A concept usage in selected preschool children.

Ernest Milton Green
University of Massachusetts Amherst

Follow this and additional works at: http://scholarworks.umass.edu/dissertations_1

Recommended Citation
http://scholarworks.umass.edu/dissertations_1/2852

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
A CONCEPT USAGE IN SELECTED PRESCHOOL CHILDREN

A Dissertation Presented
by
Ernest Milton Green

Submitted to the Graduate School of the University of Massachusetts at Amherst in partial fulfillment of the requirements for the degree of

DOCTOR OF EDUCATION

April 1974

Major Subject: Education
CONCEPT USAGE IN SELECTED PRESCHOOL CHILDREN

A Dissertation

By

Ernest Milton Green

Approved as to style and content by:

Ernest D. Washington, Chairman

David G. Coffing, Member

Alan J. Gross, Member

Juan P. Cabán, Member

Dwight W. Allen, Dean of Education

April, 1974
PURPOSE OF THE STUDY

This study assumed that where standardized norm-referenced general verbal ability tests may be considered unsuitable for verbal assessment in early education, such an assessment could be made through concept usage assessment with categorical verbal norms.

PROCEDURE

Twenty-five selected preschool children 36 to 56 months old with a Peabody Picture Vocabulary Test I.Q. group mean of 87.72 responded to 20 category descriptions with verbal instances of a category selected from the Connecticut Verbal Norms Battig and Montague (1969). Responses falling within the proper categories were tallied for an index of concept recognition. Each such response was assigned the equivalent frequency rank of the response found in the table for an index of general usage. With the Peabody Picture Vocabulary Test I.Q. and M.A. scores criterions, concept recognition scores and general usage scores were submitted to regression analysis.

RESULTS

Analysis showed that there was a high positive correlation of
performance scores on the Peabody Picture Vocabulary Test with scores of concept recognition and general usage such that as performance approached the highest possible frequency level in general usage, the higher the I.Q. and M.A. The combined criterion prediction of M.A. explained 58% of the variance accounted for with a standard error of estimate of 9.25 significant at the .001 level.

REFERENCE

TABLE OF CONTENTS

CHAPTER

I. STATEMENT OF THE PROBLEM ........................................... 1

   Introduction
   Purpose of the Study
   Hypotheses
   Limitations

II. METHODOLOGICAL CONSIDERATIONS ............................... 12

   Assessment

III. RESEARCH DESIGN .................................................. 25

   Subjects
   Procedure
   Experimental Design
   Evaluation of Concept Usage
   Variables
   Summary of Main Variables

IV. RESULTS ............................................................... 33

V. DISCUSSION AND CONCLUSIONS ................................... 38

APPENDIX ................................................................. 50

BIBLIOGRAPHY ............................................................ 56
LIST OF TABLES

TABLE

1  Stepwise Regression Prediction of I.Q. Criterion from Population of all CVN Predictor Variables ........ 33

2  Stepwise Regression Prediction of M.A. Criterion from Population of all CVN Predictor Variables ........ 34

3  Stepwise Regression Prediction of I.Q. Criterion from Population of CVN and Associative Learning Predictor Variables ........ 35

4  Residual Table for Stepwise Regression Prediction of I.Q. Criterion from Population of CVN and Associative Learning Predictor Variables ........ 36
CHAPTER I

STATEMENT OF THE PROBLEM

Introduction

As a measure of verbal performance the nature of norm-referenced standardized test construction often precludes its use as an objective standard for facilitating verbal performance characteristic of its norming population. The question posed for the present consideration is whether it is feasible that an objective criterion can be established which can serve as a basis for facilitating verbal performance in selected preschool children characteristic of mean performance on norm-referenced standardized verbal tests.

The problem of verbal assessment in preschool education is but one related to a much larger problem common to current psychometric practice in the assessment of general intelligence. Specifically, the problem arises when a norm-referenced standardized general intelligence measure is considered for defining educational objectives in terms of its own standard.

Norm-referenced tests in general by the nature of their construction are used effectively as predictive criteria without always having the benefit of a substantive reference that explains the cumulative abilities underlying the performance characteristic of their norming populations. Nevertheless, they are used as comparative and predictive measures in educational practice. Individual test scores may be compared with
standardized norms. Relative to standardized norms, such scores often to varying degrees predict future scholastic achievement better than would be expected by chance alone. That they do so, often underlies the impetus for facilitating performance that is characteristic of mean performance or better on norm-referenced standardized general intelligence measures. However, standardized norms are not designed as objective goals for performance, but are designed to serve as a reference by which to compare performance on their measures.

Providing intellectual experience to facilitate intellectual attainment characteristic of mean performance or better on norm-referenced standardized tests, if such is desired, is extremely difficult without objective criteria. In view of numerous plausible explanations for low verbal I.Q. scores during the last decade, preschool education is faced with problems. In the absence of substantive criteria, problems in facilitating and evaluating verbal attainment in preschool children appear to be even greater. Considering the variety of stated objectives reviewed by Evans (1971), it appears that ultimate evaluation of verbal development across language programs during this period was based upon performance of children on norm-referenced standardized tests. Where the norming population on such standardized tests is not representative of the population being tested, norms are of little value in interpreting performance on such measures. If such were the case, there has been little basis for objective evaluation of verbal development programs in preschool education.

Verbal intelligence test norms are meaningful when they are representative of the population for whom they are designed to reference.
A low performance by a sample population relative to such norms suggests characteristic differences between the norming population and the sample. If this is so, an alternative assessment is needed to examine the differences. Such an assessment requires an objective measure with an objective standard which by its structure provides answers to how its criterion performance is achieved relative to performance on norm-referenced standardized tests.

If the alternative verbal assessment is to be made verbally, several equally plausible explanations to account for low performance on standardized norm-referenced measures can be made in terms of the experience of children. For example, beside being unfamiliar with a particular verbal stimulus or the concept it represents, some children may be familiar with the stimulus and not be with the concept as presented by the task in terms of their experience. Also, children may be familiar with both the concept and stimulus, but have yet to make the verbal association presented by the task in terms of their experience. There is also the possibility that children may be familiar with the concept and have made other verbal associations with the concept in terms of their experience, but not the one presented as a verbal stimulus.

Three implications follow from the preceding analysis that suggest determining characteristics of a functionally objective alternative verbal assessment. First, the measure must be both verbal and conceptual if it is to assess the possible explanations for the low performance on standardized measures cited. Second, its structure must be comprehensive if it is to provide for maximum variability in experience. Third, it must be normed if it is to provide answers to how criterion
performance is achieved by individuals assessed in its own terms.

Characteristically, for an objective alternative verbal assessment to be functionally useful for facilitating preschool verbal development in its own terms suggests that it should constitute broad objective patterns of classificatory concepts representative of general usage if it is to assess conceptual as well as verbal experience. Basically, it must be broad enough to include mature conceptualizations that can develop from early childhood through maturity as a result of environmental and cultural exposure if it is to assess the full potential range of verbal developmental possibilities. Concepts constituting the pattern should be comprehensive enough in their internal structure to allow for the greatest individual variance in environmental and cultural exposure that can be normed by general usage.

Where the norm-referenced standardized intelligence measure cannot provide an objective criteria for verbal development, experimental verbal norms provide a possibility. In contrast to the standardized measure's statistical necessity of relying on frequency of general usage to achieve stability, experimental norms such as the Connecticut Verbal Norms, Battig and Montague (1969), reflect the entire spectrum of verbal behavior patterns characteristic of the basic requirements of an objective verbal assessment measure. Consequently, in terms of the magnitude of the problem, the possibility of adapting these norms as a representative alternative measure to assess verbal performance should merit investigation as a preliminary step in the determination of the feasibility of establishing an objective basis for preschool verbal development.
Purpose of the Study

As a preliminary step in ascertaining the feasibility of establishing an objective verbal standard, the present consideration of concept usage in a group of selected preschool children is an exploratory investigation into possible relationships of current standardized verbal measures with experimental measures of concept usage and learning ability. Primarily a correlation study, the exploration is principally concerned with research possibilities that could ultimately provide a substantive basis for the development of verbal attainment in selected preschool children with questionable verbal performance on standardized measures through the use of normed verbal criteria which is characteristic of mean performance or better on such measures. Providing both impetus and subjects for the preliminary investigation is an ostensibly verbally proficient group of three to four year old preschool children with a mean verbal performance approaching one standard deviation below the mean on the Peabody Picture Vocabulary Test (PPVT), Dunn (1965).

The general purpose of the study is to examine with representative measures four principal relationships thought to be relevant to the feasibility of establishing an objective standard for verbal development in selected preschool children. The first involves the relationship between I.Q. derived from the PPVT and experimental measures of concept usage derived from the Connecticut Verbal Norms (CVN), Battig and Montague (1969). The second involves the inter-relationship of learning ability defined as acquisition of new knowledge by Rohwer (1967) and
concept usage in the prediction of standardized verbal I.Q. The third involves the relationship between learning ability after Rohwer and chronological age. The fourth involves the relationship between learning ability after Rohwer and PPVT I.Q.

