AN INVESTIGATION OF THE RELATIONSHIPS BETWEEN

THE BINET VOCABULARY AND PEABODY PICTURE

VOCABULARY TESTS WHEN ADMINISTERED

WITHIN A PRESCHOOL HEAD START

POPULATION

By

KAY HAND MOUNT

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1967

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
MASTER OF ARTS
May, 1968

AN INVESTIGATION OF THE RELATIONSHIPS BETWEEN

THE BINET VOCABULARY AND PEABODY PICTURE

VOCABULARY TESTS WHEN ADMINISTERED

WITHIN A PRESCHOOL HEAD START

POPULATION

Thesis Approved:

Fred Jewell
_____
Thesis Adviser

Irvin Becke
_____

R P Locy
_____

N Durham
_____
Dean of the Graduate College

ii

ACKNOWLEDGMENTS

thesis. Without his untiring patience and encouragement, the final product of this investigation could never have been accomplished.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Investigators concerned with the language functioning of the

human organism have long realized the importance of vocabulary in the

development of linguistic skills and in turn the manipulation of the

environment. The speech pathologist is one such investigator who is

continually faced with the assessment of the language functioning of

individuals. Often a child or adult is presented in the clinical set-

ting who has little speech or language. The reason behind the language

deficit may be the result of brain damage, hearing loss, a culturally

unstimulating environment, or a series of various other difficulties.

But no matter what the etiology of the problem, the speech pathologist

is immediately faced with assessing the ramifications of the disorder.

He is particularly interested in assessing the level at which the

individual is functioning--communication-wise. In other words, what

is the language ability of the individual? What ideas is this indi-

vidual able to communicate to his listener? And in turn, what ideas

or concepts is he able to understand from the speaker?

For many years it has been recognized that one of the best ways

of assessing language functioning is through the measurement of vocabu-

lary. Thus, the speech pathologist administers a vocabulary test,

and it, along with various other types of speech assessment procedures,

enables him to determine the type and level of therapy that should

be most beneficial to the individual.

But what is vocabulary? Why is vocabulary thought to measure language ability? An attempt will be made to answer these two questions which have no doubt sprung to the mind of the reader.

Selected Review of Vocabulary Studies

Webster[1] defines vocabulary as "...a sum or stock of words employed by a language, group, individual, or work or in a field of knowledge." This is a general definition. However, it seems somewhat broad for our purposes. What connotation of the word "vocabulary" does the speech pathologist have when endeavoring to assess language functioning? The speech pathologist usually has a more specific definition in mind when using the word "vocabulary." To him one's vocabulary is the sum total of all the concepts and ideas known to the individual. The psycholinguist might speak of this in terms of symbols. The speech pathologist recognizes these symbols as words--no matter whether nouns, adjectives, or adverbs, etc., are represented. Thus, to summarize, the speech pathologist feels that the symbols which make up the language of the individual are words and that these words in turn make up the vocabulary of that person. The vocabulary of the individual then is the sum total of all the concepts known by that person, symbolized by words, and used to communicate.

However, Gray and Wise[2] warn that, "The number of words...is not necessarily an indication of the number of ideas which a child may attach to them." They explain this by saying that the child often uses many words as complete sentences and that many words are used as different parts of speech thus representing several ideas.

Myklebust[3] feels that language is divided into three types: inner, receptive, and expressive. He describes inner language as "...the use of language symbols for the purposes of inner life or throught; that is, it might be described as that language which in the individual uses for autistic purposes or for 'talking to himself.'" Receptive language is described as "...that language which an individual uses to understand others." He further states that expressive language is "...that language which the individual uses to make himself understood to others." Myklebust feels that after the child has matured to some degree expressive language as well as inner and receptive language may consist of either written or spoken symbols. If one accepts Myklebust's premise concerning language types, one must also accept the idea that there exist three types of vocabulary of the individual. There are inner, receptive, and expressive vocabularies, all consisting of different numbers of words.

Vocabulary size has always been of much interest to investigators. However, with so many types of vocabulary, it is obvious that there are some pitfalls to measurement of them. For example, Smith[4] reported 272 words for the 2 year old, 896 words for the 3 year old, 1540 words for the 4 year old, 2072 words for the 5 year old, and 2562 words for the 6 year old. Grigsby's[5] figures are somewhat higher than these. He reports the 3 year old vocabulary as 1507 words, the 4 year old as 2148, the 5 year old as 2527, and the 6 year old as 3054 words.

These estimates refer to the vocabulary of use. However, Gray and Wise,[6] Eisenson,[7] and Templin,[8] to name only a few, indicate that expressive vocabulary is much smaller than the receptive vocabulary. M. K. Smith[9] reported that the average receptive vocabulary of the

first grade child was 16,900 words with a range of 5,500 to 32,800.

It is interesting to note that these were basic words or ones which

warranted separate entries in a dictionary. When derived terms were

considered, the receptive vocabulary was increased considerably. Thus,

it is obvious that the type of vocabulary being measured must be stated.

Is the investigator measuring the vocabulary which has significance

for the individual when he hears or reads the word or the one which

he actually uses?

One must now turn to the rationale behind the measurement of

vocabulary. Of this Johnson, Darley, and Spriestersbach[10] state:

> Children's knowledge of words has long served as an
> index of their language maturity. Investigators have
> based their judgments concerning progress in language
> development upon the age at which children first begin
> speaking intelligible words, the number of words they
> appear to know at any given age, and their ability to
> define, use, or indicate understanding of selected sam-
> ples of words at various levels of difficulty.

Continuing in this vain, Johnson[11] indicates why vocabulary develop-

ment and thus its measurement is of such importance. He says that:

> ...the acquisition of a vocabulary is a process of dis-
> tinct importance to the child....The finely differentiated
> responses possible through speech are essential for the
> development of symbols that are substituted for complex
> stimulus patterns and experiences....At an early age other
> symbols than speech may be developed but word language
> offers the most rapid acquisition of the many responses
> essential for quick and efficient thinking.

Thus far the above discussion has integrated vocabulary and

language and has shown how one is related to the other. Johnson's

statement has brought the reader one step further and has introduced

"quick and efficient thinking." If this is regarded as intelligence,

a multitude of literature studying the relationship of vocabulary

and intelligence can be found. But before embarking on these studies,

one must first note the statements concerning the relationship of

the development of speech and language to intelligence made by West,

Ansberry, and Carr.[12]

> It is obvious that language must utilize intellectual
> or cognitive factors. Purely emotional cries, such as
> the animal makes, are conceivable without intelligence.
> Such cries may even have communicative significance, but
> they are not properly linguistic, for language is a proc-
> ess of concept building. Emotional cries express native,
> rather than learned, reactions, whereas language shows
> acquired responses. Delay in the development of the
> faculty of acquiring new reaction patterns will, therefore,
> retard the development of language....The faculty of re-
> training and perfecting nonnative reaction patterns is
> intelligence. Many feeble-minded children who have not
> sufficient intelligence to develop a language can still
> produce voice and imitate speech sounds accurately. But
> they have no more real speech than a parrot.

They state further that:

> The learning of speech requires a nice balance between
> experimentation and observation. In some children this
> balance is achieved early and in some late, but until it
> is achieved speech is not learned efficiently. Thus, in
> addition to a general maturation of intelligence, the on-
> set of speech requires constant maintenance of the balance
> of the basic intellectual faculties, as they develop.

With these statements in mind, it seems reasonable to assume

that the measurement of vocabulary would assess a major facet of

intellectual functioning. Terman[13] has insisted since the 1916 revi-

sion of the Simon-Binet scale that vocabulary is the best single

estimate of intellectual functioning. And a statement to this effect

has been placed in each of the manuals of this test from its 1916

revision to the present 1960 revision. In 1918 Terman[14] published

the results of an investigation which substantiated his hypothesis

that vocabulary measurement indicates the level of intellectual

functioning. Terman's study included 631 subjects who were school

students ranging in grade from first year grade school through first

year high school. Their intelligence quotients ranged from under 50
to greater than 150. In this group, he found that vocabulary mental
age as based on the vocabulary list in the 1916 revision of the Simon-
Binet scale correlated .91 with the mental age based on the entire
scale. Terman[15] states:

> It is evident that a mental age based on vocabulary score
> alone would not be far wrong in a large per cent of cases.
> We have determined the probable error of such a mental age,
> which we may call the 'vocabulary mental age,' and found
> it to be approximately 9 1/2 months.

Apart from this empirical study, statistical studies have also
shown the relationship of vocabulary and intelligence. In a study
by Kelley,[16] nine tests were given to kindergarten, third, and
seventh grade students. When a tetrad analysis was applied to the
results, he found that Spearman's[17] elemental "g" could be divided
into three traits. One of these traits was called the "verbal factor."
Schneck[18] also found the existence of a verbal factor as well as a
numerical factor using male college students as subjects. He states
that: "The vocabulary test seems to be the best instrument for the
measurement of that which is common to all the verbal tests." In
1932 Anastasi[19] replicated Schneck's findings. Also, both Schiller[20]
and Thurstone[21] have found the presence of verbal factors when analyz-
ing test results and have shown that vocabulary is an excellent measure
for these factors.

In a more recent study, Templin[22] used materials to evoke 50
consecutive responses from 480 children. She found positive cor-
relation between the number of words used and the level of intelligence.
She states that:

> Correlations between intelligence and the vocabulary
> measures were somewhat lower than reported by other

investigators. In the correlations of intelligence with
the number of different words used in 50 remarks a sharp
break in the magnitude of the correlations occurs at 5
years.

The correlations between intelligence and vocabulary measures dropped

from .57 at 4.5 years to .27 at 5 years. By 8 years the correlation

had reached a low figure of .20.[23] However, Templin[24] makes it

clear that she questions the significance of her findings. She

states:

The substantial decrease in the relationship between
the number of different words used in 50 remarks and
intellectual ability after 5 years of age may indicate
that this is not a satisfactory measure of vocabulary of
use throughout the age range tested. The number of words
in the total vocabulary of use increases with age during
the developmental period. Thus, if a sufficiently large
sample of speech were obtained, the number of different
words used would reflect the total vocabulary of use. It
is probable that 50 remarks provide a sample large enough
to reflect the number of words the young child actually
uses, but it is likely that a larger sample of speech is
necessary to be sensitive to the total number of words
used by older children. That the sample of speech is too
small is a more likely explanation than that the relation-
ship between the number of different words used and intelli-
gence decreases during the early developmental period.

This review of literature was not intended to be a comprehen-

sive survey of studies in vocabulary. It was, however, meant to

be a sample of those studies which show the relationship of vocabu-

lary, language, and intelligence. This review was compiled to show

the importance of vocabulary measurement, especially in the field of

speech pathology, and as partial justification for the present study.

For a more comprehensive review of the tremendous quantity of

vocabulary studies, the reader is referred to McCarthy[25] who presents

an excellent synopsis of various types of vocabulary research. If

one is interested in a comprehensive bibliography of vocabulary

studies, Dale and Reichert[26] have published an exhaustive book by this same name.

Purpose of the Study

This study is designed to assist the speech pathologist in evaluating more effectively the vocabulary mental age scores obtained by preschool children on the Revised Stanford-Binet Vocabulary Test (Form L-M, 1960) and the Peabody Picture Vocabulary Test (Form A). From information previously presented, the author has indicated that the speech pathologist is vitally concerned with the evaluation of the language functioning of various individuals--particularly children seen in the clinical setting. An investigation which is concerned with the comparison of two vocabulary tests which might easily be employed in the clinical setting should aid the speech pathologist in interpretation of such test scores with regard to the language capabilities of the child. This investigation was carried out within a preschool Head Start population. Although the testing was not performed in the clinical setting, it was assumed that information gained from testing in the school-like setting of the Head Start program would be applicable to testing of such children upon referral to speech and hearing clinical services within various professional environments.

The specific purpose of the study is to investigate the intra-test and inter-test relationships between the Revised Stanford-Binet Vocabulary Test (Form L-M, 1960) and the Peabody Picture Vocabulary Test (Form A) when administered within a preschool Head Start population.

For the purposes of this study, intra-test was defined as the relationship between the pre- and post-test data of the same test. Inter-test was definded as the relationship between the two tests.

As implied by the above statements, the Peabody Picture Vocabulary Test and Binet Vocabulary Test were administered on a pre- and post-test basis to a group of preschool children receiving training at the Head Start Child Development Center in Stillwater, Oklahoma, during the 1967-68 academic year. The program is administered by the Payne County Community Action Foundation, Inc. All children were considered to be culturally restricted for the purposes of this investigation on the basis of the low income level of the parents of children allowed to attend the program. Eligibility of the child to attend the program is determined by the criteria concerning income and family number established by the Office of Economic Opportunity. These criteria are listed in Appendix B.

Both the vocabulary tests were assumed to require some mental processes which were similar as well as some which were different in order to respond to the items of the instrument. Due to the nature of the responses, the Binet Vocabulary Test was thought to assess the expressive functioning of the individual while the Peabody Picture Vocabulary Test was though to assess the receptive functioning. However, the underlying aspect of vocabulary itself was considered to be the most important factor measured by either instrument. This aspect will be discussed in greater depth in Chapter III.

Specifically the study sought to investigate the following six questions. The first four concern intra-test relationships while the last two concern inter-test relationships.

