your answer.

1. Which marks would you have to get to make your favorite teacher happy?

   a. A's
   b. B's
   c. C's
   d. D's
   e. E's
   f. My favorite teacher doesn't care what marks I get in school

2. Does your favorite teacher know how you are doing in school? Pick one.

   a. She knows everything I do in school.
   b. She knows almost everything about my school work.
   c. She knows some things about my school work
   d. She only knows a little bit about my school work.
   e. She knows nothing about my school work.

3. What would be the lowest marks you could get in school and still satisfy your favorite teacher?

   a. A's
   b. B's
   c. C's
   d. D's
   e. E's
   f. My favorite teacher doesn't care what marks I get in school.

4. How far in school does your favorite teacher think you will go?

   a. quit now
   b. go to high school for a while
   c. graduate from high school
   d. go to a school to be a secretary or learn a trade
   e. go to college for a little while
   f. graduate from college
   g. more than 4 years of college

STOP

DO NOT TURN THE PAGE YET

Please write your answers to these questions.

1. If you could have any job, which one would you like to have after you finish school?

2. Sometimes the job you get is not the job you wish for. What kind of job do you think you <u>will</u> get after you finish school?

STOP

DO NOT TURN THE PAGE YET

What does your father do for a living?
(or who ever supports your family)

_____

Tell about what he does on his job.

_____

_____

_____

APPENDIX B

HOYT'S ANALYSIS OF VARIANCE COEFFICIENTS: HEARING IMPAIRED AND NON-IMPAIRED SUBJECTS' RESPONSES TO THE SCA AND SCA-FORM D SCALES

Hoyt's Analysis of Variance Coefficients for the Self-Concept of Academic Ability Scale and the Self-Concept of Academic Ability Scale-Form D: Hearing Impaired and Non-Impaired Subjects.

Scale	Non-Impaired HS Seniors (N=97)	Hearing Impaired Students (N=80)	Non-Impaired 10th Graders (N=1499)
SCA-D	.77	.83	---
SCA	---	---	.85

APPENDIX C

ABSTRACT OF PROPOSED STUDY OF SCALES AND PROCEDURES FOR ASSESSING SOCIAL-PSYCHOLOGICAL CHARACTERISTICS OF VISUALLY IMPAIRED AND HEARING IMPAIRED STUDENTS

A B S T R A C T

Title of Project:	Scales and Procedures for Assessing Social-Psychological Characteristics of Visually-Impaired and Hearing-Impaired Students
Principal Investigator:	Edsel L. Erickson, Ass't. Professor of Sociology & Education
Contracting Agency:	Western Michigan University
Federal Funds Requested:	$8,789
Beginning & Ending Dates:	May 15, 1966 - May 14, 1967

1. Objectives. The objective of this research is to develop reliable and valid parallel instruments and procedures for comparing social-psychological characteristics of visually-impaired students (including the blind and partially-sighted), hearing-impaired students, and students not known to be impaired. Major factors to be investigated are: (1) student's perceptions of how others (parents, teachers and friends) evaluate his academic ability; (2) student's self-concept of academic ability; (3) student's academic achievement; and (4) student's educational and occupational aspirations and plans. Other important variables include: students' perceptions of the expectations involved in their relationships with parents, teachers and friends; and socio-economic status. Major control variables include degree of impairment, age at onset, methods of communication used, and aptitude. Instruments to be used were first developed and standardized

under U.S.O.E. Cooperative Research Projects #845, #1636, and #2831. Modifications of these instruments for translation into sign language and finger spelling have been developed by Erickson, Joiner and Brookover for use with hearing-impaired and visually-impaired children. For the visually-impaired, these instruments will be transcribed into braille and large type. All of the other instruments are appropriate for small group testing.

2. Procedures. Administration will involve test-retest and parallel forms. Populations and samples to be studied include: (1) blind and partially-sighted students in the Michigan School for the Blind (app.N = 100); (2) hearing-impaired students living in the Michigan School for the Deaf (app.N = 100) and the Indiana School for the Deaf (N = 87); and (3) a random sample of non-impaired students drawn from the three public high schools of a midwestern city (app. N = 100). In addition, data drawn from the above populations will be compared with data from five high schools of a city collected by the Principal Investigator under Cooperative Research Project #2831 app.N = 1800 at each grade level). Analysis to determine the reliability of the major instruments will involve assessments of internal consistency, equivalence, and stability. Statistical techniques will include Guttman Scalogram Analysis, Green's Scale analysis, test-retest correlations, analysis of variance, and the "L" test for linear ranks. In addition to a theoretically based logical validation, validity determination will emphasize predictions of achievement and other hypothesized dependent variables and agreement with other measures of postulated constructs.

3. Expected Contribution. The significance of this study is that it will provide the methodological basis and tools for a more definitive study comparing the career development of visually-impaired students

and hearing-impaired students with one another and with a normative population of students not known to be impaired. A more definitive study, which would include rural-urban controls, regional differences, institutional differences, mobility factors, impaired not in institutions, etc., would be very costly. It is also expected that considerable knowledge will be developed concerning procedures for group testing of impaired populations. Finally, it will provide four graduate students with an advanced research experience in the social-psychology of impairment.

APPENDIX D

GENERAL SELF-CONCEPT OF ACADEMIC ABILITY SCALE

Michigan State University

Brookover Studies - General Self-Concept of Ability Scale

Circle the letter in front of the statement which best answers each question.

1. How do you rate yourself in school ability compared with your close friends?

   a. I am the best
   b. I am above average
   c. I am average
   d. I am below average
   e. I am the poorest

2. How do you rate yourself in school ability compared with those in your class at school?

   a. I am among the best
   b. I am above average
   c. I am average
   d. I am below average
   e. I am among the poorest

3. Where do you think you would rank in your high school graduating class?

   a. among the best
   b. above average
   c. average
   d. below average
   e. among the poorest

4. Do you think you have the ability to complete college?

   a. yes, definitely
   b. yes, probably
   c. not sure, either way
   d. probably not
   e. no

5. Where do you think you would rank in your class in college?

   a. among the best
   b. above average
   c. average
   d. below average
   e. among the poorest

Go on to next page

6. In order to become a doctor, lawyer, or university professor, work beyond four years of college is necessary. How likely do you think it is that you could complete such advanced work?

   a. very likely
   b. somewhat likely
   c. not sure either way
   d. unlikely
   e. most unlikely

7. Forget for a moment how others grade your work. In your own opinion, how good do you think your work is?

   a. My work is excellent
   b. My work is good
   c. My work is average
   d. My work is below average
   e. My work is much below average

8. What kind of grades do you think you are capable of getting?

   a. Mostly A's
   b. Mostly B's
   c. Mostly C's
   d. Mostly D's
   e. Mostly E's