Specifically, the purpose of the study is to investigate individual verbal differences in selected preschool children in an attempt to examine broader patterns of concept usage underlying verbal performance relative to I.Q. through the use of normed verbal-conceptual criteria. To facilitate verbal performance representative of mean performance on standardized measures requires an initial identification and specification of the relationships between a representative standardized measure and a representative objective alternative measure relative to the children for whom the facilitation is intended. By identification of the relationships between the two measures in terms of general usage, the feasibility of facilitating representative performance on standardized measures can be seen in terms of specific relationships found to exist between the two measures relative to the general usage of the children.

In the investigation of individual differences, the objective of the use of normed criteria is to attempt to identify characteristics or combinations of characteristics of concept usage which may lead to predictions of I.Q. in terms of general usage on a verbal level. If the alternative measure used in investigating the characteristics of concept usage is a substantively objective norm of general usage, the feasibility of facilitating representative performance on standardized measures can be seen in terms of characteristics reflected in the mean performance of the children on both measures. If this is so, the characteristics should
combine to predict I.Q. derived from the standardized norm-referenced measure.

The objective of using the Connecticut Verbal Norms is that its structure provides the possibility of independent investigations of individual characteristics of conceptual experience in terms of verbal concept recognition and characteristics of verbal experience in terms of normed general usage familiarity as they relate to mean performance on standardized measures of verbal ability. By identifying these characteristics in terms of general usage, the feasibility of facilitating representative performance on standardized measures can be seen in terms of the specific characteristics identified that could be used for such facilitation.

The primary research implication of ultimately providing a substantive basis for the development of verbal attainment in preschool children of questionable verbal performance is to examine the assumption that there is a relationship between goals for verbal performance and mean performance on standardized measures. This assumption suggests a preliminary analysis of the relationship and a test of its implications for determining the feasibility of establishing an objective criteria for verbal development in selected preschool children.

While there may be disagreement on precisely what the ultimate goal of verbal attainment is, most individuals function at some level of verbal attainment relative to the common usage of their environment. The basis and criteria for its development is the ability to function verbally in terms of their experience within their environment and in terms of the common usage of that environment, Brown (1958). If this is
so, as Brown suggests, there are only individual levels of verbal attainment in terms of common usage relative to the unique experience of an individual and his environment for which there is no other objective measurement. Consequently, common usage can be viewed as the ultimate goal of verbal attainment.

One generalization that can be made about the relationship between mean performance on standardized measures and normed verbal usage is that the individuals standardizing arbitrary test items are related to both. They share common usage and verbal experience on the particular items constituting the test. Subsequent mean performance on the standardized test items by other individuals would appear to be coincidental to the same common usage and experience. If this is correct, the commonalities of experience should be reflected in normed general usage. Where this commonality of experience does not exist on specific test items, the variability in response to these items should be reflected on both the standardized norm-referenced measure and a measure of normed general usage, but in different ways. Variability in response to the standardized measure reflects differences in common usage and experience as relatively lower scores than the normed mean scores across individuals constituting the norming group standardizing the measure. Variability in response on normed general usage reflect degrees of individual verbal attainment relative to experience and normed general usage. If this is so, through these relationships the feasibility of facilitating mean performance on standardized measures of verbal ability can be seen in terms of differences in experience reflected both on the standardized measure and normed general usage in terms of individual verbal attainment.
Hypothesis I: There is a statistically significant positive correlation between I.Q. scores derived from the PPVT and concept usage scores derived from the CVN in selected preschool children.

This hypothesis follows from the assumption that adult category verbal norms reflect variability in concept usage and general usage relative to such norms as a criterion and relative to norm-referenced standardized tests of general verbal ability to examine the relationship in the performance of a selected group of preschool children on representative measures.

Hypothesis II: There is a statistically significant positive correlation between M.A. scores derived from the PPVT and concept usage scores derived from the CVN in selected preschool children.

This hypothesis follows from the assumption that adult category verbal norms as a criterion is developmental relative to norm-referenced standardized tests of general verbal ability to examine the relationship in the performance of a selected group of preschool children on representative measures.

Hypothesis III: Learning ability after Rohwer and concept usage scores combine in a statistically significant predictive equation to account for some variation in I.Q. scores derived from the PPVT in the selected group of preschool children.

This hypothesis follows from the assumption that adult category verbal norms as a criterion is developmental relative to norm-referenced standardized tests of general verbal ability to examine this relationship relative to associative learning under experimental control in the performance of a selected group of preschool children.
**Hypothesis IV:** There is a statistically significant positive correlation between learning ability scores after Rhower and chronological age in selected preschool children.

This hypothesis follows from the assumption that associative learning is developmental to examine the relationship in the performance of a selected group of preschool children.

**Hypothesis V:** Learning ability scores after Rohwer and PPVT I.Q. scores are not significantly correlated in selected preschool children.

This hypothesis follows from the assumption that associative learning is less affected by differences in past opportunities for learning than are standardized intelligence tests to examine the relationship in the performance of a selected group of preschool children on representative measures relative to verbal ability.

**Limitations:** Systematic manipulations of multivariate natural human psychological behavior variables in experimental educational research has not produced much useful data in the areas of learning and teaching. Snow (1973) in his review of the literature suggests that in the attempt to control for factors jeopardizing the validity of experimental research design the experimenter has often acted as if the subject is an experimental tool. The experimenter often forgets that the subject is an intuitive organism with many objectives and options who is capable of reacting to an experimental stimulus in many alternative ways. Data, in Snow's terms, thus derived are molecular at most under experimental conditions and possibly the least representative of natural conditions. Further, its interpretation is difficult and of little use for generalization due to the range of options the subject can
exercise in coping with equivalent stimulus in a natural environment.

Snow further suggests that if useful information is to be gained from research efforts, representative experimental designs with descriptive generality should be given first priority in educational research. The present study is such an attempt. Consequently, it is limited to examining and describing relationships between a standardized measure and experimental measures providing the children studied the opportunity to cope with the tasks in their own way in a familiar environment under normal conditions. No attempt to generalize the subsequent findings is implied in this exploratory effort. However, the measures and methodology, where the data warrant, are exemplary of the possibilities for further research should it be deemed practical or desirable.
CHAPTER II

METHODOLOGICAL CONSIDERATIONS

Assessment

Concept usage here is defined by the tasks providing data in the assessment of verbal ability through the use of adult category verbal norms. Structurally, the categorical norms used are classificatory. The child is simply asked to respond with a single free recall response to a verbal description of a category, "clothing" for example, with an instance the description evokes. While little is known about the subjective processes of recall, Kintsch (1970) suggests that the subject is nevertheless capable of taking advantage of the structure presented in such a task by whatever subjective means that are available to him. Thus, the response elicited may be isomorphically interpreted as being derived by word association or conceptually derived. The distinction is not critical. Deese (1965) argues:

What is critical to the use of associations in the assessment of understanding, however, is the fact that they do predict linguistic and more generally, conceptual usage. (p. 172)

Structurally, categorical norms can provide a quasi-objective basis for the study of the relationships between standardized test of verbal I.Q. and general usage. Precisely which verbal norms are most appropriate for assessment of verbal performance in preschool children must await further investigation. However, categories provide a unit of measure of equivalent responses of which the standard test item may be
but a single instance. Consequently, the primary consideration in the use of categories is that they provide for maximum variability and freedom in response on a concept recognition task designed to provide quantitative and qualitative data for estimating verbal performance.

In their present use, categories can be viewed as developmental in structure in terms of quantitative and qualitative variability. Quantitatively, a measure of the number of instances of class recognition can be obtained. Subsequently, the qualitative level of usage in the responses obtained can be indexed on a developmental continuum by norms representative of general usage. The critical notion here is that category complexity establishes a developmental continuum such that a few instances of class inclusion can be expected to be recognized by young children through limited experience and exposure relative to the number that can be expected of adults. At the same time, some overlap in child and adult concept usage can be expected. Brown (1958). As Brown describes word acquisition in the very young child, concepts may be developmental from concrete to abstract as well as from abstract to concrete as a result of the variability in adult naming practices influence on the verbal-conceptual development in young children.

The implications of Brown's analysis of the effect of adult naming practices on children's word acquisition is an overriding consideration in the use of adult category norms. Brown, citing examples of the tendency of young children to over generalize their classifications of things and people, argues that the sequence in which children's words are acquired is determined more by the naming practices of adults than by the cognitive preferences of children.
The name given a thing by an adult for a child is determined by the frequency with which various names have been applied to such things in the experience of the particular adult. General referent-name counts taken from many people will predict much that the individual does, but for a close prediction, counts specific to the individual would be needed. (p. 313)

In conjunction with a developmental notion of the structure of normed category response, Brown's analysis of parental intervention in children's word acquisition suggests consideration of the use of adult category norms as a representative criteria of general usage. While concept usage may be discontinuous relative to adults and children, in terms of experience and development, categorical norms can index both in terms of word availability in free association recall. Brown is explicit in the implication and hypothesis that the parent's influence on the child's word acquisition is in conformity with the adult common general usage.