1. What is the difference between the means of the pre- and post-test scores of the Binet Vocabulary Test?

2. What is the difference between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test?

3. What is the linear correlation between the pre- and post-test scores of the Binet Vocabulary Test?

4. What is the linear correlation between the pre- and post-test scores of the Peabody Picture Vocabulary Test?

5. What is the linear correlation between the scores of the pre-test Binet Vocabulary Test and the pre-test Peabody Picture Vocabulary Test?

6. What is the linear correlation between the scores of the post-test Binet Vocabulary Test and the post-test Peabody Picture Vocabulary Test?

Scope and Limitations of the Study

This study endeavors to accumulate more normative data concerning the administration of the Binet Vocabulary and Peabody Picture Vocabulary Tests. It is hoped that this data will indicate the feasibility of use of these two instruments by the speech pathologist when evaluating the language capabilities of the culturally deprived child. This data should also indicate how confidently the speech pathologist may accept these test scores. However, it should be pointed out that the scope of this study includes only an evaluation of the relationships between the two tests--both intra-test relationships and inter-test relationships within the specified population.

The intra- and inter-test relationships were examined to show

the limitations which must be kept in mind when administering the
tests within a preschool culturally restricted population. Examination
of these relationships was also made to show the practical application
of the tests. Consideration was also given to possible limitation
variables by determining whether a significant improvement in vocabu-
lary mental age had occurred--as measured by each instrument individu-
ally--within this group of Head Start preschool children.

It can be seen from the scope of this study that there are limi-
tations concerning the significance which can be attributed to its
findings. First and foremost, the results of the study are applicable
only to a preschool culturally restricted population from a Midwestern
non-metropolitan area. One would not be justified in applying the
results to school age culturally deprived children. Nor would one
be justified in generalizing the results to all preschool culturally
restricted children. Generalization is restricted further to those
particular children who have had the opportunity to attend a Head
Start program. Finally, application of the results is limited to
an urban community in the Midwest having a population of approximately
25,000. However, despite the limitations inherent in this study,
this investigation does provide some of the much needed normative
information concerning these two instruments when administered within
a preschool culturally restricted population.

At this time there is increasing national concern regarding the
problems of the preschool culturally deprived child. Evidence of
this is seen in the abundant amount of money being expended by the
federal government on various poverty programs as well as research
projects. Through such activities much information has been gathered

concerning this child. However, one recurrent finding by noted investigators including Templin,[27] McCarthy,[28] Ausubel,[29] Bernstein,[30] and Bereiter and Engelmen[31] is the severe deficit of this child in various areas of language functioning.

Due to our increasing sophistication concerning the needs of the culturally deprived child, the speech pathologist is beginning to see more of these children than in previous years. Therefore, it seems imperative to determine whether standardized measures of vocabulary ability adequately measure this aspect of language functioning within this population. Thus, this study attempts to accumulate normative data concerning the administration of the Binet Vocabulary and Peabody Picture Vocabulary Tests within the preschool Head Start population.

FOOTNOTES

[1]Philip B. Cove, ed., Webster's Seventh New Collegiate Dictionary (7th ed.; Springfield, Mass.: G. & C. Merriam Company, Publishers, 1963), p. 996.

[2]Giles W. Gray and Claude M. Wise, The Bases of Speech (New York: Harper & Brothers, Publishers, 1959), p. 485.

[3]Helmer R. Myklebust, "Aphasia in Children--Language Development and Language Pathology," in Handbook of Speech Pathology, ed. by Lee E. Travis (New York: Appleton-Century-Crofts, Inc., 1957) pp. 504-505.

[4]Madora E. Smith, An Investigation of the Development of the Sentence and the Extent of the Vocabulary of Young Children, University of Iowa Studies in Child Welfare, No. 5 (Iowa City: University of Iowa Press, 1926) pp. 3, 5.

[5]Olive John Grigsby, "An Experimental Study of the Development of Concepts of Relationship in Preschool Children as Evidenced by Their Expressive Ability," Journal of Experimental Education, I (December, 1932), pp. 144-162.

[6]Gray and Wise, pp. 486-487.

[7]Jon Eisenson, et al., The Psychology of Communication (New York: Appleton-Century-Crofts, 1963), p. 210.

[8]Mildred C. Templin, Certain Language Skills in Children: Their Development and Interrelations, Institute of Child Welfare Monograph Series, No. 26 (Minneapolis: The University of Minnesota Press, 1957), p. 105.

[9]Mary K. Smith, "Measurement of the Size of General Vocabulary through the Elementary Grades and High School," Genetic Psychology Monographs, XXIV (November, 1941), pp. 311-345.

[10]Wendell Johnson, Frederic L. Darley, and D. C. Spriestersbach, Diagnostic Methods in Speech Pathology (New York: Harper & Row, Publishers, 1952), pp. 173-174.

[11]Buford J. Johnson, Child Psychology (Springfield, Ill.: Charles C. Thomas, 1932), p. 204.

[12]Robert West, Merle Ansberry, and Anna Carr, The Rehabilitation of Speech (New York: Harper & Row Publishers, 1957), pp. 54-55.

[13]Lewis M. Terman, The Measurement of Intelligence (Boston: Houghton Mifflin Company, 1916), p. 230.

[14]Lewis M. Terman, "The Vocabulary Test as a Measure of Intelligence," Journal of Educational Psychology, IX (October, 1918), pp. 452-466.

[15]Ibid., p. 454.

[16]Truman L. Kelley, Crossroads in the Mind of Man (Stanford University: Stanford University Press, 1928), p. 107.

[17]Charles Spearman, The Abilities of Man (New York: Macmillian Company, 1927), pp. 87-97.

[18]M. R. Schneck, "The Measurement of Verbal and Numerical Abilities," Archives of Psychology, No. 107 (1929), pp. 1-49.

[19]Anne Anastasi, "Further Studies on the Memory Factor," Archives of Psychology, No. 142 (1932), pp. 1-60.

[20]B. Schiller, "Verbal, Numerical, and Spatial Abilities of Young Children," Archives of Psychology, No. 161 (1934), pp. 1-69.

[21]L. L. Thurston, "Primary Mental Abilities," Psychometric Monographs, No. 1 (1938), pp. 1-121.

[22]Templin, p. 120.

[23]Ibid., p. 118.

[24]Ibid., p. 119.

[25]Dorothea McCarthy, "Language Development in Children," in Manual of Child Psychology, ed. by Leonard Carmichael (New York: Chapman and Hill, Limited, 1954), pp. 526-536.

[26]Edgar Dale and Donald Reichert, Bibliography of Vocabulary Studies (Columbus: Ohio State University, 1957).

[27]Templin, pp. 147-149.

[28]Dorothea McCarthy, The Language Development of the Preschool Child, Institute of Child Welfare Monograph Series, No. 4 (Minneapolis: University of Minnesota Press, 1930), pp. 57, 90, 106, 123.

[29]David P. Ausubel, "Effects of Cultural Deprivation on Learning Patterns," Audiovisual Instruction, X (January, 1965), pp. 10-12.

[30]Basil Bernstein, "Social Class and Linguistic Development: A Theory of Social Learning," in Education, Economy, and Society, ed. by A. H. Halsey, Jean Floud, and C. Arnold Anderson (New York: Free Press of Glencoe, 1961).

[31]Carl Bereiter and Siegfried Englemann, _Teaching Disadvantaged Children in the Preschool_ (Englewood Cliffs, N. J.: Prentice Hall, 1966).

CHAPTER II

METHOD AND PROCEDURE

This chapter includes a description of the conditions under which
the present study was conducted and the procedures which were employed.
The subjects involved in the study and the test instruments will be
discussed. Also, included will be the six hypotheses to be tested in
this investigation. Finally the treatment of the data will be de-
scribed.

Subjects

The subjects included the 43 children between the ages of 3 years,
6 months and 6 years of age who were receiving training at the Head
Start Child Development Center in Stillwater, Oklahoma, during the
1967-68 school year. This was assumed to be a random sample of sub-
jects in that they represent 43 such pupils within this age range
from all such children enrolled in Head Start programs throughout the
state. To be included in this study, the child's birthdate must have
been on or before April 1, 1964. Thus, all children were 3 years,
6 months or older before the pre-testing began.

Three restrictions were placed upon selection of subjects. The
first was failure to respond to any one of the speech frequencies
(500 Hz, 1000 Hz, and 2000 Hz) in either ear at 20 dB HL ISO 1964 as
assessed by an audiometric sweep check. Referral was made to the

Oklahoma State University Speech and Hearing Clinic for pure-tone threshold testing for those children who failed the initial screening. The outcome of the pure-tone threshold testing determined inclusion in the study. The children were not excluded from the investigation simply because of failure on the screening since it was thought that children of this age may, on first assessment of hearing acuity, be too naive to understand what is required of them in a hearing test. On the other hand, passing the screening warranted automatic inclusion in the study. No children were excluded from the investigation on the basis of poor auditory acuity.

It was anticipated that some children either would not or could not communicate with the investigator. Thus, a second and mandatory selection restriction was exclusion of all such children. Only one child was restricted from the group for this reason.

Finally, only those children who participated in both pre- and post-testing were included as subjects for this investigation. Three subjects who were unable to attend the Head Start program after the initial pre-testing period were omitted from the study.

Thus, four children were eliminated on the basis of the above criteria. The remaining 39 children comprised the subjects for this investigation.

It should further be stated that no restriction was placed on inclusion concerning race or sex. And as mentioned previously, all children represented families from low socio-economic levels and were considered to be culturally restricted for the purposes of this investigation.

The subjects included 21 males and 18 females. Race was

represented in the form of 12 Caucasoids, 1 Mongoloid (American Indian), and 26 Negroids. Nineteen of the children had attended the program the previous year, while 20 were attending for the first time. Of the 19 who had attended previously, 16 were Negroes and 3 were Caucasians. Although no statistical analysis was performed on the above data, it is presented here for purposes of describing the sample.

Choice of Test Instruments

The two tests involved in this study were the Peabody Picture Vocabulary Test (Form A) and the Binet Picture Vocabulary as well as the Vocabulary Words from the 1960 revision of the Stanford-Binet Intelligence Scale (Form L-M). Both tests were scored according to manual instructions. Each test yielded a vocabulary mental age score for the child. This score was designated in terms of months.

There were four reasons for selection of these two instruments. First, the investigator was interested in determining the relationship between two recognized measures of vocabulary. As mentioned before, both vocabulary tests were thought to require some mental processes which were similar as well as some which were different in order to respond to the items of the instrument. Due to the nature of the responses, the Binet Vocabulary Test was thought to assess the expressive vocabulary functioning while the Peabody Picture Vocabulary Test was thought to assess the receptive functioning. The investigator was primarily interested in determining the relationships between two tests which assessed two forms of vocabulary ability.

Second, there is little information concerning administration of

these two instruments within a culturally restricted population.
Such information is necessary when evaluating the results of these
tests when administered to such children in the clinical setting.

The third reason for choosing these two tests concerns a matter
of longevity. The Peabody Picture Vocabulary Test was first published
by Dunn in 1959. Thus, this test is still under ten years of age.
The vocabulary test of the Stanford-Binet revision has been in exist-
ence in similar form to the one of today from its first introduction
by Terman in the 1916 Stanford-Binet Intelligence Scale. With the
1937 revision came the Picture Vocabulary of that test. Many of the
same words which appeared in this test in 1937 are contained in the
1960 form. Comparatively speaking, the Peabody Picture Vocabulary
Test is quite new and the reliability and validity data concerning
this test is still being gathered. Investigations concerning the
reliability and validity of the Binet Vocabulary Test, and the entire
scale for that matter, are also being made at this time. However,
the Binet Vocabulary Test has been in existence for many years and
has been shown to have high correlation with the entire Revised
Stanford-Binet Intelligence Scale as well as with other tests of
intelligence. Thus, the Binet Vocabulary Test serves in one sense
as a criterion measure in lieu of the full scale against which the
Peabody Picture Vocabulary Test may be compared.

The fourth and final reason for choosing these two instruments
involves frequency of use. The Peabody Picture Vocabulary Test is
used frequently by the speech pathologist as a measure of language
functioning. However, the Binet Vocabulary Test is rarely used during
speech evaluation to assess language functioning. The Peabody Picture

Vocabulary Test takes a considerable amount of time to administer--
especially if the child is above average in intelligence. The Binet
Vocabulary Test, on the other hand, takes considerably less time to
administer to the preschool child. If these two instruments assess
vocabulary ability equally well, it seems strange that one should
always use a longer test. When time is limited, or the young child
is unusually restless, it would be helpful to know whether these
tests were interchangeable. Thus, the relationship between these tests
needs to be determined in order to decide whether one could be used
in lieu of the other.

These four reasons prompted the choice of the two test instru-
ments and contributed to the framework of the problem. The specific
purpose--to investigate various relationships between the two tests--
was an outgrowth of the ideas and questions underlying these four
considerations.

Administration of Tests

The Binet Vocabulary Test (Form L-M, 1960) and the Peabody
Picture Vocabulary Test (Form A) were administered on a pre- and post-
test basis to 39 Head Start children ranging in age from 3 years, 6
months to 6 years. The pre-testing was done during the fourth, fifth,
and sixth weeks of class work. The post-testing was done during the
twenty-third, twenty-fourth, and twenty-fifth weeks of classes. Thus,
all children had received approximately 19 weeks or 5 months of train-
ing in the Head Start program during the intervening period. During
the pre-testing an audiometric sweep check was also administered to
assess hearing acuity.

The tests were administered in a quiet room of the Washington
School where the Head Start Center of Stillwater, Oklahoma, is located.
Only the child and the examiner were present in the room during the
testing. All the tests were administerd by the same examiner--the
author.