The names provided by parents for children anticipate the functional structure of the child's world. This is not, of course, something parents are aware of doing. Each thing has its own name, just one, and that is what we give to a child. The one name is, of course, simply the usual name for us. Naming each thing in accordance with local frequencies, parents unwittingly transmit their own cognitive structures. (p. 318)

Class recognition tasks for obtaining data in the assessment of verbal ability follows from the methods of investigating organizational processes in recall. It has been long held that free associations help recover material lost in memory as witnessed by its application in psychoanalysis. It is also well known that there is a high degree of commonality in responses given to a stimulus word when a large group of subjects is asked to give a free association word response (Jenkins and Russell, 1952). These factors have produced a variety of applications
of free association and methods of investigating organizational processes in recall which in turn has produced a variety of normative data.

Normative data on word association have provided the basis for most of the investigation on the effects of associative structure in recall. Kent and Rosanoff (1910) provided the initial normed frequency response count on 100 familiar English words upon which most of the early investigations were based. Subsequently, Palermo and Jenkins (1964) have provided extensive word association norms used by most recent workers. Deese (1965) in the meantime has provided 18 tables of relative frequency with which words tend to elicit each other in free association tests. Prior to 1969 limited normative data on categorical relationships in the investigation of clustering had been provided by Cohen, Bousfield, and Whitmarsh (1957).

The extensive verbal norms alluded to by Brown (1958) had to await the development of the study of categorical relationships in learning for their ultimate existence. Bousfield (1953) studied the recall of subjects who learned a 60 word list comprising 15 instances of 4 conceptual categories. The results demonstrated that the subjects were grouping their responses according to conceptual categories. Bousfield had selected the words for his categories according to his best judgment. Most later investigators have made use of the Connecticut Category Norms compiled by Cohen, Bousfield, and Whitmarsh (1957) to investigate category clustering in free recall learning and in memory. Cohen, et al compiled 4 responses from 400 university students in 43 categories.
Battig and Montague (1969) replicated and extended the Connecticut norms to 56 categories. Each category has a ranked frequency distribution of 26 or more responses occurring ten or more times. In addition, for each response the mean rank of that response, as it occurred on the lists eliciting it, is given. Thus, the Category Norms For Verbal Items in 56 Categories provides normative data for the total frequency distribution of responses to a category and mean frequency distributions among the responses elicited.

Consideration of specific tasks in the use of the Connecticut verbal norms follow from the area of study for which the norms were developed. The class recognition task here follows in principle from Tulving and Pearlston (1966) who found that if categorized lists are given to two different groups the one given the category names will recall more words than the other group not given category names. Based on their findings the application here is: given a category name, the children can respond with an instance of that category from memory if a relationship can be established and recognized as such. If no response is elicited, the child is given the most frequent normed response from the category as an example of an instance and is asked to respond with another instance. From these tasks samples of concept usage and learning behavior are obtained.

Rohwer (1967) addressed the problem of distinguishing familial retardation from unequal learning opportunities as measured by performance test of intelligence or school achievement. Rohwer's major underlying assumption in making such a distinction was to test subjects on a learning task primarily requiring acquisition of new knowledge not dependent upon
previously acquired knowledge to measure their learning efficiency.

Using an experimental designed pictoral paired-associate task, Rohwer assessed three sample populations of known mean mental age (MA) under four conditions assumed to facilitate paired-associate learning. Two samples consisted of children. The third sample consisted of adult institutionalized retardates with no known organic defects. Performance on the task by kindergarten children with a mean MA of 4.7 and third grade children with a mean MA of 8.6 was compared to the performance of the institutionalized retardates with a mean MA of 9.6. The results of the assessment found that learning was more efficient in each of the children samples than in the sample of institutionalized retardates.

Rohwer suggests the following interpretation of the results:

Among Ss who are classified accurately as familially retarded, even optimal conditions of learning, as represented by the PA task used here, are not sufficient to improve performance to the level of that observed in equal- or lower-MA normals. In contrast, under these same conditions, the performance of lower-strata children, inaccurately classified as slow learners on the basis of standardized test performance, belies the assumption that they cannot be proficient learners. (p. 80)

In a similar subsequent study, Rohwer (1967) examined the relationship of the pictorial PA task to an intelligence test with preschool children’s performance. Performance of children varying in ages from 36 to 65 months was assessed on the experimental PA task and on the PPVT to estimate the degree of relationship between their performance on the two measures. Two populations of cooperatively operated preschools were samples. Mental age and I.Q. were found to predict learning efficiency in one sample, but not in the other. The interaction of sample with verbalization was significant, favoring
the sample for which MA and I.Q. predict learning efficiency.

The relevance of the Rohwer studies to methodological considerations in the present study is two-fold. First, he has demonstrated that the PA task as a diagnostic instrument is less affected by differences in past opportunities for learning than are standard intelligence tests. Consequently, an alternative means of investigating the learning processes in preschool children has been provided. Its use in the present study is an attempt to assess learning ability as an alternative measure of ability inferred by standard intelligence tests. Second, the data suggest that differences found between the preschool samples is verbal. While less affected by differences in past opportunities for learning than standard intelligence tests, the PA task appears to discriminate verbal ability. If this is correct, the implications of the PA task that discriminates verbal ability should be considered in the design of alternative measures of verbal ability.

Kintsch (1970) suggests that in a paired-associate task the subject must learn two distinct things. He must learn what the response terms are. Then he must learn to associate each stimulus term with the response term. If the experimenter employs unfamiliar words, response learning would be quite difficult. This would be a problem of response integration. However, when familiar words are the response terms, response integration is less a problem. Where the response term is a familiar word, it is assumed to be already available in the subject's memory. If this is so, variability in verbal ability among samples of very young children's performance on a PA task may confound the interpretation of the learning efficiency construct.
One implication of variability in verbal ability in assessing the verbal performance of very young children is the method of presentation. Martin (1967) and Bernbach (1967) have demonstrated that recognition of the stimulus term is a necessary condition for establishing stimulus-response connections. Martin's subjects learned 8 trigram-number pairs. Subjects were required to make two responses. First, they told whether they had recognized the stimulus. Then, they gave the stimulus response. Where the recognition response was correct, Martin found that the probability of recall increased over trials as a learning curve should. In contrast, where the recognition response was incorrect, no learning occurred at all. Bernbach's data in a similar study support these findings. Moreover, both Bernbach and Martin found that regardless of how often the subject has previously given the correct recall response to a particular item, if that item is not recognized on a given trial, recall performance is no better than chance. It remains to explicate the implication of these findings to methodological considerations in the presentation of verbal tasks in the assessment of verbal performance in very young children.

Task presentation mode may have an effect on verbal variability in assessing learning proficiency in preschool children. While trigram-number pairs and pictorial pairs in paired-associate learning cannot be equated, the effect of verbal variability in very young children on the sentence elaborated pictorial PA task is analogous to a dichotomy of these two modes of presentation. The very young child without readily available words in memory for the pictorial stimulus and response terms is confronted with a problem dissimilar to the problem confronting the
child with words readily available in memory for the stimulus and response terms in such a task. In addition to response learning and response integration, the subject without readily available words in memory for the stimulus and response terms is faced with the problems similar to those present in the trigram-number PA learning task. Under such circumstances, a subject's performance in terms of learning efficiency cannot be equated with the performance of other subjects' performance on dissimilar tasks due to initial differences in verbal ability.

It is imperative to methodological considerations in assessing verbal ability to minimize the possibility of initial individual differences in establishing a reliable index of ability where possible. To minimize such a possibility, as Snow (1973) suggests, care must be taken to provide for the exercise of individual options in stimulus recognition and recall response in the design of assessment tasks for preschool children. This consideration is of singular importance where language differences may be suspected to exist as suggested by Cole and Bruner (1967); Labov (1970); Labov and Cohen (1967), among others. To make sociological inferences from empirical data, as does Rohwer to "strata", serves no empirical purpose except possibly to distract from the chance that the task assumed to measure learning efficiency may be a better index of verbal ability than of learning ability.

The review above of Rohwer's findings suggests indexing verbal variability while assessing concept usage as an alternative measure of learning. The strength of the PA task as an alternative measure of learning ability, Rohwer assumes, is that the assumption of
equivalent previous opportunities for learning is unnecessary. However, considering the possible effects of variability in verbal ability in view of Kintsch's (1970) analysis of the PA task together with the findings of Martin (1967) and Bernbach (1967) on the prerequisite conditions for PA learning, raises some questions about such an assumption in a sentence elaborated PA learning task. It would appear that possibly the cumulative effect of verbal ability may not be adequately controlled in such a PA task presentation mode. Further, the finding of significant differences in the main effect and interaction effect between preschool sample and I.Q. suggests the possibility that they may be explained in terms of cumulative verbal ability which is generally assumed to be the best single predictor of I.Q., Dale and Reichert, (1957).

As an alternative to the PA assessment of learning ability that possibly indexes variability in verbal ability indirectly, the present study attempts to provide a criterion of general usage to index variability in verbal ability while assessing concept usage as a measure of learning capacity. By definition, the cumulative effect of learning results in a cumulative capacity to learn. The cumulative effect of learning is implicit in discrimination, for example, as a capacity relative to generalization and concept formation. The consideration here is that given a criterion relative to a suitable verbal learning task, the cumulative effect of verbal ability as well as variability in verbal ability could reflect in the examination of the relationships under study.
Ebel (1972) concerned with the limitations of criterion-referenced measures in the pursuit of excellence in educational achievement holds that:

The essential difference between norm-referenced and criterion-referenced measurements is in the quantitative scales used to express how much the individual can do. In norm-referenced measurement the scale is usually anchored in the middle on some average level of performance for a particular group of individuals. The units on the scale are usually a function of the distribution of performances above and below the average level. In criterion-referenced measurement the scale is usually anchored at the extremities, . . . (p. 144)

Popham and Husek (1969) concerned with some of the implications of criterion-reference measures view the comparative basis of individual performance as the basic distinction between normed-referenced measures and criterion-referenced measures. In their view the basic distinction lies in the decisions to be made about individual performance and the implications of such purposes on test design.

Most standardized tests of achievement or intellectual ability, represented here by the Peabody Picture Vocabulary Test, can be classified as norm-referenced measures by the foregoing distinctions. They compare individual performance on the same measurement device compared to a normative group. A criterion-referenced test, represented here by a measure derived from the Connecticut Verbal Norms, assesses individual performance with respect to these adult norms as a performance standard. While the experimental measure is normed-referenced, the critical distinction is that it serves to index what the individual does verbally, not how he stands with others.

The implication in the purpose of a test and its design provides a pivotal distinction between norm-referenced and criterion-referenced
tests. Generally, the purpose of the normed-referenced test is to make decisions about individuals. It selects on the basis of comparative performance. Consequently, the test is designed to spread individuals out. In contrast, the criterion-referenced test is generally for the purpose of making decisions about individuals and treatments relative to a performance standard.

Purposes and uses of normed-referenced or criterion-referenced tests are not mutually exclusive, but their design is distinctive. Commonality of response, discussed in the previous chapter, in test item selection is relative to normed-referenced test stability, but the variability in the scores these selected items produce on a test is at the core of the basic difference between normed-referenced and criterion-referenced test design. For the normed-referenced test, the more variability in the scores the better. With criterion-referenced tests, variability is irrelevant. Popham and Huresk (1969) suggest that the normed-referenced test writer in promoting variability disdains items that are too easy or too hard and tries to increase the allure of wrong answer options. In contrast, whether an item is difficult or easy; discriminating or indiscriminate, the criterion-referenced test item writer's goal is an accurate reflection of the criterion behavior. The critical notion here is that while normed the CVN is the criterion behavior that is construed to provide the test items for the experimental measure and the standard of performance for all verbal measures in the present study.

In summary, the implication of criterion-referenced performance here, relative to the performance on normed-referenced verbal
intelligence tests, is that the criterion provides a means to examine the relationship between performance on each type of sampling device when examining the relationship between the normed-referenced measure and the criterion. An examination is attempted of the inter-relationship between the variability on a normed-referenced test and the variability on a criterion-referenced test in the performance of selected preschool children.
CHAPTER III

RESEARCH DESIGN

Subjects

Twenty-five children age 36 through 56 months attending a low income housing development preschool in Springfield, Massachusetts served as subjects. Their demographic and verbal characteristics are presented in Appendix I.

Procedure

First, all subjects were administered the Peabody Picture Vocabulary Test (PPVT) Dunn (1965). Next, they were administered a study developed verbal prompting test. Then they were given a learning test after Rohwer (1967). All tests were individually administered over a three-month period during the fall of 1973. The tests were administered during scheduled learning sessions before noon. All tests were administered by the same administrator. The administrator was known by the subjects as a member of the preschool staff. Tests were touted as games. Token rewards of individually wrapped bubble gum and candy were given to the subjects to take home after each test participation.

Testing was conducted in a quiet staff lounge adjacent to the class rooms. The testing arrangement situated the subject and the test administrator at eye-level. The temperature in the room where the tests
were administered was comfortable. The room was without windows. Lighting was bright. The door was left ajar. Parents and staff could observe the testing.

The order of testing the children followed the school's roster in general. However, each of the children was asked if they would like to play the special games in the lounge by their teacher before they were tested. Few, if any, declined the opportunity. Occasionally, if a child had a "cold" or was unsociable at time the child was to be tested, the child's teacher would wait to ask the child. That is to say, there was an effort by the staff to have the children tested under the most favorable conditions possible. All tests were administered in fifteen minutes or less. There was ample time for friendly conversation between the subject and the administrator before and after testing.

Experimental Design

The study is a single sample multiple-regression design with 12 variables. Criterion prediction is by step-wise multiple linear regression with concomitant use of analysis of variance.

Evaluation of Concept Usage

Three types of ability measures were involved in the study comprised of PPVT measures, study developed measures and an associative learning
measure after Rohwer (1967).

The PPVT was selected as a representative standardized measurement of verbal ability on the basis of its measurement validity, administrative and normed technical excellence evaluations provided by Hoepfner, Stern, and Nummedal (1971). It was constructed and standardized on 300 out of an initial pool of 3,885 words found in the dictionary the meaning of which could be illustrated by line drawings. These were illustrated and age-leveled where 40 to 60 percent of the appropriate group passed. Categories for the plates included in the test comprise: man-made objects, animals, birds, human actions, nature scenes, plants, flowers, inanimate objects, adverbs, articles in a home, adjectives, musical instruments, occupations, scientific materials, parts of a house, and wearing apparel.

Criteria for selection of a word to be used in making up the test were: (1) all words were found at the leveling and pre-test stages to be of the same level of difficulty and (2) all words demonstrated good linear growth curves in terms of percent passing at successive age levels.

The study developed instrument presented in Appendix III was constructed to sample concept usage, verbal usage and concept learning. Primarily verbal prompting tasks, adult Connecticut Verbal Norms serve as a representative criterion reference for these measures and as an index of general usage. If a subject's response on any of the measures is equivalent to a response within the CVN stimulus category, it is scored as a criterion equivalent response as a measure of concept usage. Then the response is indexed with the equivalent frequency rank of the
normed response as a measure of general usage.

Twenty categories from the Connecticut Verbal Norms were used as concept class recognition test items. A description of each category used in the test was designed and pre-tested heuristically on a similar population at the school where the present study was conducted. The item arrangement on the test follows from considerations involving task orientation. Item number one on the test was found to elicit the least responses during the test item evaluation. As the first item on the test it was intended to immediately alert subjects to the possible difficulty of the task unobtrusively. That is, the item was intended as a vehicle to shift from informal conversation to the task. The test has no formal introduction except: "Let's play a game; here's how we play." The first test item serves as means of focusing the childrens' attention rather sharply on the task. However, the next test item, the first of the concept learning items, provides the child with assurance that there will be help in the accomplishment of the overall task. It appears no further instructions are necessary with this arrangement.

Test item number twenty on the test had been found to elicit the most responses during item pre-testing. As the last item on the test, it was intended to provide the subjects maximum opportunity to be successful as the test was completed. Also, it provides some indication that the subject has understood the task. The last ten items are not alternated with prompted test items that help verbally mediate the entire test.

The first ten items on the study developed instrument were selected to match categories with test items found on the PPVT for
cross-referenced measures of verbal performance by mode of presentation. The PPVT represents a single-response type test; the study developed measure represents an equivalent-response type test. Categories were selected from the adult Connecticut Verbal Norms where PPVT items fell within the category one or more times. The PPVT items selected for matching the categories were selected on the basis of an item falling within the range of the highest expected level of performance for the age group studied. That is, a ceiling of the first seventy-five age-leveled test items on the PPVT with six errors produces an I.Q. score of 140 for a subject age 60 months. As a consequence of the matching, the first ten test items can be construed as being age-leveled.

The "Name me or tell me something about" the stimuli format of the class recognition task presentation was heuristically derived from a test evaluation sample. The format was selected on the basis of the ability of the format to elicit responses from the very youngest subjects in the test evaluation group. Ten alternate prompted test items making up concept learning task are cued with the most frequent response in the stimulus category.

To assess associative learning, the experimental learning ability measure here follows from Rohwer's (1967) preschool experiment in principle and design essentially replicating the "Sentence elaborated" mode of presentation with one modification. In Rohwer's experiment subjects learned one set of 5 pairs in this mode of presentation of the pictorial paired-associate task in two trials. Subjects here learn two different sets of 5 pairs on two trials. Otherwise, the same presentation rate of 4 seconds with 1 second intervals between pairs
was followed. On the test trial here a stimulus was presented at 10 second intervals. Materials are from Rohwer with slides adapted from motion picture film by Coffing (1971). The common audio and slide presentation is presented in Appendix IV.