The children were chosen at random for testing on each given
day of the test period. One-half of the children were given the Binet
Vocabulary Test first and the Peabody Picture Vocabulary Test second.
The order of the testing was reversed for the other half. This pro-
cedure was followed during both pre- and post-test sessions. This
was done to reduce the chance of practice effects of the first test
which might affect the results of the second instrument. Both tests
were administered during a single test session.

The audiometric screening was always administered last during
the pre-test sessions. It was felt that a test of this nature might
frighten the child if given initially, thus reducing the adequacy
of the results on the vocabulary tests. It was felt that the vocabu-
lary testing was more suitable to establish rapport. After the child
had gained confidence in the examiner, it was thought that a more
adequate screening of hearing acuity could be obtained.

The following are the instructions given by the examiner when
introducing the two vocabulary tests to the child. The instructions
given were those suggested in the manual of each test.

The Peabody Picture Vocabulary Test was introduced in this
manner:

> ...I want to play a picture game with you. Turn to Example
> A and say: See all the pictures on this page. (Indicate
> this by pointing to each in turn.) I will say a word, then

> I want you to put your finger on the picture of the word
> I have said. Let us try one. Put your finger on 'bed.'
> When a subject makes the desired response, turn to Example
> B saying: That's fine. Now put your finger on 'fish.'
> Then turn to Example C saying: Good! Show me 'butterfly.'
> Then say: Fine! Now I am going to show you some other
> pictures. Each time I say a word, you find the picture
> of it. When we get along further in the book you may not
> be sure you know the word, but I want you to look carefully
> at all of the pictures anyway and choose the one you think
> is right. Point to _____.[1]

The Binet Picture Vocabulary Test was introduced in the follow-

ing manner. The child's attention was called to the first picture.

He was then asked: "...What's this? What do you call it?"[2] From

these instructions it should be noted that the Binet Picture Vocabu-

lary Test unlike the Peabody Picture Vocabulary Test has no trial

series, and the test is begun immediately.

When the child was able to respond correctly to 14 or more of

the 18 words of the Binet Picture Vocabulary, the Binet Vocabulary

Words requiring definitions were administered. The following are

the instructions suggested in the manual preceding the administra-

tion of the Vocabulary Words.

> ...I want to find out how many words you know. Listen,
> and when I say a word, you tell me what it means. What
> is an orange? Vary the form of the question to avoid a
> stilted manner of presentation, e.g., What does......
> mean? Tell me what a.......is, or give just the word with-
> out further question. If S hesitates, urge him to try by
> saying, Just tell me in your own words; say it any way you
> please. All I want to know is whether you know what a.......
> is. Or You know what a.......is! Tell me what is a.......?
> If the child can read, give him the vocabulary card and
> let him look at each word as you read it.
>   If the child's meaning is not clear, that is, if his
> response can't be scored either plus or minus without
> further explanation, say, Tell me what you mean, or Tell
> me more about it.[3]

Treatment of the Data

This study was concerned with the intra- and inter-test rela-
ships between the Binet Vocabulary Test and the Peabody Picture
Vocabulary Test. Intra-test was defined as the relationship between
the pre- and post-test data of the same test. Inter-test was defined
as the relationship between the two tests. This study sought to
investigate six questions--four concerning the intra-test relation
ships and two concerning the inter-test relationships.

The four questions concerning the intra-test relationships were:

1. What is the difference between the means of the pre- and post-
   test scores of the Binet Vocabulary Test?

2. What is the difference between the means of the pre- and post-
   test scores of the Peabody Picture Vocabulary Test?

3. What is the linear correlation between the pre- and post-test
   scores of the Binet Vocabulary Test?

4. What is the linear correlation between the pre- and post-test
   scores of the Peabody Picture Vocabulary Test?

   The two questions concerning the inter-test relationships were:

5. What is the linear correlation between the scores of the pre-
   test Binet Vocabulary Test and the pre-test Peabody Picture
   Vocabulary Test?

6. What is the linear correlation between the scores of the post-
   test Binet Vocabulary Test and the post-test Peabody Picture
   Vocabulary Test?

   The null and alternate hypotheses to be tested on the basis of
the above six questions were these:

1. <u>Null Hypothesis</u>: There is no significant difference between the means of the pre- and post-test scores of the Binet Vocabulary Test (H_O: $\mu_1 = \mu_2$).

   <u>Alternate Hypothesis</u>: There is a significant difference between the means of the pre- and post- test scores of the Binet Vocabulary Test (H_O: $\mu_1 \neq \mu_2$).

2. <u>Null Hypothesis</u>: There is no significant difference between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test (H_O: $\mu_1 = \mu_2$).

   <u>Alternate Hypothesis</u>: There is a significant difference between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test (H_O: $\mu_1 \neq \mu_2$).

3. <u>Null Hypothesis</u>: The pre- and post-test scores of the Binet Vocabulary Test are not significantly correlated (H_O: $\rho = 0$).

   <u>Alternate Hypothesis</u>: The pre- and post-test scores of the Binet Vocabulary Test are significantly correlated (H_1: $\rho \neq 0$).

4. <u>Null Hypothesis</u>: The pre- and post-test scores of the Peabody Picture Vocabulary Test are not significantly correlated (H_O: $\rho = 0$).

   <u>Alternate Hypothesis</u>: The pre- and post-test scores of the Peabody Picture Vocabulary Test are significantly correlated (H_1: $\rho \neq 0$).

5. <u>Null Hypothesis</u>: The pre-test scores of the Binet Vocabulary Test and the pre-test scores of the Peabody Picture Vocabulary Test are not significantly correlated (H_O: $\rho = 0$).

   <u>Alternate Hypothesis</u>: The pre-test scores of the Binet Vocabulary Test and the pre-test scores of the Peabody Picture Vocabulary

Test are significantly correlated (H_o: $\rho \neq 0$).

6. **Null Hypothesis:** The post-test scores of the Binet Vocabulary
   Test and the post-test scores of the Peabody Picture Vocabulary
   Test are not significantly correlated (H_o: $\rho = 0$).

   Alternate Hypothesis: The post-test scores of the Binet Vocabulary
   Test and the post-test scores of the Peabody Picture Vocabulary
   Test are significantly correlated (H_o: $\rho \neq 0$).

   Two statistical procedures were utilized to test the signficance
of these hypotheses. An application of the t test on paired observa-
tions was employed to test those hypotheses concerned with differences
between means. The obtained value of t was then compared with tabu-
lated t at the .01 level of significance for 38 degrees of freedom
and a two-tailed test. A Pearson product-moment correlation coef-
ficient was computed for each hypothesis involving linear correlation.
The significance of this value was then compared with the critical
value of the correlation coefficient for 37 degrees of freedom and
a two-tailed test at the .01 level of significance. More informa-
tion concerning these statistical procedures will be presented in
Chapter V.

FOOTNOTES

[1]Lloyd M. Dunn, _Manual for the Peabody Picture Vocabulary Test_ (Minneapolis: American Guidance Service, 1959), p. 7.

[2]Lewis M. Terman and Maud A. Merrill, _Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M_ (Boston: Houghton Mifflin, 1960), p. 68.

[3]_Ibid._, p. 83.

CHAPTER III

THE TEST INSTRUMENTS

This chapter will present a description of the Binet Vocabulary
and Peabody Picture Vocabulary Tests. This description will include
information concerning standardization of the tests as well as admin-
istration and scoring. These two test instruments will also be dis-
cussed in terms of the type of vocabulary ability assumed to be mea-
sured. Specifically, the rationale for assuming that the Binet
Vocabulary Test measures expressive vocabulary ability while the
Peabody Picture Vocabulary Test measures receptive ability will be
presented.

Peabody Picture Vocabulary Test

The Peabody Picture Vocabulary Test consists of 150 numbered
plates. Each plate contains four heavy-line drawings representing
an object, activity, or state of being. Three example plates are
also included. Two forms (A and B) of this test have been con-
structed.

According to Dunn[1] the test "...is designed to provide an esti-
mate of a subject's verbal intelligence through measuring his hearing
vocabulary." He states further that:

> ...the scale may be given to any English speaking resident
> of the United States between 2 years, 6 months and 18 years
> who is able to hear words, see the drawings, and has the

facility to indicate 'yes' and 'no' in a manner which com-
municates.[2]

Test construction of the Peabody Picture Vocabulary Test began
with an initial group of 3,885 words from Webster's New Collegiate
Dictionary. Several series of plates were constructed from these
words. Each series was administered to several hundred subjects.
Item difficulty was determined for each plate at each age level.
Finally, 150 plates and two forms of the test--Form A and Form B
--were developed. Thus, only the "best" 300 words were retained
as determined by the item difficulty of each plate during the initial
testing period.

The test was then standardized on 4,012 white children from the
Nashville, Tennessee area.[3] The two forms of the test were adminis-
tered in counterbalanced order with three to seven days intervening
between tests. High correlation coefficients were found between the
two forms. These correlations ranged from .67 at the 6 year level
to .84 at both the 17 and 18 year level. Thus, reliability between
the two forms was considered to be sufficient. A synopsis of many
reliability studies concerning this test is presented in the 1965
edition of the manual. Many studies concerning the reliability of
this instrument are also presented in Chapter IV. Thus, no further
discussion of this aspect will be presented at this time.

Statistical validity was determined following test publication
in 1960. Several of these studies were presented in the 1959 edition
of the manual. Considerable information concerning the validity of
the test has been gathered since that time. Brief summaries of such
studies are also presented in the 1965 edition of the manual. Several

studies of this nature are also presented in Chapter IV. Thus, detailed description of such investigations will not be presented at this time. Suffice it to say that high positive correlations have been found between the Peabody Picture Vocabulary Test and various other tests including the Revised Stanford-Binet Intelligence Scale, Wechsler Intellignece Scale for Children, and the Columbia Test of Mental Maturity to name only a few.[4]

The scale is designed for administration over the critical range of items for the particular individual. In other words, the starting point, basal, and ceiling items differ from one individual to another. The child is requested to point to or indicate in some way an oral stimulus word given by the examiner. The method of introducing the test is indicated in detail in Chapter II, thus, discussion will be eliminated here. After sufficient examples have been presented to enable the child to understand the nature of the task and the required response, the examiner begins the test at the appropriate level for the child's age. This is determined by a table provided in the test manual which indicates the age level of the child and the plate number upon which the examination should begin.[5] For example, if the child is below 3 years, 3 months, he should begin on plate number 1. If he is between the ages 3 years, 3 months and 4 years, 2 months, he should begin on plate number 15, etc. When the child is suspected of functioning at a subnormal level of intelligence, the examiner may begin the test at a level lower than that recommended, taking into consideration the mental age of the child.

The examiner works forward from the starting point until the subject makes an error. When eight consecutive correct responses

have not been made preceding this error, the examiner works backward
from the starting point until eight such responses have been made.
This is the basal level--the first response in the series of eight
consecutive responses. The experimenter continues the test forward
from the first incorrect response until the child incorrectly identi-
fies six out of eight consecutive responses. The last item presented
is considered to be the ceiling item. Thus, the test is continued
only until the basal and ceiling have been established.

One is interested in finding the raw score or the number of
correct responses. The raw score is determined by subtracting the
number of errors made between the basal point and the ceiling item
from the number of the ceiling item or the last word presented.

The raw score can then be converted to mental age scores, intel-
ligence quotients, or percentiles by using the appropriate tables
established by Dunn from data gathered from the normative population.
In the present study, only mental age scores were considered. Accord-
ing to Dunn these age equivalents provide an index of the level of
development for a given subject.[6]

The above system of administration and scoring was utilized
in this investigation. No changes were needed due to inadequacies
of the test or idiosyncrasies of the subjects.

Binet Vocabulary Test

This study involves the use of the vocabulary subtest of the
Revised Stanford-Binet Intelligence Scale (Form L-M, 1960). Thus,
we must speak of standardization of the vocabulary subtest in terms
of the full scale Binet. There are of course three revisions of the

Binet-Simon scale. The first was in 1916, the second in 1937, and the third in 1960.

For the purposes of this study, the revision and standardization of the 1960 scale seems most germane. The following provides a brief summary of information pertinent to this revision and standardization.

The 1960 Stanford Revision retains the basic characteristics of the Binet type scale. These features include the use of age standards of performance and the concept of measurement in terms of general mental adaptability. The best subtests were taken from the 1937 scales, Forms L and M. Thus, the 1960 scale embodies a single form known as the L-M.

Subtests to be included in the 1960 scale were selected on the basis of records of tests administered during the years from 1950 to 1954. The subjects utilized in the evaluation of the subtests consisted of 4,498 subjects ranging in age from 2 years, 6 months to 18 years. The test records of this group were compared to the standardization group of the 1930's upon which the 1937 revision was based. In order to determine the changes in difficulty of subtests, a comparison was made between the per cent passing the individual tests in the 1950's and the per cent passing in the 1930's. Selection of test items was then based on: (1) an increase in the per cent of children passing at succeeding mental age levels, and (2) the validity based on the biserial correlation of test items with the total score.