Variables

One demographic and eleven ability measures were involved in the study. They were:

1. **Age.** This measure is the chronological age of the subject computed to the nearest month at the time the PPVT was administered.

2. **I.Q.** This measure is the PPVT score thought to define vocabulary hearing recognition.

3. **M.A.** This measure is the PPVT score thought to define vocabulary hearing recognition mental age.

4. **Associative Learning.** This measure is a score of correctly learned sentence elaborated pictorial paired-associates thought to define learning efficiency by Rohwer (1967).

5. **Concept Recognition.** This measure is a score of correctly identified concept categories on the experimental test here thought to define concept usage ability.

6. **General Usage.** This measure is the cumulative score of individually criterion indexed concept category recognition responses. Each response is scored by the frequency rank
of the corresponding response found in the Connecticut Verbal Norms category tables. The scoring metric is a truncated descending scale ranging from 1 to 26 with one being the highest score per response and 26 equivalent to zero. This scale accommodates assigning equivalent scores across categories which vary in the number of responses per category with 25 being the least.

7. **Concept Usage.** This measure is a combined score of Concept Recognition and Concept Learning measures.

8. **Criterion Usage.** This measure is a combined score of General Usage and Concept Learning Usage measures.

9. **Concept Learning.** This measure is the score on the test here thought to define the ability to learn conceptual relationships through verbal examples of class inclusion.

10. **Concept Learning Usage.** This measure is the cumulative score of individually criterion indexed Concept Learning responses.

11. **Across Measure Recognition.** This measure is a score of Concept Recognition on test items matched with test items on the PPVT.

12. **Across Measure Usage.** This measure is a score of General Usage indexing responses to test items matched with test items on the PPVT.
## Summary of Main Variables

1. Age
2. I.Q.
3. M.A.
4. Associative Learning
5. Concept Recognition
6. General Usage
7. Concept Usage
8. Criterion Usage
9. Concept Learning
10. Concept Learning Usage
11. Across Measure Recognition
12. Across Measure Usage
CHAPTER IV

RESULTS

Hypothesis I

This hypothesis states that there is a statistically significant positive correlation of I.Q. scores derived from the PPVT with concept usage scores derived from the CVN in selected preschool children.

Correlation Analysis. All CVN test scores were significantly and positively correlated with PPVT I.Q. scores at the .05 level.

Regression Analysis. The results of the step-wise regression analysis examining the relationship between CVN predictors and the I.Q. criterion variable is displayed in Table 1.

Table 1 Stepwise Regression Prediction of I.Q. Criterion From Population of all CVN Predictor Variables.

<table>
<thead>
<tr>
<th>Step Number</th>
<th>Multiple R</th>
<th>R SQ.</th>
<th>RSQ Increase</th>
<th>F</th>
<th>Label of Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>.67018</td>
<td>.44914</td>
<td>.44914</td>
<td>18.75</td>
<td>General Usage</td>
</tr>
<tr>
<td>2.</td>
<td>.70539</td>
<td>.49758</td>
<td>.04844</td>
<td>10.89</td>
<td>Concept Learning</td>
</tr>
<tr>
<td>3.</td>
<td>.73155</td>
<td>.53517</td>
<td>.03759</td>
<td>8.05</td>
<td>Criterion Usage</td>
</tr>
<tr>
<td>4.</td>
<td>.75902</td>
<td>.57611</td>
<td>.04094</td>
<td>6.79</td>
<td>Concept Usage</td>
</tr>
<tr>
<td>5.</td>
<td>.77214</td>
<td>.59621</td>
<td>.02010</td>
<td>5.61</td>
<td>Concept Recognition</td>
</tr>
<tr>
<td>6.</td>
<td>.79322</td>
<td>.62920</td>
<td>.03299</td>
<td>5.09</td>
<td>Cross Measure Usage</td>
</tr>
<tr>
<td>7.</td>
<td>.79589</td>
<td>.63344</td>
<td>.00425</td>
<td>4.19</td>
<td>Cross Measure Recognition</td>
</tr>
<tr>
<td>8.</td>
<td>.79649</td>
<td>.63440</td>
<td>.00095</td>
<td>3.47</td>
<td>Concept Learning Usage</td>
</tr>
</tbody>
</table>
Hypothesis I is confirmed as can be seen in the table. The explanation of variance accounted for reached .63 with multiple R, .80 significant at the .05 level. The standard deviation of residuals for the final I.Q. criterion prediction is 12.72.

Hypothesis II

This hypothesis states that there is a statistically significant positive correlation between M.A. scores derived from the PPVT and concept usage scores derived from the CVN in selected preschool children.

Correlation Analysis. All of the CVN derived test scores were significantly and positively correlated with the PPVT MA scores at the .05 level.

Regression Analysis. The results of the step-wise regression analysis examining the relationship between CVN predictors and the PPVT MA criterion variable is displayed in Table 2.

Table 2 Stepwise Regression Prediction of M.A. Criterion From Population of all CVN Predictor Variables

<table>
<thead>
<tr>
<th>Step Number</th>
<th>Multiple R</th>
<th>R SQ.</th>
<th>RSQ Increase</th>
<th>F</th>
<th>Label of Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>.74924</td>
<td>.56136</td>
<td>.56136</td>
<td>29.43</td>
<td>Concept Usage</td>
</tr>
<tr>
<td>2.</td>
<td>.76426</td>
<td>.58409</td>
<td>.02274</td>
<td>15.44</td>
<td>Cross Measure Usage</td>
</tr>
<tr>
<td>3.</td>
<td>.77736</td>
<td>.60429</td>
<td>.02019</td>
<td>10.68</td>
<td>Concept Learning</td>
</tr>
<tr>
<td>4.</td>
<td>.78791</td>
<td>.62080</td>
<td>.01651</td>
<td>8.18</td>
<td>General Usage</td>
</tr>
<tr>
<td>5.</td>
<td>.79328</td>
<td>.62930</td>
<td>.00850</td>
<td>6.45</td>
<td>Concept Learning Usage</td>
</tr>
<tr>
<td>6.</td>
<td>.79462</td>
<td>.63143</td>
<td>.00213</td>
<td>5.13</td>
<td>Concept Recognition</td>
</tr>
<tr>
<td>7.</td>
<td>.79706</td>
<td>.63531</td>
<td>.00388</td>
<td>4.23</td>
<td>Cross Measure Recognition</td>
</tr>
<tr>
<td>8.</td>
<td>.79730</td>
<td>.63568</td>
<td>.00038</td>
<td>3.48</td>
<td>Criterion Usage</td>
</tr>
</tbody>
</table>
Hypothesis II is confirmed as can be seen in the table. The explanation of variance accounted for reached .63 with multiple \( R \), .80 significant at the .05 level. The standard deviation of residuals for the final M.A. criterion prediction is 10.15.

Hypothesis III

This hypothesis states that learning ability (Associative Learning), after Rohwer, and concept usage scores combine in a statistically significant predictive equation to account for some variation in I.Q. scores derived from the PPVT in the selected group of preschool children.

Regression Analysis. The results of the step-wise regression analysis examining the combined relationship of the CVN and associative learning predictors with the criterion PPVT I.Q. variable is displayed in Table 3.

Table 3 Stepwise Regression Prediction of I.Q. Criterion From Population of CVN and Associative Learning Predictor Variables

<table>
<thead>
<tr>
<th>Step Number</th>
<th>Multiple R</th>
<th>R SQ.</th>
<th>RSQ Increase</th>
<th>F</th>
<th>Label of Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>.67018</td>
<td>.44914</td>
<td>.44914</td>
<td>18.75</td>
<td>General Usage</td>
</tr>
<tr>
<td>2.</td>
<td>.70539</td>
<td>.49758</td>
<td>.04844</td>
<td>10.89</td>
<td>Concept Learning</td>
</tr>
<tr>
<td>3.</td>
<td>.73155</td>
<td>.53517</td>
<td>.03759</td>
<td>8.05</td>
<td>Criterion Usage</td>
</tr>
<tr>
<td>4.</td>
<td>.75902</td>
<td>.57611</td>
<td>.04094</td>
<td>6.79</td>
<td>Concept Usage</td>
</tr>
<tr>
<td>5.</td>
<td>.77214</td>
<td>.59621</td>
<td>.02010</td>
<td>5.61</td>
<td>Concept Recognition</td>
</tr>
<tr>
<td>6.</td>
<td>.79322</td>
<td>.62920</td>
<td>.03299</td>
<td>5.09</td>
<td>Cross Measure Usage</td>
</tr>
<tr>
<td>7.</td>
<td>.79972</td>
<td>.63955</td>
<td>.01035</td>
<td>4.30</td>
<td>Associative Learning</td>
</tr>
<tr>
<td>8.</td>
<td>.80334</td>
<td>.64535</td>
<td>.00580</td>
<td>3.63</td>
<td>Concept Learning Usage</td>
</tr>
</tbody>
</table>
Hypothesis III is confirmed as can be seen in Table 3. The explanation of variance accounted for reached .65 with multiple R, .80 significant at the .05 level. The standard deviation of residuals for the final I.Q. criterion prediction is 12.92. A table of residuals for the final I.Q. criterion prediction is presented in Table 4.

Table 4  Residual Table for Stepwise Regression Prediction of I.Q. Criterion Table 3.

<table>
<thead>
<tr>
<th>Observation</th>
<th>Y Value</th>
<th>Y Estimate</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>122</td>
<td>113.277448</td>
<td>8.722552</td>
</tr>
<tr>
<td>2.</td>
<td>112</td>
<td>105.710771</td>
<td>6.2892286</td>
</tr>
<tr>
<td>3.</td>
<td>113</td>
<td>113.530196</td>
<td>-0.530196</td>
</tr>
<tr>
<td>4.</td>
<td>92</td>
<td>92.3139560</td>
<td>-2.3139560</td>
</tr>
<tr>
<td>5.</td>
<td>77</td>
<td>77.1519739</td>
<td>-10.1519739</td>
</tr>
<tr>
<td>6.</td>
<td>90</td>
<td>90.8013355</td>
<td>-8.8013355</td>
</tr>
<tr>
<td>7.</td>
<td>89</td>
<td>99.7644829</td>
<td>-10.7644829</td>
</tr>
<tr>
<td>8.</td>
<td>93</td>
<td>79.5727584</td>
<td>13.4272416</td>
</tr>
<tr>
<td>9.</td>
<td>81</td>
<td>74.8246329</td>
<td>6.1753671</td>
</tr>
<tr>
<td>10.</td>
<td>84</td>
<td>80.4022842</td>
<td>3.5977158</td>
</tr>
<tr>
<td>11.</td>
<td>82</td>
<td>83.7438363</td>
<td>-1.7438363</td>
</tr>
<tr>
<td>12.</td>
<td>88</td>
<td>83.7714858</td>
<td>4.2285142</td>
</tr>
<tr>
<td>13.</td>
<td>126</td>
<td>99.0499823</td>
<td>26.9500177</td>
</tr>
<tr>
<td>14.</td>
<td>93</td>
<td>107.5019284</td>
<td>-14.5019284</td>
</tr>
<tr>
<td>15.</td>
<td>75</td>
<td>83.1448602</td>
<td>-8.1448602</td>
</tr>
<tr>
<td>16.</td>
<td>69</td>
<td>80.1620267</td>
<td>-11.1620267</td>
</tr>
<tr>
<td>17.</td>
<td>79</td>
<td>71.4449503</td>
<td>7.5550497</td>
</tr>
<tr>
<td>18.</td>
<td>80</td>
<td>95.3288020</td>
<td>-15.3288020</td>
</tr>
<tr>
<td>19.</td>
<td>93</td>
<td>81.7045910</td>
<td>11.2954090</td>
</tr>
<tr>
<td>20.</td>
<td>84</td>
<td>75.4977519</td>
<td>8.5022481</td>
</tr>
<tr>
<td>21.</td>
<td>79</td>
<td>84.8479216</td>
<td>-5.8479216</td>
</tr>
<tr>
<td>22.</td>
<td>63</td>
<td>75.5712930</td>
<td>-12.5712930</td>
</tr>
<tr>
<td>23.</td>
<td>106</td>
<td>101.1197765</td>
<td>4.8802235</td>
</tr>
<tr>
<td>24.</td>
<td>58</td>
<td>57.9999999</td>
<td>0.0000001</td>
</tr>
<tr>
<td>25.</td>
<td>85</td>
<td>84.7609580</td>
<td>0.2390420</td>
</tr>
</tbody>
</table>
Hypothesis IV

This hypothesis states that there is a statistically significant positive correlation between learning ability scores (Associative Learning), after Rohwer, and chronological age in selected preschool children.

**Correlation Analysis.** The correlation between Associative Learning scores and chronological age is positive with \( r = .55 \) significant at the .006 level.

**Regression Analysis.** In the regression analysis examining the relationship between Associative Learning scores and chronological age the explanation of variance accounted for was .30 with \( F = 9.98953 \) significant at the .006 level. Thus, Hypothesis IV was confirmed.

Hypothesis V

This hypothesis states that learning ability scores (Associative Learning), after Rohwer, and PPVT I.Q. scores are not significantly correlated in selected preschool children.

**Correlation Analysis.** The correlation between Associative Learning scores and PPVT I.Q. scores is positive and statistically significant at the .005 level.

**Regression Analysis.** In the regression analysis examining the relationship between Associative Learning scores and PPVT I.Q. scores the explanation of variance accounted for was .30 with \( r = .55 \) significant at the .005 level. Hypothesis V is rejected.
CHAPTER V

DISCUSSION AND CONCLUSIONS

The question that has been posed for the present consideration is whether it is feasible that an objective criterion can be established which can serve as a basis for facilitating verbal performance that is characteristic of mean performance on norm-referenced standardized verbal tests in selected preschool children. Using the *Connecticut Verbal Norms (CVN)* Battig and Montague (1969) as a representative criterion for concept usage and general usage; the *Peabody Picture Vocabulary Test (PPVT)* Dunn (1965) as providing representative measures of I.Q. and mental age (M.A.); and an associative learning experiment after Rohwer (1967), four relationships have been examined that are thought to be prerequisite to the clarification of the question.

The general purpose of the study has been to examine with representative measures four principal relationships thought to be relevant to the feasibility of establishing an objective standard for verbal development in selected preschool children with questionable verbal performance on norm-referenced standardized general verbal intelligence tests. The first involves the relationship between I.Q. derived from the PPVT and experimental measures of concept usage derived from the CVN. The second involves the inter-relationship of learning ability (associative learning) defined as acquisition of new knowledge by Rohwer (1967) and concept usage in the prediction of norm-referenced standardized verbal I.Q. The third involves the relationship between

...
learning ability after Rohwer and chronological age (C.A.). The fourth involves the relationship between learning ability after Rohwer and PPVT I.Q.

Verbal intelligence test norms are generally considered to be meaningful in educational practice when they are representative of the population for whom they are designed to reference. Such norms are not ordinarily considered as goals of attainment. They are comparative measures primarily designed to classify individual performance on a continuum of relative performance established by a norming group under the same conditions. Among other things, a low performance by a sample population relative to such norms suggests characteristic differences between the norming population and the sample population. The initial identification of these differences empirically rather than by sociological inference raises a series of problems for educational research prerequisite to the clarification of the feasibility of ultimately establishing a substantive standard which can serve as a basis for reconciling characteristic differences between such populations.

As a preliminary step to evaluate the feasibility of establishing an objective verbal standard capable of identifying characteristic differences in verbal performance on general verbal ability standardized norm-referenced tests and adaptable to reconciliation of the differences found, this exploratory study has examined approaches to a series of related problems involved in determining the feasibility. These problems were devising a method of examining characteristic differences in verbal performance on such measures, evaluating these differences, and identifying a representative standard by which the evaluation could be made.
The CVN have been suggested as characteristically representative of a substantive criterion for verbal performance in very young children because it was thought that the CVN's categorical structure could be shown to be developmental. They were used to facilitate equivalent-response type testing as a method of examining variability in response to categorical concepts relative to single-response type testing which is generally characteristic of norm-referenced tests. As a basis for examining the relationships between performance on norm-referenced measures of general verbal ability and criterion-referenced measures, responses to the equivalent-response concept category labels used for stimuli were indexed according to the general usage of college freshmen. The general usage of college freshmen who have by some means achieved scholastic success was seen as a representative goal for ultimate verbal attainment underlying mandatory education. Prompting tasks using twenty CVN category verbal labels with the highest normed frequency response as a cue were used as stimuli to elicit the subject's response from memory to provide data for analysis in examining the relationships involved in the study.

The first relationship examined followed from the assumption that adult category verbal norms reflect variability in concept usage and verbal general usage relative to such norms as a criterion and also relative to norm-referenced standardized verbal ability tests. It was hypothesized that there was a positive correlation between I.Q. scores derived from the PPVT and concept usage scores derived from the CVN in selected preschool children.

The CVN scores are derived from the children's performance on two equivalent-response type tasks which were thought to provide four
measures of concept usage with concommitant indices of verbal usage. One task is a straightforward free recall task that asks the child to tell or name something that is an instance of a class label or a description of a CVN category with a single word. "Name me something we wear on our feet," is an example of how the normed CVN category of A Kind of Footgear is labeled as a stimulus. The verbal responses to this task is called Concept Recognition. When the child's response falls within the stimulus category it is scored as a correct response for a measure of Concept Recognition. This response is then indexed by the rank of its equivalent ranked frequency response on the CVN as a score for a measure of General Usage. Ten of the twenty categories constituting the test have instances of such categories as single-response type test items on the PPVT within the expected range of performance on that measure for a 60 month old child attaining a PPVT derived I.Q. of 145. The category test item scores are tallied separately for measures of Across Measures Recognition and Across Measures Usage respectively.