Changes incorporated in the 1960 revision included: (1) elimination or relocation of those tests which were found to have significantly changed in difficulty level since the 1937 standardization; (2) elimination or substitution of tests found unsuitable on the

bases of cultural changes; (3) clarification of scoring procedures and test administration; and (4) correction of inadequacies of struc- ture in the 1937 scale by introducing adjustments to equate average mental age and average chronological age at each age level and by revising and extending IQ tables to compromise for atypical varia- bility of IQ at various age levels. The latter adjustment allows for comparison of standard score IQ's at all age levels.[7]

Little information is presented in the test manual concerning standardization of the vocabulary subtest. However, the above dis- cussion provides a general survey of the techniques used in the revi- sion of this subtest. We do know from the test manual, however, that high positive biserial correlation was found between the vocabulary subtests and the total score at all age levels. These correlations ranged from a high of .96 at the superior adult level to a low of .61 at the 2 year, 6 month level.[8]

The vocabulary subtest consists of a picture form for young chil- dren and a word form for older children and adults. The Picture Vocabulary form consists of eighteen picture plates. All words represented by these pictures are nouns. All eighteen plates are presented to the child regardless of the number missed. The child is asked to simply give the name of the picture. The purpose of this test as stated by Termin and Merrill is to "...determine whether the sight of a familiar object in a picture provokes recognition and calls up the appropriate name."[9]

The Vocabulary Word form consists of forty-five words. Most of these are nouns but some verbs and adjectives are included. These words are presented on cards to the adult and orally to the child.

The purpose of this test is "...to determine whether the subject knows the meaning of the word, not whether he can give a completely logical definition. Awkwardness of expression is disregarded."[10] In this investigation termination of the test was based on six consecutive incorrect responses. This criterion for termination is provided in the test manual. Terman and Merrill[11] report that there "...is little likelihood of success beyond the point where six words have been failed."

Scoring standards are presented for each form of the vocabulary subtest. However, on the Picture Vocabulary it can be generally stated that a minus is scored for descriptions or responses in terms of use. A minus is also scored for naming only part of the pictured item. For example, responding with _toes_ for _foot_ would be scored minus. No general statements can be applied to the Vocabulary Word form. However, determining whether the meaning is comprehended receives primary concern.

Standards for passing the Binet Vocabulary Test at each age level are presented in the test manual. However, these standards encompass such a wide age range that interpolation of the scores was found necessary for this investigation.

Other modifications were found to be necessary due to this wide age range. The scoring manual requires a score of 14 correct responses on the Picture Vocabulary to obtain a mental age of 4 years. No further scoring standards are presented until the 6 year level. To obtain this mental age requires six correct responses on the Vocabulary Words. Due to this, two procedures were necessary. First, the scoring standards were interpolated downward from the 4 to 6 year

old levels. Interpolation downward was instigated at all other levels
as well. Second, it seemed imperative to set fourteen correct
responses on the Picture Vocabulary as the criterion for continuing
the test on the Vocabulary Word form. It seemed possible on the
basis of the scoring standards that a child who had scored 14 or
more on the Picture Vocabulary might achieve a mental age of 6 years
on the Vocabulary Words if given the opportunity. Thus, the criterion
was established that all children obtaining fourteen or more correct
responses on the Picture Vocabulary would also be allowed to perform
on the Vocabulary Words. If the child was able to score at the 6
year level or above, he was credited with this mental age. If he
was allowed to perform on the Vocabulary Words but did not obtain
a score at the 6 year level or beyond, he was credited with the high-
est mental age he received on the Picture Vocabulary Test.

The scoring standards used on both the Picture Vocabulary Test
and the Vocabulary Words are presented in Table I. Only those scor-
ing standards necessary for this investigation are presented.

TABLE I

STANDARDS FOR PASSING THE BINET VOCABULARY TEST

Mental Age Level	No. Correct
*2 yrs.	3 plus
2 yrs., 1 mo.	4 plus
2 yrs., 2 mo.	5 plus
2 yrs., 3 mo.	6 plus
2 yrs., 4 mo.	7 plus
*2 yrs., 5 mo.	
*2 yrs., 6 mo.	8 plus
2 yrs., 9 mo.	9 plus
3 yrs.	10 plus
3 yrs., 3 mo.	11 plus

TABLE I (Continued)

Mental Age Level	No. Correct
3 yrs., 6 mo.	12 plus
*3 yrs., 9 mo.	13 plus
4 yrs.	14 plus
4 yrs., 6 mo.	15 plus
*5 yrs.	16 plus
6 yrs.	6 plus on Voc. Words
*7 yrs.	7 plus on Voc. Words
8 yrs.	8 plus on Voc. Words

*These scoring standards were presented in the test manual. All others are interpolated.

Vocabulary Ability Measured by the Two Instruments

Previously, it has been mentioned that the Binet Vocabulary Test and the Peabody Picture Vocabulary tests were assumed to require some mental processes which were similar as well as some which were different in order to respond to the items of the instrument. It has also been stated that the Binet Vocabulary Test was thought to assess the expressive functioning of the individual while the Peabody Picture Vocabulary Test was thought to assess receptive functioning. At this time further information regarding these assumptions will be presented.

To determine the type of vocabulary ability measured by each instrument, one must analyze the incoming stimulus, the response of the child, and the intellectual processes of the child. The primary incoming stimulus of the Peabody Picture Vocabulary Test is visual in nature in the form of a picture plate. The child is asked to point to a certain picture. Thus, the response is in terms of a gesture. Between the stimulus input and the response output, the

child must comprehend the meaning of the word. Intellectual function-
ing here may be thought of in terms of the language process of decoding
proposed by Osgood.[12] Thus, the primary vocabulary ability measured
here is visual-motor-decoding according to McCarthy and Kirk's[13] modi-
fication of Osgood's language model. Decoding according to McCarthy
and Kirk[14] is "...the sum total of those habits required to ultimately
obtain meaning from either auditory or visual linguistic stimuli...."
The channels of communication according to them are the stimulus input
and the response output. Thus, one sees that the reception and com-
prehension of the linguistic unit is the primary ability assessed by
this vocabulary test. In other words, receptive vocabulary ability
is measured here.

The Binet Vocabulary Test can be analyzed similarly. The incom-
ing stimuli for the picture form is primarily visual. The response
is vocal in nature. During the time between the stimulus and the
response, the child must comprehend the word and go through whatever
processes are necessary in order for him to name the word. Thus the
primary vocabulary ability measured by this test is visual-vocal-
encoding. The ability measured by the Vocabulary Word form when the
word is presented orally would be auditory-vocal-encoding due to the
change in the incoming stimuli. Encoding according to McCarthy and
Kirk[15] is "...the sum total of those habits required to ultimately
express oneself in words or gestures...." Thus, one sees that the
primary ability assessed by this test is expressive vocabulary.

Expressive and receptive vocabulary are used here in the same
terms that Myklebust uses expressive and receptive language. His
ideas were presented in Chapter I and will therefore not be repeated

at this time.

A thorough description of Osgood's language model and McCarthy and Kirk's modification of this model is beyond the scope of this discussion. However, their model provides an excellent method of presenting the above information.[16]

As one readily sees from the information above, encoding first requires comprehension or cognition. On the other hand comprehension and cognition are the end-products of decoding.[17] Thus, it is evident that some similar mental operations are involved in each of these tests. However, the differing nature of the stimuli and responses probably indicate differing mental operations also.

Although the basic aspect of vocabulary functioning is probably more important than the particular form of vocabulary ability assessed, the above information was presented to shed light on the results of this study. However, it must be remembered that the above deals with assumptions analyzed in terms of a theoretical model. Thus, any conclusions based on this discussion are merely speculative.

FOOTNOTES

[1] Lloyd M. Dunn, Expanded Manual for the Peabody Picture Vocabulary Test (Minneapolis: American Guidance Service, Inc., 1965), p. 25.

[2] Ibid.

[3] Ibid., p. 27. To obtain information concerning the sampling techniques utilized in order to allow application of the normative data throughout the United States, the reader is referred to the manual.

[4] Ibid., pp. 32-42.

[5] Ibid., p. 8.

[6] Ibid., p. 10.

[7] This information was paraphrased from Lewis M. Terman and Maud A. Merrill, Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M (Boston: Houghton Mifflin, 1969), pp. 39-40.

[8] Ibid., pp. 342-347.

[9] Ibid., p. 126.

[10] Ibid., p. 233.

[11] Ibid., p. 84.

[12] Charles E. Osgood, et al., Contemporary Approaches to Cognition (Cambridge: Harvard University Press, 1957), pp. 75-118.

[13] James J. McCarthy and Samuel A. Kirk, The Construction, Standardization and Statistical Characteristics of the Illinois Test of Psycholinguistic Abilities (n.p.: By the Authors, 1963), pp. 1-13.

[14] Ibid., p. 2.

[15] Ibid.

[16] It should be noted that neither Osgood's model nor McCarthy and Kirk's modification of it is presented in its complete form.

[17] See Osgood, p. 77, for information concerning organizational levels for greater depth of these concepts.

CHAPTER IV

REVIEW OF THE LITERATURE

This chapter is devoted to a selected review of the previous

investigations concerning the Peabody Picture Vocabulary Test (PPVT)

and the Stanford-Binet Vocabulary Test (BVT).

Peabody Picture Vocabulary Test

Since the Peabody Picture Vocabulary Test was first published

by Dunn in 1959, many studies have been concerned with reliability

and validity of the test. Six studies which were completed soon

after the publication of the test served as investigations of relia-

bility and validity with the following populations: "normal" American,

mentally retarded, cerebral palsied, and "normal" English. These

studies were reported in the original test manual. In these investi-

gations test-retest reliability was examined. Also, alternate forms

of the test were used over short intervals of time. The following

is a brief synopsis of these studies.

Norris, Hottel, and Brooks[1] compared the Peabody Picture Vocabu-

lary scores of 60 fifth graders under group and individual administra-

tion. They found that the "...form of the test, type of administration,

testing order, nor any combination of these factors produced a change

in PPVT scores which could not be attributed to chance at the .05

level of significance."

Tempero and Ivanoff[2] also administered the PPVT on a group basis
to 150 seventh grade pupils. Both forms of the test were given. A
reliability coefficient of .75 between the two forms was found. Evi-
dence of congruent validity was also found when the PPVT was compared
with an achievement test and two intelligence tests. Correlation with
the Henmon-Nelson Tests of Mental Ability ranged from .61 to .64.
Correlations with the California Test of Mental Maturity ranged from
.34 on the non-language portion to .65 on the language portion. Cor-
relations with the California Achievement Test ranged from .40 to .63.

Dunn and Brooks[3] conducted a study in 1960 to collect data on
the ability of the PPVT to measure the verbal intelligence of educable
mentally retard children. Using 371 subjects ranging in age from 6
to 18 years, a reliability coefficient of .83 between Forms A and B
was found. When PPVT scores were compared with Revised Stanford-
Binet (RS-B, 1937) scores, PPVT MA's were on the average 4.5 months
lower than RS-B MA's. IQ's on the PPVT averaged 2.1 points above
RS-B IQ's. A validity coefficient of .76 was found when MA's on
the PPVT and RS-B were compared.

Dunn and Hottel[4] administered both forms of the PPVT to 220 train-
able mentally retarded subjects ranging in age from 6 to 16 years.
A correlation coefficeint of .84 was found when the mental age scores
on the two forms were compared. When MA scores on the PPVT and 1937
revision of the RS-B were correlated, a correlation coefficient of
.66 was found. MA scores on the PPVT compared with teachers's ratings
of reading and writing showed correlation coefficients ranging from
.39 to .60.

Dunn and Harley[5] compared the PPVT with the Van Alstyne Picture

Vocabulary Test, the Columbia Mental Maturity Scale, and the Ammons

Full-Range Picture Vocabulary Test in a group of 20 cerebral palsied

children ranging in age from 7 years, 1 month to 16 years, 2 months.

A reliability coefficient of .97 for the alternate forms of the PPVT

was found. Intercorrelations between the scores of the four tests

exceeded .80.

A study concerning the clinical utility of the PPVT with 101

English children aged 6 to 9 years was conducted by Moss and Edmonds.[6]

They found no differences between the mean scores of the English chil-

dren and the American children of the same age on whom the test was

standardized.

Since these first six studies, many more investigations have

been carried out concerning the reliability and validity of the test

as well as its clinical utility with various populations. A sampling

of twelve such studies was made for inclusion in this chapter. This,

by no means, exhausts the number of studies concerning the PPVT. But

these twelve studies seem to be a representative sample of the work

done. The studies to be reviewed have been grouped according to the

population investigated, since this seemed to be the only feasible

means of classification. The populations studied include: the mentally

retarded--8 studies, "normal" subjects--3 studies, and subjects requir-

ing psychological referral--1 study. As indicated by this enumeration,

most of the investigations have been concerned with the use of the

PPVT with mentally subnormal children. A brief review of the results

of these studies will be given.

Tobias and Gorelick[7] studied the validity of the PPVT with 107

retarded adults and found that the PPVT tends to over-rate the MA

and IQ of such subjects. The authors hypothesized that "...vocabu-lary skills of retardates may continue to mature beyond that of other intellectual factors measured by existing instruments."

Similar to the study by Tobias and Gorelick, an investigation by Allen, Haupt, and Jones[8] indicated that the PPVT overestimates the intellectual ability of retarded children with severe visual perceptual difficulties. These results suggest that the test should not be used with such populations.

Kimbrel,[9] also like Tobias and Gorelick, showed the PPVT IQ overestimates the IQ on the RS-B as well as on the Ammons Full-Range Picture Vocabulary Test when investigating 63 adolescent educable mental retardates.