The prompted equivalent-response type task provides a measure of what was thought to be concept learning. In addition to the description or category label provided in the Concept Recognition task, the child is provided with an example of class inclusion with the cue in the example being the highest frequency ranked response in the stimulus category. The child is asked to respond with another instance of the category. For example, "A shoe is something we wear on our feet, name me or tell me something else we wear on our feet." Shoe is the cue in this instance. The child's response was scored and indexed in the same way as in the Concept Recognition task and the variables called Concept Learning.
and Concept Learning Usage respectively. The combined Concept Recognition and Concept Learning scores constitute the Concept Usage variable with the combined usage scores from these measures constituting the Criterion Usage Variable.

Statistically, as a meaningful general reference for the subsequent interpretation of the relative strengths, magnitudes or closeness in the relationship between the variables examined in this study sample of twenty-five, a minimum correlation coefficient of .396 is required to provide 95 percent confidence that the relationship is statistically significant. "Statistically significant" indicates only that it is unlikely that the true correlation in the total population is zero.

For Hypothesis I, the correlation analysis showed all the CVN derived variables positively and significantly correlated with the PPVT derived I.Q. variable as predicted. In the order of magnitude or closeness found in these relationships considered individually the coefficient for Concept Usage was highest of the task variables at .648 followed by Concept Learning at .644 then Across Measures Recognition at .604 with Concept Recognition lowest at .594. The order of magnitude found in the correlations between PPVT I.Q. scores and the scores of the concept usage verbal usage indicies was General Usage as the highest at .670 followed by Across Measures Usage at .602 then Concept Learning at .545 with Criterion Usage lowest at .461.

The strongest statistical relationship found by the correlation analysis between I.Q. and the CVN derived variables was the relationship between I.Q. and General Usage the verbal usage index of Concept
Recognition. In other words, the variability in general usage with the CVN as criterion has the closest relationship to the variability of I.Q. shown by the analysis in terms of scores on the measures used. When considering individual relationships between CVN derived variables and PPVT I.Q. it must be remembered that the tasks are not the same. The PPVT task is a single-response type test; the CVN derived tasks are equivalent-response type tests. Consequently, the performance scores are reflecting different values on the different scales in the relationship.

In terms of tasks, on the I.Q. measure the children's scores were determined solely by the number of correct answers given which provides the variability on this measure. On the Concept Recognition measure the children's score was determined by the number of correct equivalent-response type answers given also. However, it is the value of the indexed response in terms of verbal usage with the CVN as a criterion that provides the verbal variability on this measure having the strongest relationship with the variability of I.Q. on the PPVT scale.

In terms of scale, the relationship between General Usage and I.Q. was found to be such that with the CVN as a criterion: as the value of Concept Recognition responses approach criterion on its scale, I.Q. scores tend to rise on the PPVT scale relative to the measures used. That is to say, relative to the number of test items attempted on the PPVT which varies from child to child, such has been found to be the case in the relationship between General Usage and I.Q.

Regression analysis showed that General Usage and Concept Learning variables combine in the explanation of variance in I.Q. accounted for contribute nearly fifty percent toward the explanation. With this the
case, the conclusion drawn from the interpretation of the statistical analysis is that alternative plausible explanations not withstanding Hypothesis I is statistically and logically confirmed for the exploratory purpose of the study.

In the examination of the relationships above, the relationships were examined on the basis of comparative performance among children of the same chronological age; Hypothesis II attempts to provide a framework to examine the relationship between the CVN and representative age norms thought to be an index of intellectual development expressed in terms of mental age. The hypothesis followed from the assumption that adult category verbal norms as a criterion were developmental relative to norm-referenced standardized test of general verbal ability. It stated that there was a statistically significant positive correlation between mental age scores derived from the PPVT and concept usage scores derived from the CVN in selected preschool children.

For Hypothesis II, the correlation analysis showed all the CVN variables positively and significantly correlated with the PPVT derived M.A. as predicted. The analysis also showed stronger correlations in each of the M.A. relationships with CVN variables than was shown with I.Q. Regression analysis showed that Concept Usage and Across Measure Usage combine as the best predictors of M.A. in the group of selected preschool children studied on the measures used. This being the case, the interpretation of the statistical significance of the relationships found with the CVN as a criterion the relationship between the developmental nature of the CVN and the PPVT is such that, alternative plausible explanations not withstanding, Hypothesis II is statistically
and logically confirmed for the exploratory purpose of the study.

Hypothesis III predicted that Associative Learning scores and scores derived from the CVN variables would combine significantly to predict I.Q. scores on the PPVT. Assuming the CVN to be developmental as a criterion for verbal ability, the hypothesis as such was principally to examine the contribution of what was thought to be primarily a learning task by Rohwer (1967) towards the prediction of I.Q.

The correlation analysis showed Associative Learning to be positively and significantly correlated with all the CVN derived variables. In order of magnitude, the correlation between Associative Learning and CVN task scores was shown to be relatively higher than the correlation between the CVN derived indices. The regression analysis showed that Associative Learning combines with CVN derived variables to predict I.Q. scores in the selected group of preschool children to confirm Hypothesis III. Associative Learning contributed better than one percent (.01035) toward the explanation of variance in I.Q. scores accounted for when combined with the CVN variables. The total explanation of variance in I.Q. scores accounted for reached 64 percent. Alternative plausible explanations notwithstanding, the conclusion drawn from the analysis is that while contributing to the prediction of I.Q. scores, Associative Learning scores correlate with Concept Learning scores such that the Concept Learning task is sufficiently sensitive alone in reflecting developmental relationships as the analysis showed relative to M.A.
Hypothesis IV predicted that Associative Learning would be positively and significantly correlated with chronological age. The hypothesis follows from the assumption that a primarily learning task that was assumed to be less affected by previous opportunities to learn by Rohwer (1967) would reflect learning ability developmentally. The correlation analysis showed the correlation between chronological age in months and Associative Learning scores to be positive and significant. The correlation was .5428. Thus, the hypothesis was confirmed. A tentative conclusion drawn from the analysis is that this relationship suggests that the variance found in the analysis of Hypothesis III may be explained in terms of the developmental nature or structure of the CVN related tasks after examining the relationship between Associative Learning and I.Q. independently.

Following from the same assumptions by Rohwer above, Hypothesis V predicted that Associative Learning would not be significantly correlated with I.Q. relative to verbal ability with the CVN as criterion in the group of selected preschool children. It must be remembered that Rohwer (1967) was investigating ways of facilitating verbal ability as measured on norm-referenced standardized tests of general verbal ability such as the PPVT. The analysis showed that the correlation between Associative Learning and PPVT I.Q. scores was positive and statistically significant. The correlation was .5424. Hypothesis V is rejected.

While Hypothesis V is rejected, in examining the correlation between Concept Learning and Associative Learning in the prediction of I.Q. in the present study the analysis suggests difficulties encountered in the interpretation of comparative sample studies similar
to the Rohwer (1967) study in preschool children. In Rohwer's own words:

The main conclusions of the present experiment are clear. First, the pictorial PA task is quite appropriate for use in investigation learning processes in preschool children. Reliable differences in learning efficiency as a function of experimental conditions can be detected even in this young population. Second, the PA task promises to have sufficient reliability to warrant its conversion into a test instrument for the purpose of classifying children with respect to learning proficiency. (p. 87)

With these clear conclusions and reliable measures Rohwer could generalize from these findings that I.Q. predicts learning efficiency modestly well for higher-strata children, but was unrelated to learning efficiency in lower-strata without suggesting a practical reason why. Specifically, Rohwer concluded:

In the pre-school age range (three to five years of age) inferior performance among lower-strata children is observed even under optimal conditions of learning. (p. 102)

Optimal conditions of learning as a function of experimental conditions can be contrasted in the present study with conditions of learning as a function of representative experimental design in terms of descriptive generality. That is in terms of similar tasks, in the experimentally controlled Associative Learning task the subject's ability to respond to the stimulus can be limited by his verbal usage in performing the verbal task involved in the experiment. If such happens to be the case, the only way this can be reflected is in the score of correct answers. In contrast, on the verbally similar Concept Learning equivalent-response type task the child's verbal ability is not limited by verbal usage in quite the same way. The child is free to respond at
least in terms of experimentally controlled responses. The child can relate whatever the child chooses to the stimulus provided, but in doing so does not obviate the fact that the child must still judge whether the response is appropriate. Relative to magnitudes of the correlations involved, the I.Q. prediction analysis tends to support the suggestion that differences in verbal usage between the two samples studied by Rohwer accounts for the differences in the ability of I.Q. to predict learning efficiency in both samples equally as well.