Budoff and Purseglove[10] studied the performance of 46 retarded adolescents from 16 to 18 years of age on the PPVT. For the entire sample, they found high correlation with the 1937 and 1960 RS-B and high correlations between the alternate forms of the test. However, when the moderately and severely retarded subjects were considered separately, their scores on the PPVT underestimated that on the RS-B. This finding was in sharp contrast to that of Tobias and Gorelick.

Kahn[11] also administered the PPVT to adolescent and young adult retardates. However, he was interested in the long-term reliability of the test. He administered Form A annually for four years to 141 subjects. Correlation coefficients ranged from .71 to .87, thus, comparable to short term test-retest intervals.

More information concerning the validity of the PPVT was found by Throne, Kasper, and Schulman[12] who administered the test to 35 mentally retarded boys. When compared with the WISC, RS-B (1937),

and the Goodenough Draw A Person Test, significant correlations were obtained. However, no correlation was found between the Metropolitan Achievement Test and the PPVT.

Mein[13] used the PPVT with 80 mentally defective patients in England ranging in age from 10 to 30 years. The investigator found a correlation of .71 between MA scores on the 1937 RS-B and PPVT. However, item analysis showed that some plates were misplaced in order of difficulty. The author concluded that the PPVT was also a valid and reliable vocabulary test for the English mental defective.

Shipe, Cromwell, and Dunn[14] investigated a somewhat different problem than is customary with the PPVT. They analyzed the responses of 60 institutionalized retarded children and adults on the PPVT items. All the subjects had more difficulty on those items with human content. However, those who were emotionally disturbed had significantly more difficulty than the nondisturbed on the items with human content. The authors speculated that previous disturbed interpersonal relationships interfered with performance on such items.

A review of the three studies concerning "normal" subjects will now be presented. Weeks[15] made one such study. His subjects included 240 college students. Group administration was used. The test was found to be effective as a verbal intelligence screening device for these subjects since all scored above the established norms for 18 year olds.

A second study done on a "normal" population was that of Fargo, et al.,[16] whose subjects included 135 third, fourth, and fifth grade students. Group television and individual administration of the PPVT was compared. The investigators found no significant difference

between individual and group administration scores on either Form A or B and concluded that television administration was feasible.

Milgram and Ozer[17] administered the PPVT and RS-B (1960) to two groups of low socio-economic status Negro children from a large Eastern city. The 116 subjects of this study ranged in age from 3 to 6 years of age. When the PPVT scores were compared with the RS-B scores, it was found that the PPVT scores were consistently lower. The authors concluded that the PPVT may be more susceptible to environmental impoverishment than the RS-B.

The following is a study involving subjects referred to a psychological clinic. This study was done by Gage and Naumann[18] for the purpose of determining if the PPVT was reliable and valid as a measure of intelligence of those children referred for psychological evaluation. The subjects included 30 children ranging in age from 5 to 15 years. A correlation of .68 was found between the WISC and the PPVT. The authors concluded that the instruments were measuring different things. They suggest that the PPVT is a valuable measure of "understanding" vocabulary but should not be substituted as a measure of intelligence.

As indicated by the above review of research, only one study to the author's knowledge has been specifically concerned with the performance of low socio-economic status preschool children on the PPVT. This points further to the need for the present investigation.

Binet Vocabulary Test

Terman included an extensive vocabulary test in his first revision of the Binet-Simon scale for measuring intelligence. Terman[19]

felt that his vocabulary test had "...a far higher value than any other single test in the scale." Reference has already been made to Terman's[20] 1918 investigation of the vocabulary test which he conducted when the test had been criticized by other psychologists. He concluded in this investigation that the vocabulary test was a reliable measure of intelligence. In 1927 Lincoln[21] verified the results found in Terman's 1918 study.

Terman[22] and Whipple[23] have shown that the test is influenced by sex. Merrill[24] and Sunne[25] have shown it to be influenced by race. Cuff[26] also indicates that socio-economic status affects test results.

Thus, one sees that much investigation was done on the 1916 vocabulary list. Another such study was that by Mahan and Witmer[27] who administered the test to 269 children referred to a psychiatric hospital and clinic. A regression equation was then used on the obtained data to determine a mental age score for each vocabulary raw score. Their study was an attempt to facilitate scoring of the test.

Louden[28] was also interested in Terman's vocabulary test and investigated the relative difficulty of the two vocabulary lists. She concluded that it was desirable to use the combined lists if a true estimate of mental ability were to be obtained.

In 1937 Terman and Merrill published a new revision of the Sanford-Binet scale. Soon afterward Atwell[29] compared the vocabulary scores of the 1916 revision with that of the new 1937 revision. When both tests were administered to 100 adults from a psychiatric hospital, a correlation between the two instruments of .86 was found.

A higher average mental age rating was found on the 1937 revision than on the 1916 revision.

Carlton[30] investigated the effect of chronological age on the RS-B vocabulary score at the moron and imbecile levels. The study was undertaken to determine whether chronological age affected the vocabulary scores of such an atypical population. From the results of this study, Carlton concluded that chronological age along with mental age is important in the explanation of scores on the RS-B Vocabulary Test.

Spache[31] found that the MA's determined from a combined administration of the Picture Vocabulary and Vocabulary Word Test predicted scale results with fair accuracy. However, he warned that the MA scores found from such a combination of items underestimates the MA's obtained on the full scale. This information is particularly interesting with regard to the present investigation.

A more recent study concerning the relation of the MA scores on the vocabulary test and the scale as a whole was done by Dunsdon and Roberts.[32] The subjects included 450 English children ranging in age from 7 years to 14 years, 2 months. Correlations between MA scores on the vocabulary test and the entire scale ranged from .60 to .90. The lower correlations were found during the seventh, eighth, and ninth years of life.

In 1954 Cureton[33] felt that a table of mental age equivalents should be provided for clinicians administering the Vocabulary Word Test of forty-five items. Thus, he reviewed McNemar's[34] statistical study of the RS-B and on the basis of this data calculated mental age equivalents for the vocabulary test. That he regards the vocabulary

test as something less than perfect is seen in the following remark:

> Any inaccuracies in the table, whether due to the method
> of its derivation or to the limitations of the original
> standardization data, are small in comparison to the
> unreliability of the test.[35]

However, he estimated the correlation between the vocabulary mental

age and "true" mental age to be .825.

May and Perry[36] compared the RS-B Vocabulary Test (Form L) and

the Columbia Mental Maturity Scale in a group of 51 mentally retarded

children. A correlation coefficient of .43 was obtained between the

two measures. The authors concluded that this was an inadequate

degree of relationship to allow interchangeable use of the two tests.

Dundsdon and Roberts[37] made a study of four vocabulary tests

with a population of 2,000 English children between the ages of 5

years and 14 years, 2 months. Their report of the investigation

includes a table of norms for the RS-B Vocabulary Word Test. This

table indicates the number of words a child of a specific age would

be expected to know. This information along with the MA scores derived

by Cureton gives much helpful information to the investigator using

the vocabulary test.

In 1958 Levinson[38] became concerned that a culture change had

made certain Binet tasks more familiar than others. Thus, he decided

to reevaluate the use of the RS-B Vocabulary Test (Form L) as a test

of intelligence. His subjects included 640 native and foreign born

kindergarten and primary school children ranging in age from 4 years

to 9 years, 11 months. A statistically significant difference was

found in the MA scores of the two groups. This difference favored

the native born child. The mean vocabulary MA was found to over-

estimate the MA on the full scale for the native born child while

underestimating that of the foreign born child. He concludes that
while the Binet Vocabulary Test is still a valid index of intelligence,
it should not be used alone as a short verbal scale of intelligence.

This review has shown the variety of studies concerning the Binet
Vocabulary Test from its first inclusion in the Stanford-Binet Intel-
ligence Scale in 1916 to the present time. To the author's knowledge,
no studies have been published concerning the 1960 revision of the
Stanford-Binet Vocabulary Test which is most important to the present
investigation. However, the 1937 revision and the 1960 revision of
the vocabulary tests--both picture and word--are almost identical.
Thus, any information concerning the 1937 revision should have signifi-
cant implications for the present form of the test.

FOOTNOTES

[1]Raymond C. Norris, John V. Hottel, and Sadye T. Brooks, "Comparability of Peabody Picture Vocabulary Test Scores Under Group and Individual Administration," Journal of Educational Psychology, LI (April, 1960), pp. 87-91.

[2]Howard Tempero and John Ivanoff, "Effectiveness of the Peabody Picture Vocabulary Test with Seventh-grade Pupils," Journal of Educational Research, LVIII (May-June, 1965), pp. 412-415.

[3]Lloyd M. Dunn and Sadye T. Brooks, "Peabody Picture Vocabulary Test Performance of Educable Mentally Retarded Children," Training School Bulletin, LVII (August, 1960), pp. 35-40.

[4]Lloyd M. Dunn and Hottel, "Peabody Picture Vocabulary Test: Performance of Trainable Mentally Retarded Children," American Journal of Mental Deficiency, LXV (March, 1961), pp. 448-452.

[5]Lloyd M. Dunn and Randall K. Harley, "Comparability of the Peabody, Ammons, Van Alstyne, and Columbia Test Scores with Cerebral Palsy Children," Exceptional Children, XXVI (October, 1959), pp. 70-74.

[6]James W. Moss and Phyllis Edmonds, "The Peabody Picture Vocabulary Test with English Children," British Journal of Educational Psychology, XXX (February, 1960), p. 82.

[7]Jack Tobias and Jack Gorelick, "The Validity for the Peabody Picture Vocabulary Test as a Measure of Intelligence of Retarded Adults," Training School Bulletin, LVIII (November, 1961), pp. 92-98.

[8]Robert Allen, Thomas Haupt, and R. Wayne Jones, "A Suggested Use and Non-use for the Peabody Picture Vocabulary Test with the Retarded Child," Psychological Reports, XV (October, 1964), pp. 421-422.

[9]Don Kimbrel, "Comparison of PPVT, FRVT, RS-B, and Academic Achievement Scores Among Institutionalized Educable Mental Retardates," Perceptual and Motor Skills, XXIII (December, 1966) p. 1178.

[10]Milton Budoff and Eleanor Purseglove, "Peabody Picture Vocabulary Test: Performance of Institutionalized Mentally Retarded Adolescents," American Journal of Mental Deficiency, LXVII (May, 1963), pp. 756-760.

[11]Harris Kahn, "Evidence for Long-term Reliability of the PPVT with Adolescent and Young Adult Retardates," American Journal of Mental Deficiency, LXX (May, 1966), pp. 895-898.

[12]Frances M. Throne, Joseph C. Kasper, and Jerome L. Schulmann, "Peabody Picture Vocabulary Test in Comparison with Other Intelligence Tests and an Achievement Test in a Group of Mentally Retarded Boys," Educational and Psychological Monographs, XXV (Summer, 1965), pp. 589-595.

[13]R. Mein, "Use of the Peabody Picture Vocabulary Test with Severely Subnormal Patients," American Journal of Mental Deficiency, LXVII (September, 1962), pp. 269-273.

[14]Dorothy Shipe, Rue Cromwell, and Lloyd Dunn, "Responses of Emotionally Disturbed and Non-disturbed Retardates to PPVT Items of Human Versus Nonhuman Content," Journal of Consulting Psychology, XXX (October, 1966), pp. 439-443.

[15]Richard W. Weeks, "Effectiveness of the Peabody Picture Vocabulary Test with College Students," Journal of Educational Research, LVII (October, 1963), pp. 109-110.

[16]George Fargo, et al., "Comparability of Group Television and Individual Administration of the Peabody Picture Vocabulary Test: Implications for Screening," Journal of Educational Psychology, LVIII (June, 1967), pp. 137-140.

[17]Norman Milgram and Mark Ozer, "Peabody Picture Vocabulary Test Scores of Preschool Children," Psychological Reports, XX (July, 1967), pp. 779-784.

[18]Gerald Gage and Theodore Naumann, "Correlation of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children," Journal of Educational Research, LVIII (July, 1965), pp. 466-468.

[19]Lewis M. Terman, The Measurement of Intelligence (Boston: Houghton Mifflin Company, 1916), p. 230.

[20]Lewis M. Terman, "The Vocabulary Test as a Measure of Intelligence," Journal of Educational Psychology, IX (October, 1918), pp. 452-466.

[21]Edward Lincoln, "The Reliability of the Stanford-Binet Scale and the Consistency of the Intelligence Quotients," Journal of Educational Psychology, XVIII (December, 1927), pp. 621-626.

[22]Terman, The Measurement of Intelligence, pp. 68-72.

[23]Guy M. Whipple, Manual of Mental and Physical Tests (Baltimore: Warwick and York, Inc., 1920), p. 211.

[24]Maud A. Merrill, "Mental Differences in Children Referred to the Psychological Clinic," Journal of Applied Psychology, X (December, 1926), pp. 470-486.

[25]Dagne Sunne, "Comparison of White and Negro Children in Verbal and Nonverbal Tests," School and Society, XIX (April, 1924), pp. 469-472.

[26]N. B. Cuff, "Social Status and Vocabulary," Journal of Genetic Psychology, XLVI (March, 1935), pp. 226-229.

[27]Harry Mahan and Louise Witmer, "A Note on the Stanford-Binet Vocabulary Test," Journal of Applied Psychology, XX (April, 1936), pp. 258-263.

[28]Mary Louden, "Relative Difficulty of Vocabulary Lists in the Stanford-Binet Scale," Journal of Educational Research, XXVI (April, 1933), pp. 601-607.

[29]C. R. Atwell, "Comparison of Vocabulary Scores on the Stanford-Binet and the Revised Stanford-Binet," Journal of Educational Psychology, XXX (September, 1939), pp. 467-469.

[30]Theodore Carlton, "The Effect of Chronological Age on Revised Stanford-Binet Vocabulary Score at the Moron and Imbecile levels," Journal of Genetic Psychology, LXI (December, 1942), pp. 321-326.

[31]George Spache, "The Vocabulary Tests of the Revised Stanford-Binet as Independent Measures of Intelligence," Journal of Educational Research, XXXVI (March, 1943), pp. 512-516.

[32]M. I. Dunsdon and J. A. Roberts, "The Relation of the Terman-Merrill Vocabulary Test to Mental Age in a Sample of English Children," British Journal of Statistical Psychology, VI (November, 1953), pp. 61-70.

[33]Edward E. Cureton, "Mental Age Equivalents for the Revised Stanford-Binet Vocabulary Test," Journal of Consulting Psychology, XVIII (October, 1954), pp. 381-383.

[34]Quinn McNemar, The Revision of the Stanford-Binet Scale (Boston: Houghton Mifflin, 1942).

[35]Cureton, p. 383.

[36]Theodore May and Harold Perry, "The Relationship Between the Stanford-Binet (Form L) Vocabulary and the Columbia Mental Maturity Scale in a Group of Mentally Retarded Children," American Journal of Mental Deficiency, LXII (September, 1957), pp. 330-333.

[37]M. I. Dunsdon and J. A. Roberts, "A Study of the Performance of 2,000 Children on Four Vocabulary Tests: Norms, with Some Observations on the Relative Variability of Boys and Girls," British Journal of Statistical Psychology, X (May, 1957), pp. 1-16.

[38]Boris M. Levinson, "Reevaluation of the Revised Stanford-Binet Scale, Form L Vocabulary as a Test of Intelligence for the Kindergarten and Primary School Child," Journal of Genetic Psychology, XCIII (September, 1958), pp. 237-248.

CHAPTER V

PRESENTATION AND ANALYSIS OF DATA

This chapter is concerned with presentation of the data accumu-
lated in this investigation. The data will be analyzed in terms of
the intra- and inter-test relationships.

The purpose of this study was to investigate the intra-test
and inter-test relationships between the Revised Stanford-Binet
Vocabulary Test (Form L-M, 1960) and the Peabody Picture Vocabulary
Test (Form A) when administered within a preschool Head Start popu-
lation. These two instruments were administered on a pre- and post-
test basis to 39 Head Start children ranging in age from 3 years,
6 months to 6 years.

The study sought to investigate six questions--four concerning
the intra-test relationships and two concerning the inter-test rela-
tionships. These questions were listed in Chapter II, p. 23. There-
fore, they will not be repeated here. From these six questions, six
hypotheses were formulated. This chapter will present and analyze
the data concerning each hypothesis. All hypotheses will be tested
at the .01 significance level.

To facilitate analysis of the data, each hypothesis will be pre-
sented separately along with the statistical procedures employed to
test it. Discussion of the statistical method used as well as pre-
sentation of the data and statistical computations necessary for

testing each hypothesis will reveal whether a particular hypothesis should be accepted or rejected.

The first four hypotheses are concerned with the intra-test relationships between the pre- and post-test scores of the two test instruments. The following is the first such hypothesis considered in this investigation.

Null Hypothesis: There is no significant difference between the means of the pre- and post-test scores of the Binet Vocabulary Test (H_0: $\mu_1 = \mu_2$).

Alternate Hypothesis: There is a significant difference between the means of the pre- and post-test scores of the Binet Vocabulary Test (H_1: $\mu_1 \neq \mu_2$).

To test this hypothesis an application of the t test on paired observations was used. The experiment utilized paired observations since the investigation was concerned with how the same individual performed on a single test given twice--in a pre- and post-test fashion. Thirty-nine children participated in the experiment. Thus, there were thirty-nine pairs of scores. As mentioned previously, the pre-test was given at the beginning of the training period in the Head Start program while the post-test was administered approximately five months later.

Pairing the scores on the pre- and post-tests of each individual serves to minimize the experimental error. The extraneous variance which exists from pair to pair is eliminated by computing the variance of the differences between the pairs of observations rather than the variance among the individuals within the sample.[1]

Li[2] illustrates this point by showing that if a certain number,

say, 10, is added to each of two observations of the first pair and

20 is added to the two observations of the second pair, etc., the

differences between the pairs remain the same. Thus, the value of

t remains unaffected. Li[3] continues with the following statements:

> ...the IQ's of the children may vary a great deal from
> one pair to another, yet the accuracy of the experiment
> is not affected. Moreover, if the teaching methods are
> tried on children of different IQ's, the experiment has
> the advantage of having a wide inductive basis. A con-
> clusion based on children of various IQ's is more apt to
> be generally valid than the one based on children with
> a particular IQ.

From Li's statements we understand that wide variations can exist

but do not affect the accuracy of the experiment when a pairing

procedure has been utilized.

With this information concerning paired observations, one may

proceed with analysis of the data. The reader is referred to Table II

where the test scores of each individual on the pre- and post-test

Binet Vocabulary Test are presented. Also included is the necessary

information for the calculation of t.

With reference to Table II, D is the difference between the

vocabulary mental age scores of each individual on the pre- and post-

test. D^2 is self-explanatory as the square of these differences.

The thirty-nine differences shown in Table I constitute a sample

drawn from a parent population which consists of an infinite number

of potential observations. Any future experiments of the same kind

will be further samples drawn from the same parent population. Thus,

our null hypothesis (H_0: $\mu_1 = \mu_2$) in terms of paired variates reduces

to testing that the mean of the population of differences is zero

(H_0: $\mu_D = 0$). The alternate hypothesis is that the mean of the

TABLE II

VOCABULARY MENTAL AGE SCORES IN MONTHS
ON THE BINET VOCABULARY TEST

Subject Number	Vocabulary Mental Age Scores		Difference	
	Pre-BVT X_1	Pst-BVT X_2	D $(X_1 - X_2)$	D^2 $(X_1 - X_2)$
1	42	54	12	144
2	72	84	12	144
3	45	48	3	9
4	39	45	6	36
5	45	48	3	9
6	60	60	0	0
7	42	42	0	0
8	42	54	12	144
9	39	60	21	441
10	42	54	12	144
11	39	45	6	36
12	60	54	-6	36
13	48	54	6	36
14	39	39	0	0
15	39	54	15	225
16	42	54	12	144
17	33	39	6	36
18	36	30	-6	36
19	48	60	12	144
20	42	48	6	36
21	36	48	12	144
22	33	36	3	9
23	39	42	3	9
24	33	42	9	81
25	42	48	6	36
26	42	48	6	36
27	27	27	0	0
28	36	42	6	36
29	27	36	9	81
30	39	60	21	441
31	45	48	3	9
32	33	33	0	0
33	39	39	0	0
34	30	33	3	9
35	36	42	6	36
36	72	84	12	144
37	33	39	6	36
38	45	42	-3	9
39	42	45	3	9
n=39	ΣX_1=1623	ΣX_2=1860	ΣD=237	ΣD^2=2925
	$\bar{x}_1$=41.6153	$\bar{x}_2$=47.6923	$\bar{d}$=6.0769	
	s_1=9.9541	s_2=12.0157		

TABLE II (Continued)

H_0:	$\mu_1 = \mu_2$		$s_{\bar{d}}^2$	$=$	$\dfrac{\Sigma D^2 - \dfrac{(\Sigma D)^2}{n}}{n(n-1)}$	$=$	1.002
$n =$	39						
$\Sigma D =$	237		$s_{\bar{d}}$	$=$	$\sqrt{s_{\bar{d}}^2}$	$=$	1.001
$(\Sigma D)^2 =$	56,169				$\dfrac{\Sigma D}{n}$	$=$	6.077
$\Sigma D^2 =$	2,925		$\bar{d}$	$=$			

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{6.077}{1.001} = 6.071 \text{ with } 38 \text{ df}$$

population of differences is not zero (H_1: $\mu_D \neq 0$).

The lower portion of Table II presents the necessary information for the computation of the statistic t. This information includes both formulas and calculations. Since this information is presented in such detail, little explanation seems necessary at this point. Thus, a recapitulation of the computation of t will not be presented. However, further explanation concerning the symbols used in this computation does seem appropriate. While most of these symbols can be readily understood, two of them require definition. Thus, it should be said that $s_{\bar{d}}^2$ is the variance of the mean differences while $s_{\bar{d}}$ is the standard deviation of the mean differences.

The statistic $t = \dfrac{\bar{d}}{s_{\bar{d}}}$ is symmetrically distributed about a mean of zero and corresponds to the tabulated distribution of Student's t. Thus, the test criterion is distributed as t when the null hypothesis is true and the assumption is correct that the differences are normally distributed. In other words, when the null hypothesis is false,

a value from another distribution has been computed--not Student's t.

When the sample t is greater than the tabulated t value with which it is compared, the null hypothesis is rejected and the alternate hypothesis is accepted. With reference to the present problem, tabulated $t_{.01}$ for 38 degrees of freedom (n - 1) and a two-tailed test is 2.711. Since the computed value of t is 6.071, the null hypothesis is rejected. Thus, the alternate hypothesis that there is a significant difference between the means of the pre- and post-test scores of the Binet Vocabulary Test at the 1% level is accepted. This statement is accepted with 99% confidence.

The second hypothesis considered in this investigation was:

<u>Null Hypothesis</u>: There is no significant difference between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test (H_0: $\mu_1 = \mu_2$).

<u>Alternate Hypothesis</u>: There is a significant difference between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test (H_1: $\mu_1 \neq \mu_2$).

To test this hypothesis an application of the t test on paired observations was again utilized. As before there were thirty-nine pairs of scores. Again, our null hypothesis (H_0: $\mu_1 = \mu_2$) in terms of paired variates reduces to testing that the mean of the population of differences is zero (H_0: $\mu_D = 0$). The alternate hypothesis is that the mean of the population of differences is not zero (H_1: $\mu_D \neq 0$).

Table III presents the pre- and post-test scores of each individual on the Peabody Picture Vocabulary Test. The lower portion of this table presents the necessary formulas and calculations for the

TABLE III

VOCABULARY MENTAL AGE SCORES IN MONTHS ON THE
PEABODY PICTURE VOCABULARY TEST

Subject Number	Vocabulary Mental Age Scores Pre-PPVT Y_1	Pst-PPVT Y_2	Difference D $(Y_1 - Y_2)$	D^2 $(Y_1 - Y_2)^2$
1	42	44	2	4
2	65	71	6	36
3	82	75	-7	49
4	51	53	2	4
5	53	55	2	4
6	43	61	18	324
7	39	39	0	0
8	35	44	9	81
9	55	69	14	196
10	45	58	13	169
11	42	45	3	9
12	44	49	5	25
13	34	41	7	49
14	35	48	13	169
15	42	49	7	49
16	46	59	13	169
17	38	47	9	81
18	39	42	3	9
19	47	71	24	576
20	42	43	1	1
21	39	43	4	16
22	32	32	0	0
23	34	46	12	144
24	38	49	11	121
25	58	59	1	1
26	45	55	10	100
27	24	26	2	4
28	43	47	4	16
29	30	32	2	4
30	32	39	7	49
31	48	58	10	100
32	31	34	3	9
33	45	48	3	9
34	27	27	0	0
35	40	51	11	121
36	61	78	17	289
37	27	36	9	81
38	44	46	2	4
39	49	49	0	0
n=39	ΣY_1=1666	ΣY_2=1918	ΣD=252	ΣD^2=3072
	$\bar{y}_1$=42.7180	$\bar{y}_2$=49.1795	$\bar{d}$=6.4615	
	s_1=11.2249	s_2=12.5087		

TABLE III (Continued)

H_0:	$\mu_1 = \mu_2$		$s_{\bar{d}}^2 =$	$\dfrac{\Sigma D^2 - \dfrac{(\Sigma D)^2}{n}}{n(n-1)} =$	.974
$n =$	39				
$\Sigma D =$	252		$s_{\bar{d}} =$	$\sqrt{s_{\bar{d}}^2} =$	.987
$(\Sigma D)^2 =$	63,504				
$\Sigma D^2 =$	3,072		$\bar{d} =$	$\dfrac{\Sigma D}{n} =$	6.462

$$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{6.462}{.987} = 6.547 \text{ with 38 df}$$

computations of the statistic t.

As before, tabulated $t_{.01}$ for 38 degrees of freedom and a two-tailed test is 2.711. The computed value of t is 6.547. Since the computed value of t is greater than the tabulated value of t, one may accept the alternate hypothesis that there is a significant difference between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test at the 1% level. This statement is accepted with 99% confidence.

The third hypothesis considered in the investigation was:

Null Hypothesis: The pre- and post-test scores of the Binet Vocabulary Test are not significantly correlated (H_0: $\rho = 0$).

Alternate Hypothesis: The pre- and post-test scores of the Binet Vocabulary Test are significantly correlated (H_1: $\rho \neq 0$).