Learning efficiency or Associative Learning in the present study and Concept learning scores were found to be highly correlated (.75) and correlate with I.Q. scores. The magnitudes of the other relationships in the performance of the children on both tasks suggest that in the prediction of I.Q. scores in the same children it can be demonstrated statistically that Concept Learning is a better predictor with its CVN indexed measure of general usage than Associative Learning. In terms of descriptive generality, Concept Learning with its index Concept Learning Usage which is referenced to a known representative index of general usage predicts I.Q. scores with a practical explanation better than Associative Learning as a function of experimental conditions.

If the foregoing relationships are found to hold across comparative samples, the finding in terms of methodology is significant further. Equivalent-response testing and response indexing with a suitable criterion provide practical information as Snow (1973) has suggested.
For the exploratory purpose of the study the summary conclusion is that representative measures and methodology here have demonstrated statistically significant relationships between verbal performance on such measures and performance on a representative norm-referenced standardized test of general verbal intelligence with categorical norms as a criterion. The magnitude of the relationships suggest that the feasibility of the ultimate use of categorical verbal norms as a basis for verbal performance objectives in reconciling differences in general usage in selected preschool children has not been precluded by the findings of the present study.

If the investigation is to be carried further, the profile of individual performance on the Concept Learning task provides a basis for facilitating learning experiences that is more specific than that obtained from the results of norm-referenced standardized tests. Concept elaboration with high frequency usage should reflect in higher performance on standardized tests. If this can be shown, the feasibility of ultimately establishing a substantive reference for verbal performance in preschool children will have been demonstrated.
APPENDIX I

STUDY SAMPLE POPULATION
DEMOGRAPHIC AND VERBAL CHARACTERISTICS

Study Sample Population Demographic And Verbal Characteristics

Age

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Row</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Column Total

| 2 | 1 | 2 | 1 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 1 | 3 | 2 |

Study Sample Population Demographic and Verbal Characteristics
Crosstabulation of PPVT I.Q. by Age in Months
APPENDIX II

CONNECTICUT VERBAL NORMS TEST PROMPTS AND CATEGORIES

Test Prompts

1.1 Name me a piece of furniture.
1.2 A chair is a piece of furniture, name me another piece of furniture.
2.1 Name me something we go places in.
2.2 A car is something we go places in, name me another thing we go places in.
3.1 Name me an animal.
3.2 A dog is an animal, name me another animal.
4.1 Name me anything we wear on our feet.
4.2 We wear shoes on our feet, name me something else we wear on our feet.
5.1 Now, name me anything we wear.
5.2 A shirt is something we wear, name me another thing we wear.
6.1 Name me a toy.
6.2 A doll is a toy, name me another toy.
7.1 Name me a bug.
7.2 A fly is a bug, name me another bug.
8.1 Name me a kind of money.
8.2 A dollar is a kind of money, name me another kind of money.

Categories

An Article of Furniture
A Type of Vehicle
A Four-Footed Animal
A Type of Footgear
An Article of Clothing
A Toy
An Insect
A Kind of Money
9.1 Name me something to make music with.

9.2 A piano is something to make music with, name me another thing to make music with.

10.1 Name me something to read.

10.2 A book is something to read, name me another thing to read.

11. Name me a part of a house.

12. Name me a vegetable.

13. Name me a part of your body.

14. Name me a weapon.

15. Name me a flower.

16. Name me something about weather.

17. Name me a kind of tree.

18. Name me a color.

19. Name me a kind of bird.

20. Name me something to drink
APPENDIX III

ASSOCIATIVE LEARNING COMMON PRESENTATION

Welcome to our picture game. Here's how you play. Try to remember the things you see together, so when you see only one thing you can answer out loud what went with it. Listen to learn what the things are. Let's play a game.

Here are some things to remember together.
Remember these things together.
The blanket covers the tree.
The car hits the wagon.
The elephant kicks the clock.
The needle pops the balloon.
Now give your answers out loud. What went with:
The car
The elephant
The blanket
The needle
Let's play the game again. Here's some things to remember together.
The milk fills the bowl.
The rock breaks the bottle.
The fire burns the bed.
The fork cuts the cake.
The pencil tears the paper.
Now give your answers out loud. What went with:

The rock
The pencil
The fire
The milk
The fork

Now let's play our last game. Here's some things to remember together.

The spoon rolls the egg.
The hand hits the hat.
The bat breaks the cup.
The man bends the pole.
The axe cuts the wood.

Now, give your answers out loud. What went with:

The bat
The spoon
The man
The hand
The axe

Thank you very much for playing our game.
## APPENDIX IV

### CORRELATION COEFFICIENTS

<table>
<thead>
<tr>
<th></th>
<th>CA</th>
<th>MA</th>
<th>LA</th>
<th>CVNR</th>
<th>CVNF</th>
<th>CVNRT</th>
<th>CVNFT</th>
<th>PCVNR</th>
<th>PCVNF</th>
<th>PTCVNR</th>
<th>PTCVNF</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ</td>
<td>.20</td>
<td>.55</td>
<td>.54</td>
<td>.46</td>
<td>-.39</td>
<td>.53</td>
<td>-.42</td>
<td>.56</td>
<td>-.44</td>
<td>.46</td>
<td>-.44</td>
</tr>
<tr>
<td>IQ</td>
<td>.89</td>
<td>.55</td>
<td>.59</td>
<td>-.67</td>
<td>.65</td>
<td>-.46</td>
<td>.64</td>
<td>-.54</td>
<td>.60</td>
<td>-.60</td>
<td></td>
</tr>
<tr>
<td>MA</td>
<td>.64</td>
<td>.68</td>
<td>-.74</td>
<td>.75</td>
<td>-.60</td>
<td>.73</td>
<td>-.60</td>
<td>.67</td>
<td>-.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LA</td>
<td>.65</td>
<td>-.60</td>
<td>.73</td>
<td>-.46</td>
<td>.75</td>
<td>-.53</td>
<td>.71</td>
<td>-.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVNR</td>
<td>-.92</td>
<td>.95</td>
<td>-.70</td>
<td>.72</td>
<td>-.51</td>
<td>.93</td>
<td>-.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVNF</td>
<td>-.91</td>
<td>.83</td>
<td>-.74</td>
<td>.62</td>
<td>-.88</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVNRT</td>
<td>-.74</td>
<td>.90</td>
<td>-.72</td>
<td>.89</td>
<td>-.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVNFT</td>
<td>-.64</td>
<td>.63</td>
<td>-.67</td>
<td>.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCVNR</td>
<td>-.89</td>
<td>.72</td>
<td>-.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCVNF</td>
<td>-.54</td>
<td>.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTCVNR</td>
<td>-.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Variable Name

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA</td>
<td>Chronological Age in months</td>
</tr>
<tr>
<td>IQ</td>
<td>Peabody Picture Vocabulary Test I.Q.</td>
</tr>
<tr>
<td>MA</td>
<td>Peabody Picture Vocabulary Test Mental Age</td>
</tr>
<tr>
<td>LA</td>
<td>Associative Learning</td>
</tr>
<tr>
<td>CVNR</td>
<td>Concept Recognition</td>
</tr>
<tr>
<td>CVNF</td>
<td>General Usage</td>
</tr>
<tr>
<td>CVNRT</td>
<td>Concept Usage</td>
</tr>
<tr>
<td>CVNFT</td>
<td>Criterion Usage</td>
</tr>
<tr>
<td>PCVNR</td>
<td>Concept Learning</td>
</tr>
<tr>
<td>PCVNF</td>
<td>Learning Usage</td>
</tr>
<tr>
<td>PTCVNR</td>
<td>Cross Measures Recognition</td>
</tr>
<tr>
<td>PTCVNF</td>
<td>Cross Measures Usage</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


Brown, R. How shall a thing be called? Psychological Review, 1958, 65(1).


Kent, G. H., and Rosanoff, A. J. A study of association in insanity, 

Kintsch, W. Learning, Memory, and Conceptual Processes. New York: John 

Labov, W. The logical non-standard English. In Language and poverty. 

Labov, W., Cohen, P., Robins, C., and Lewis, J. A. A study of the non-
standard English of Negro and Puerto Rican speakers of New York 
City. Vol. I and II, Final Report on Contract OEC-6-10-059, 
Cooperative Project 3288, 1968.

Martin, E. Relation between stimulus recognition and paired-associate 

Palermo, D. S., and Jenkins, J. J. Word Association Norms: Grade School 
through College, University of Minnesota Press: Minneapolis, 
1964.

Popham, W. J., and Husek, T. R. Implications of criterion-referenced 
measurement. In Perspectives in Educational and Psychological 
Measurement. Edited by G. Bracht, K. Hopkins, and J. Stanley, 

Rohwer, W. D. Social class differences in the role of linguistic 
structures in paired associate learning. Office of Education 
contract 6-10-273, University of California, Berkeley, 1967.

Snow, R. E. Representative and Quasi-Representative Designs for 
Research on Teaching. Paper presented at the Annual Convention 
of the American Educational Research Association, New Orleans, 

Tulving, E. and Pearlstone, A. Availability versus accessibility of 
information in memory for words, *Journal of Verbal Learning and 