This hypothesis was tested through the use of Pearson Product-Moment correlation. To arrive at the linear correlation coefficient, the following formula was employed: $r = \dfrac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$. Item number two

of Table IV presents all the necessary information for computing this statistic. In this case, the linear correlation was found to be .85443. To test the significance of this value of r, the statistic was compared with tabulated values of r. These tabulated values of r represent the value of the correlation coefficient needed to meet the requirements of significance at various levels for a two-tailed test and for samples of various sizes. The appropriate degrees of freedom available for the sample size are represented by the formula n - 2, where n is the number of pairs of observations. Thus, in this case, one is interested in the critical values of r at the .01 level for 37 degrees of freedom. When the computed r is greater than or equal to the tabulated r, the null hypothesis is rejected and the alternate is accepted. With reference to the present problem, the tabulated value of r at the .01 level of significance for 37 degrees of freedom and a two-tailed test is .408. The computed value of r is .85443. The computed value is much greater than the tabulated value. Thus, the null hypothesis is rejected. The alternate hypothesis that the pre- and post-test scores of the Binet Vocabulary Test are significantly correlated at the .01 level is accepted.

The following is the fourth and final hypothesis considered under the classification of the intra-test relationships:

Null Hypothesis: The pre- and post-test scores of the Peabody Picture Vocabulary Test are not significantly correlated (H_0: $\rho = 0$).

Alternate Hypothesis: The pre- and post-test scores of the Peabody Picture Vocabulary Test are significantly correlated (H_1: $\rho \neq 0$).

Item number one of Table IV shows that the linear correlation

between the pre- and post-test scores of the Peabody Picture Vocabulary Test is .87058. Since the tabulated value of r for 37 degrees of freedom at the .01 level of significance and a two-tailed test is .408, one may accept the alternate hypothesis that the pre- and post-test scores of the Peabody Picture Vocabulary Test are significantly correlated. This statement can be accepted with 99% confidence.

TABLE IV

INTRA- AND INTER-TEST CORRELATIONS

Tests Compared	Σxy	Σx^2	Σy^2	r*
1. Pre- vs Pst-PPVT	4644.97461	4787.89746	5945.74414	.87058
2. Pre- vs Pst-BVT	3883.38477	3765.23145	5486.30859	.85443
3. Pre-PPVT vs Pre-BVT	2551.76953	4787.89746	3765.23145	.60100
4. Pst-PPVT vs Pst-BVT	4193.15430	5945.74414	5486.30859	.73417

$$*r = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \Sigma y^2}}$$

The fifth and sixth hypotheses tested in this investigation are listed under the classification of inter-test relationships. The following is the first of these hypotheses concerned with the relationship between the two test instruments:

Null Hypothesis: The pre-test scores of the Binet Vocabulary Test and the pre-test scores of the Peabody Picture Vocabulary Test are not significantly correlated (H_0: $\rho = 0$).

Alternate Hypothesis: The pre-test scores of the Binet Vocabulary Test and the pre-test scores of the Peabody Picture Vocabulary Test are significantly correlated (H_1: $\rho \neq 0$).

By referring to Table IV, item three, it can be seen that the computed r is .60100. Since this value is greater than the tabulated value of r at the .01 level of significance for 37 degrees of freedom and a two-tailed test, the alternate hypothesis is accepted.

The final hypothesis considered in this study is presented below:

Null Hypothesis: The post-test scores of the Binet Vocabulary Test and the post-test scores of the Peabody Picture Vocabulary Test are not significantly correlated (H_0: $\rho = 0$).

Alternate Hypothesis: The post-test scores of the Binet Vocabulary Test and the post-test scores of the Peabody Picture Vocabulary Test are not significantly correlated (H_1: $\rho \neq 0$).

Reference to Table IV, item four, shows that the linear correlation is .73419. This value is also greater than the tabulated value of r at the .01 level of significance for 37 degrees of freedom and a two-tailed test. Thus, the alternate hypothesis is accepted with 99% confidence.

FOOTNOTES

[1]Robert G. D. Steel and James H. Torrie, _Principles and Procedures of Statistics_ (New York: McGraw-Hill Book Company, Inc., 1960), p. 78.

[2]Jerome C. R. Li, _Statistical Inference I_ (Ann Arbor: Edwards Brothers, Inc., 1964) p. 111.

[3]Ibid.

CHAPTER VI

SUMMARY AND CONCLUSIONS

This chapter will be devoted to restatement of the problem,
recapitulation of the results, discussion and interpretation of
results, and recommendations for further research.

Restatement of the Problem

The purpose of this study was to investigate the intra-test and
inter-test relationships between the Revised Stanford-Binet Vocabulary
Test (Form L-M, 1960) and the Peabody Picture Vocabulary Test (Form A)
when administered within a preschool Head Start population. The
subjects included 39 children between the ages of 3 years, 6 months
and 6 years who were receiving training at the Head Start Child
Development Center in Stillwater, Oklahoma during the 1967-68 aca-
demic year. The tests were administered on a pre- and post-test basis
with nineteen weeks intervening between the first and second adminis-
trations.

Recapitulation of Results

Six hypotheses were tested in this investigation. Four concerned
intra-test relationships while two concerned inter-test relationships.
To test those hypotheses concerned with differences between means, an
application of the t test on paired observations was utilized. The

computed value of t was then compared with the tabulated value of t at the .01 level of significance for 38 degrees of freedom and a two-tailed test. To test those hypotheses involving linear correlation, a Pearson product-moment correlation coefficient was first computed. To test its significance, this value was then compared with the critical value of the correlation coefficient for 37 degrees of freedom and a two-tailed test at the .01 level of significance.

Analysis of the intra-test relationships revealed the following results:

1. There was a significant difference between the means of the pre- and post-test scores of the Binet Vocabulary Test.

2. There was a significant difference between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test.

3. The pre- and post-test scores of the Binet Vocabulary Test were significantly correlated.

4. The pre- and post-test scores of the Peabody Picture Vocabulary Test were significantly correlated.

Analysis of the inter-test relationships revealed the following results:

5. The pre-test scores of the Binet Vocabulary Test and the pre-test scores of the Peabody Picture Vocabulary Test were significantly correlated.

6. The post-test scores of the Binet Vocabulary Test and the post-test scores of the Peabody Picture Vocabulary Test were significantly correlated.

Discussion and Interpretation of Results

The finding that there is a significant difference between the
means of the pre- and post-test scores of the Binet Vocabulary Test
indicates a change in the vocabulary mental age scores between the
first and second administration of the test. This change was in the
form of improvement in the scores. If the test measures what it
purports to measure, or in other words is valid, this change in the
form of improvement might be thought of as vocabulary growth. Whether
this change, improvement, or growth in vocabulary is due to matura-
tion of the child and/or experience in the training program cannot
be determined on the basis of this investigation. It can merely be
stated that improvement in terms of vocabulary mental age scores was
shown.

A significant difference between the means of the pre- and post-
test scores of the Peabody Picture Vocabulary Test was also found.
Thus, the above statements apply to the interpretation of this find-
ing.

The finding that the pre- and post-test scores of the Binet
Vocabulary Test are significantly correlated seems to imply that the
test is reliable in terms of yielding similar results from one test
administration to another. The term "similar results" as it is used
here means that those persons who score high on the pre-test also tend
to score high on the post-test. It follows that those who score low
on the pre-test also tend to score low on the post-test.

The pre- and post-test scores of the Peabody Picture Vocabulary
Test were also significantly correlated. Therefore, this test

instrument also seems to yield reliable scores upon testing and
retesting.

The pre-test scores of the Binet Vocabulary Test are signifi-
cantly correlated with the pre-test scores of the Peabody Picture
Vocabulary Test. There was also a significant linear correlation
between the post-test scores of these instruments. These findings
imply that those children who scored high on one test also scored
high on the other test. The same holds true for low scorers.

It will be remembered that comparison of the pre-test scores of
the two instruments resulted in a correlation coefficient of .60.
Comparison of the post-test scores of the two instruments resulted in
a correlation coefficient of .73. These high positive correlations
may be due to the underlying factor of vocabulary which both instru-
ments purport to measure. However, the decrease in the magnitude of
these correlations from perfect positive correlation may possibly be
due to the differences in mental operations and responses required
by the two instruments. The latter two statements are of course
beyond the scope of the statistical analysis and should be considered
merely speculative.

At this point one wonders whether these two tests could be used
interchangeably. In other words, could one test be used in lieu of
the other? This question cannot be answered on the basis of this
study. This investigation allows us to say only that there was high
positive correlation between the two instruments and that the children
of this study who scored high or low on one test tended to score simi-
larly on the other.

The critical reader may also be concerned about the discrepancy

between the correlation coefficients obtained when comparing the pre-tests of the two instruments and the post-tests of both instruments. As stated above, these were .60 and .73, respectively. Again, the statistical analysis cannot answer this question concerning the cause of this discrepancy. However, one may speculate concerning the factors which would appear to affect the magnitude of these correlation coefficients. The maturation of the child and the experience of the training program are two such factors. Due to both of these elements, attention span and tolerance for the test situation may have increased, thus, increasing the correlation coefficient obtained when comparing the post-tests of both instruments. Other factors, of course, may be operating as well.

Conclusions

A significant difference was found between the means of the pre- and post-test scores of the Binet Vocabulary Test. Thus, this test appears to measure the change or improvement in vocabulary mental age scores within this population. A significant difference was also found between the means of the pre- and post-test scores of the Peabody Picture Vocabulary Test. Thus, this instrument also seems to measure the change or improvement in vocabulary mental age scores within this population. However, the cause(s) of this improvement cannot be determined on the basis of this investigation.

Significant correlations were found between: (1) pre- and post-test scores of the Binet Vocabulary Test, (2) pre- and post-test scores of the Peabody Picture Vocabulary Test, (3) pre-test scores of the Binet Vocabulary Test and Peabody Picture Vocabulary Test, and

(4) post-test scores of the Binet Vocabulary Test and Peabody Picture Vocabulary Test. Thus, significant correlations would be expected among these variables within the parent population. These correlation coefficients were all high and positive. Thus, these tests seem to provide reliable test scores upon testing and retesting within this population. It might also be expected on the basis of this investigation that those children within this population who score high on one test also tend to score high on the other. The same holds true for low scorers.

Recommendations for Further Research

Several recommendations are in order for further research in this area. First, replication of this study utilizing a larger sample of children drawn from several Head Start programs throughout the state would provide for more adequate generalization of the results to the parent population. Second, replication of the study utilizing a control group who did not attend the Head Start program might aid in determining what effects the training program has on the vocabulary mental age scores obtained by culturally restricted children.

Third, validity studies on these two instruments need to be carried out within this population. If these tests are going to be used to assess the vocabulary ability of this child, it is necessary to determine as adequately as possible whether these instruments measure his vocabulary.

One important factor concerning the validity of the two tests is the aspect of cultural bias. One wonders if the culturally restricted child is somehow penalized by test items which are outside the realm

of his experiences. Thus, item analysis of these two instruments seems most appropriate.

The list of recommendations could extend endlessly. These are merely three such recommendations which appear to be important if more information is to be gained concerning the administration of these two instruments.

Many more studies concerning vocabulary measurement within this population are needed in order that the speech pathologist may adequately interpret the score of the culturally restricted child on standardized tests of vocabulary. It is hoped that his study provides a basis for this larger body of research.

BIBLIOGRAPHY

Allen, Robert; Haupt, Thomas; and Jones, Wayne R. "A Suggested Use
    and Non-use for the Peabody Picture Vocabulary Test with the
    Retarded Child." Psychological Reports, XV (October, 1964),
    pp. 421-422.

Anastasi, Anne. "Further Studies on the Memory Factor." Archives
    of Psychology, No. 142 (1932), pp. 1-60.

Atwell, C. R. "Comparison of Vocabulary Scores on the Stanford-Binet
    and the Revised Stanford-Binet." Journal of Educational Psy-
    chology, XXX (September, 1939), pp. 467-469.

Ausubel, David P. "Effects of Cultural Deprivation on Learning
    Patterns." Audiovisual Instruction, X (January, 1965), pp. 10-
    12.

Bereiter, Carl, and Engleman, Siegfreid. Teaching Disadvantaged Chil-
    dren in the Preschool. Englewood Cliffs, N. J.: Prentice Hall,
    1966.

Bernstein, Basil. "Social Class and Linguistic Development: A Theory
    of Social Learning." Education, Economy, and Society. Edited
    by A. H. Halsey, Jean Floud, and C. Arnold Anderson. New York:
    Free Press of Glencoe, 1961.

Budoff, Milton, and Purseglove, Eleanor M. "Peabody Picture Vocabu-
    lary Test: Performance of Institutionalized Mentally Retarded
    Adolescents." American Journal of Mental Deficiency, LXVII (May,
    1963), pp. 756-760.

Carlton, Theodore. "The Effect of Chronological Age on Revised
    Stanford-Binet Vocabulary Score at the Moron and Imbecile Levels,"
    Journal of Genetic Psychology, LXI (December, 1942), pp. 321-326.

Cove, Philip B., ed. Webster's Seventh New Collegiate Dictionary.
    7th ed. Springfield, Mass.: G. & C. Merriam Company, Publishers,
    1963.

Cuff, N. B. "Social Status and Vocabulary." Journal of Genetic Psy-
    chology, XLVI (March, 1935), pp. 226-229.

Cureton, Edward E. "Mental Age Equivalents for the Revised Stanford-
    Binet Vocabulary Test." Journal of Consulting Psychology, XVIII
    (October, 1954), pp. 381-383.

Dale, Edgar, and Reichert, Donald. Bibliography of Vocabulary Studies. Columbus: Ohio State University, 1957.

Dunn, Lloyd M. Expanded Manual for the Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service, Inc., 1965.

Dunn, Lloyd M. Manual for the Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service, Inc., 1959.

Dunn, Lloyd M., and Brooks, Sadye T. "Peabody Picture Vocabulary Test Performance of Educable Mentally Retarded Children." Training School Bulletin, LVII (August, 1960), pp. 35-40.

Dunn, Lloyd M., and Harley, Randall. "Comparability of the Peabody, Ammons, Van Alstyne, and Columbia Test Scores with Cerebral Palsy Children." Exceptional Children, XXVI (October, 1959), pp. 70-74.

Dunn, Lloyd M., and Hottel, John V. "Peabody Picture Vocabulary Test: Performance of Trainable Mentally Retarded Children." American Journal of Mental Deficiency, LXV (March, 1961), pp. 448-452.

Dunsdon, M. I., and Roberts, J. A. "A Study of the Performance of 2,000 Children on Four Vocabulary Tests: Norms, with Some Observations on the Relative Variability of Boys and Girls." British Journal of Statistical Psychology, X (May, 1957), pp. 1-16.

Dunsdon, M. I., and Roberts, J. A. "The Relation of the Terman-Merrill Vocabulary Test to Mental Age in a Sample of English Children." British Journal of Statistical Psychology, VI (November, 1953), pp. 61-70.

Edwards, Allen L. Statistical Methods for the Behavioral Sciences. New York: Rinehart & Company, Inc., 1954.

Eisenson, Jon, et al. The Psychology of Communication. New York: Appleton-Century-Crofts, 1963.

Fargo, George, et al. "Comparability of Group Television and Individual Administration of the Peabody Picture Vocabulary Test: Implications for Screening." Journal of Educational Psychology, LVIII (June, 1967), pp. 137-140.

Gage, Gerald, and Naumann, Theodore. "Correlation of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children." Journal of Educational Research, LVIII (July, 1965), pp. 466-468.

Gray, Giles W., and Wise, Claude M. The Bases of Speech. New York: Harper & Brothers, Publishers, 1959.

Grigsby, Olive John. "An Experimental Study of the Development of Concepts of Relationship in Preschool Children as Evidenced by

Their Expressive Ability." Journal of Experimental Education, I (December, 1932), pp. 144-162.

Johnson, Buford J. Child Psychology. Springfield, Ill.: Charles C. Thomas, 1932.

Johnson, Wendell; Darley, Frederic L.; and Spriestersbach, D. C. Diagnostic Methods in Speech Pathology. New York: Harper & Row, Publishers, 1952.

Kahn, Harris. "Evidence for Long-term Reliability of the PPVT with Adolescent and Young Adult Retardates." American Journal of Mental Deficiency, LXX (May, 1966), pp. 895-898.

Kelley, Truman L. Crossroads in the Mind of Man. Stanford University: Stanford University Press, 1928.

Kimbrell, Don. "Comparison of PPVT, FRVT, RS-B, and Academic Achievement Scores Among Institutionalized Educable Mental Retardates." Perceptual and Motor Skill, XXIII (December, 1966), p. 1178.

Levinson, Boris M. "Reevaluation of the Revised Stanford-Binet Scale, Form L Vocabulary as a Test of Intelligence for the Kindergarten and Primary School Child." Journal of Genetic Psychology, XCIII (September, 1958), pp. 237-248.

Li, Jerome C. R. Statistical Inference I. Ann Arbor: Edwards Brothers, Inc., 1964.

Lincoln, Edward. "The Reliability of the Stanford-Binet Scale and the Consistency of the Intelligence Quotients." Journal of Educational Psychology, XVII (December, 1927), pp. 621-626.

Louden, Mary. "Relative Difficulty of Vocabulary Lists in the Stanford-Binet Scale." Journal of Educational Research, XXVI (April, 1933), pp. 601-607.

Mahan, Harry, and Witmer, Louise. "A Note on the Stanford-Binet Vocabulary Test." Journal of Applied Psychology, XX (April, 1936), pp. 258-263.

May, Theodore, and Perry, Harold. "The Relationship Between the Stanford-Binet (Form L) Vocabulary and the Columbia Mental Maturity Scale in a Group of Mentally Retarded Children." American Journal of Mental Deficiency, LXII (September, 1957), pp. 330-333.

McCarthy, Dorothea. "Language Development in Children." Manual of Child Psychology. Edited by Leonard Carmichael. New York: Chapman and Hill, Limited, 1954.

McCarthy, Dorothea. The Language Development of the Preschool Child. Institute of Child Welfare Monograph Series, No. 4. Minneapolis: University of Minnesota Press, 1930.

McCarthy, James J., and Kirk, Samuel A. The Construction, Standardization and Statistical Characteristics of the Illinois Test of Psycholinguistic Abilities. n.p.: By the Authors, 1963.

McNemar, Quinn. The Revision of the Stanford-Binet Scale. Boston: Houghton Mifflin, 1942.

Mein, R. "Use of the Peabody Picture Vocabulary Test with Severely Subnormal Patients." American Journal of Mental Deficiency, LXVII (September, 1962), pp. 269-273.

Merrill, Maud A. "Mental Differences in Children Referred to the Psychological Clinic." Journal of Applied Psychology, X (December, 1926), pp. 470-486.

Milgram, Norman A., and Ozer, Mark N. "Peabody Picture Vocabulary Test Scores of Preschool Children." Psychological Reports, XX (July, 1967), pp. 777-784.

Moss, James, and Edmonds, Phyllis. "The Peabody Picture Vocabulary Test with English Children." British Journal of Educational Psychology, XXX (February, 1960), p. 82.

Myklebust, Helmer R. "Aphasia in Children--Language Development and Language Pathology." Handbook of Speech Pathology. Edited by Lee E. Travis. New York: Appleton-Century-Crofts, Inc., 1957.

Norris, Raymond C.; Hottel, John V.; and Brooks, Sadye T. "Comparability of Peabody Picture Vocabulary Test Scores Under Group and Individual Administration." Journal of Educational Psychology, LI (April, 1960), pp. 87-91.

Osgood, Charles E., et al. Contemporary Approaches to Cognition. Cambridge: Harvard University Press, 1957.

Schiller, B. "Verbal, Numerical and Spatial Abilities of Young Children." Archives of Psychology, No. 161 (1934), pp. 1-69.

Schneck, M. R. "The Measurement of Verbal and Numerical Abilities." Archives of Psychology, No. 107 (1929), pp. 1-49.

Shipe, Dorothy; Cromwell, Rue; and Dunn, Lloyd. "Responses of Emotionally Disturbed and Non-disturbed Retardates to PPVT Items of Human Versus Nonhuman Content." Journal of Consulting Psychology, XXX (October, 1966), pp. 439-443.

Smith, Madora E. An Investigation of the Development of the Sentence and the Extent of the Vocabulary of Young Children. University of Iowa Studies in Child Welfare, No. 5. Iowa City: University of Iowa Press, 1926.

Smith, Mary K. "Measurement of the Size of General English Vocabulary Through the Elementary Grades and High School." Genetic Psychology Monographs, XXIV (November, 1941), pp. 311-345.

Spache, George. "The Vocabulary Tests of the Revised Stanford-Binet as Independent Measures of Intelligence." Journal of Educational Research, XXXVI (March, 1943), pp. 512-516.

Spearman, Charles. The Abilities of Man. New York: Macmillian Company, 1927.

Steel, Robert G. D. and Torrie, James H. Principles and Procedures of Statistics. New York: McGraw-Hill Book Company, Inc., 1960.

Sunne, Dagne. "Comparison of White and Negro Children in Verbal and Non-verbal Tests." School and Society, IXX (July, 1924), pp. 469-472.

Tempero, Howard, and Ivanoff, John. "Effectiveness of the Peabody Picture Vocabulary Test with Seventh-grade Pupils." Journal of Educational Research, LVIII (May-June, 1965), pp. 412-415.

Templin, Mildred C. Certain Language Skills in Children: Their Development and Interrelations. Institute of Child Welfare Monograph Series, No. 26. Minneapolis: The University of Minnesota Press, 1957.

Terman, Lewis M. The Measurement of Intelligence. Boston: Houghton-Mifflin Company, 1916.

Terman, Lewis M. "The Vocabulary Test as a Measure of Intelligence." Journal of Educational Psychology, IX (October, 1918), pp. 452-466.

Terman, Lewis M., and Merrill, Maud A. Measuring Intelligence. Boston: Houghton Mifflin Company, 1937.

Terman, Lewis M., and Merrill, Maud A. Stanford-Binet Intelligence Scale: Manual for the Third Revision, Form L-M. Boston: Houghton Mifflin Company, 1960.

Throne, Frances; Kasper, Joseph C.; and Schulman, Jerome L. "Peabody Picture Vocabulary Test in Comparison with Other Intelligence Tests and an Achievement Test in a Group of Mentally Retarded Boys." Educational and Psychological Measurement, XXV (Summer, 1965), pp. 589-595.

Thurston, L. L. "Primary Mental Abilities." Psychometric Monographs, No. 1 (1938), pp. 1-121.

Tobias, Jack, and Gorelick, Jack. "The Validity for the Peabody Picture Vocabulary Test as a Measure of Intelligence of Retarded Adults." Training School Bulletin, LVIII (November, 1961), pp. 109-110.

U.S. Office of Economic Opportunity. Head Start Child Development Programs Manual of Policies and Instruction, September, 1967. Washington, D.C.: Government Printing Office, 1967.

Weeks, Richard W. "Effectiveness of the Peabody Picture Vocabulary Test with College Students." Journal of Educational Research, LVII (October, 1963), pp. 109-110.

West, Robert; Ansberry, Merle; and Carr, Anna. The Rehabilitation of Speech. New York: Harper and Row, Publishers, 1957.

Whipple, Guy M. Manual of Mental and Physical Test. Baltimore: Warwick and York, Inc., 1920.

APPENDIX A

TABLE V

VOCABULARY MENTAL AGE SCORES ON THE
BVT AND PPVT IN CONJUNCTION
WITH CHRONOLOGICAL AGE

Subject Number	CA	Vocabulary Mental Age Scores			
		Pre-BVT	Pst-BVT	Pre-PPVT	Pst-PPVT
	Z	X_1	X_2	Y_1	Y_2
1	48	42	54	42	44
2	59	72	84	65	71
3	54	45	48	82	75
4	57	39	45	51	53
5	51	45	48	53	55
6	56	60	60	43	61
7	48	42	42	39	39
8	59	42	54	35	44
9	59	39	60	55	69
10	57	42	54	45	58
11	49	39	45	42	45
12	67	60	54	44	49
13	67	48	54	34	41
14	44	39	39	35	48
15	58	39	54	42	49
16	49	42	54	46	59
17	44	33	39	38	47
18	48	36	30	39	42
19	59	48	60	47	71
20	51	42	48	42	43
21	43	36	48	39	43
22	43	33	36	32	32
23	56	39	42	34	46
24	46	33	42	38	49
25	48	42	48	58	59
26	54	42	48	45	55
27	52	27	27	24	26
28	48	36	42	43	47
29	42	27	36	30	32
30	46	39	60	32	39
31	52	45	48	48	58
32	43	33	33	31	34
33	53	39	39	45	48
34	43	30	33	27	27
35	56	36	42	40	51
36	55	72	84	61	78
37	45	33	39	27	36
38	57	45	42	44	46
39	49	42	45	49	49
n=39	ΣZ=2015	ΣX_1=1623	ΣX_2=1860	ΣY_1=1666	ΣY_2=1918
	$\bar{z}$=51.6667	$\bar{x}_1$=41.6153	$\bar{x}_2$=47.6923	$\bar{y}_1$=42.7180	$\bar{y}_2$=49.1795
	s=6.4902	s_1=9.9541	s_2=12.0157	s_1=11.2249	s_2=12.5087

APPENDIX B

TABLE VI

ELIGIBILITY GUIDELINES FOR THE HEAD START
CHILD DEVELOPMENT PROGRAM

Family Size	Non-Farm	Farm
1	$1,600	$1,100
2	2,000	1,400
3	2,500	1,700
4	3,200	2,200
5	3,800	2,600
6	4,200	3,000
7	4,700	3,300
8	5,300	3,700
9	5,800	4,000
10	6,300	4,400
11	6,800	4,700
12	7,300	5,100
13	7,800	5,400

Source: U.S., Office of Economic Opportunity, Head Start
Child Development Programs Manual of Policies and
Instruction, September, 1967 (Washington, D.C.:
Government Printing Office, 1967), p. 7.

VITA

Kay Hand Mount

Candidate for the Degree of

Master of Arts

Thesis: AN INVESTIGATION OF THE RELATIONSHIPS BETWEEN THE BINET
        VOCABULARY AND PEABODY PICTURE VOCABULARY TESTS WHEN
        ADMINISTERED WITHIN A PRESCHOOL HEAD START POPULATION

Major Field: Speech

Biographical:

    Personal Data: Born in Okmulgee, Oklahoma, July 11, 1945, the
        daughter of Mr. and Mrs. Forrest S. Hand. Married Kimmie
        T. Mount, June 4, 1965.

    Education: Graduated from Beggs High School, Beggs, Oklahoma,
        in May, 1963; received the Bachelor of Science degree from
        Oklahoma State University in May, 1967, as a Senior Year
        Trainee under an Office of Education Fellowship, with a
        major in Speech Pathology; completed requirements for the
        Master of Arts degree at Oklahoma State University in May,
        1968, under an Office of Education Fellowship.

Honorary Organizations: Member of the Honor Society of Phi Kappa
        Phi.