AN ABSTRACT OF THE THESIS OF

Laura S. Hansen for the degree of Master of Arts in Interdisciplinary Studies in the co-departments of Speech Communication, Psychology and Education presented on August 7, 1984.

Title: Inter-test Comparison of Three Preschool Language Tests: SICD, PLS, and PPVT-R

Abstract approved: Redacted for privacy

Emery V. Hildebrandt

Thirty preschoolers were tested for language skills using the Preschool Language Scale (PLS), Peabody Picture Vocabulary Test-Revised (PPVT-R) Form L, and Sequenced Inventory of Communication Development (SICD). Each child's age scores on one test were compared to his/her scores on the other two tests. As a group, the children achieved consistently higher age scores on the PLS than on the PPVT-R, and in turn, higher age scores on the PPVT-R than on the SICD. In terms of age scores, the widest inter-test difference was between the receptive subtests of the PLS and SICD, with a mean difference score of 13 months. A paired t-test applied to the within-subject difference scores was significant for every inter-test comparison. Correlation coefficients were not high enough to warrant prediction of an individual child's actual scores on one test from his scores on another test. In the author's opinion, the tendency of one test to produce higher scores than another should be kept in mind during test selection and interpretation.

Analysis of age scores showed that this group of children achieved slightly better scores on the expressive portion of the SICD than on the receptive portion of the SICD. However, the children achieved slightly better scores on the receptive half of the PLS than on the expressive half of the PLS.

Other studies involving inter-test comparisons are cited in the literature review. Many studies report correlation coefficients but do not include measures of central tendency. It is important that both kinds of data be reported in future studies.

Inter-test Comparison of Three Preschool Language Tests: SICD, PLS, and PPVT-R

by

Laura S. Hansen

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Arts in Interdisciplinary Studies

Completed August 7, 1984

Commencement June 1985

APPROVED:

Redacted for privacy

Professor of Speech Communication in charge of major

Redacted for privacy

Professor of Education in charge of co-field

Redacted for privacy

Professor of Psychology in charge of co-field

Redacted for privacy

Chairman of Department of Speech Communication

Redacted for privacy

Dean of Graduate School

Date thesis is presented August 7, 1984

Typed by researcher for Laura S. Hansen

ACKNOWLEDGEMENTS

First of all, I feel lucky to have received my graduate training at Oregon State University under Dr. Emery Hildebrandt and Cathryn Westbrook. Many times Dr. Hildebrandt taught by coaxing information out of me that I only half-knew. I have worked hard to try to meet his standards. Outside the classroom, he is a clinician who genuinely cares about and enjoys his clients. He never loses sight of the "big picture"--effective communication--and he never seems to forget any details, either!

Cathryn Westbrook provided the original idea for this project, and taught me (long ago) how to administer the tests. She is a crack diagnostician with many ideas for enticing preschoolers into displaying their communication skills. Both she and Dr. Hildebrandt require a sound theoretical rationale for clinical decision-making, whether one is choosing an evaluation instrument or planning therapy. It is our joint hope that this project and others like it will help clinicians choose and interpret tests more judiciously.

Five day care centers and one parent education class allowed my intrusions into the daily routine and provided space for testing. Teachers and directors often helped me get to know my subjects as we began testing. Corvallis Community Day Care Center deserves special mention for being the "first". Director Judy Marcus was quick to suggest procedural matters that smoothed the way for all.

Thanks also go to my 30 real-live "subjects" plus seven

"practice-subjects" and their parents. Thank you for extending yourselves and making the project possible.

I was greatly assisted in research design and choice of statistical tests by Dr. Carol Saslow. More tutoring in statistics was willingly offered by John Van Sickle.

Oregon State University Library provided a free computer-assisted bibliographical search. The OSU Computer Center provided computer time.

I was able to accelerate my graduate studies and preparation for the thesis thanks to a generous scholarship from the Benton County March of Dimes (Ann Ellenburg Scholarship).

My husband Bruce Frey helped me talk to a computer, and was a willing editor. His never-ending faith in me was supportive (when I felt confident too) but sometimes unnerving (when I knew I was behind schedule). My mother and sister also offered encouragement at key points.

A very special thanks goes to Nacha Tola, Teresa Hartlerode, and Edith Poblete, babysitters for my children through my graduate student years. They and I know it would have been impossible without them. Each day I could go to school knowing that my children were going to enjoy themselves in a warm, loving home. I can't thank them enough. In addition, Campus Day Care was a very special place for Peter during the past year.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTER-TEST COMPARISON OF THREE LANGUAGE TESTS: SICD, PLS, AND PPVT-R

INTRODUCTION

Focus of the Study

This study is concerned with the assessment of language skills in preschoolers. More specifically, it compared children's test scores on three tests which are commonly used to determine a child's level of language development: the Preschool Language Scale (PLS), Peabody Picture Vocabulary Test-Revised (PPVT-R), and Sequenced Inventory of Communication Development (SICD). Each of these tests yields an "age score" so that a clinician may determine that a child understands and/or expresses himself like, for example, a typical 3 1/2 year old.

Language age scores are important because they help the clinician answer the following questions:

1. Does the child have a problem with language-learning?
2. What is the severity of the language problem?
3. How does the child's receptive language (understanding of language addressed to him) compare to his expressive language (the child's spoken language)?

A child's obtained language age score depends on a number of variables, including the child's actual skills, variables of the testing

situation, and variables of test design. In order to use a language age score with confidence, a clinician must be familiar with sources of variation, as well as with research studies regarding the reliability and validity of the test scores. These help the clinician to interpret language age scores with the proper degree of caution.

The present study is an inter-test reliability study. The three tests (SICD, PLS, and PPVT-R) were chosen because of their widespread use in the language evaluation of preschoolers. All three tests were given to each of 30 preschoolers, permitting the investigator to compare a child's scores on one test with his/her scores on the other two tests. The goals of the study were to determine:

1. Are these tests equivalent in their estimates of language ages?
2. Is any of the three tests so similar to another, that one could be used in the place of the other?
3. If the three tests are not equivalent, can the examiner use a correction factor or some other means to predict a child's score on one test, from his/her score on another test?

The results of this study demonstrate the tendency of one test to score a child lower or higher than the other tests. This data should be of practical value for choosing among and interpreting the tests.

The data yielded by this study also permit intra-test comparisons for the two tests which include receptive and expressive language subtests, the PLS and the SICD. A common use of these tests is to

determine if a child's receptive and expressive language skills are equivalent. Some children present with disabilities in both understanding and in speaking, while others appear to have adequate understanding but inadequate abilities to generate words, sentences, etc., correctly and appropriately. Comparing scores from the receptive and expressive subtests of the PLS or SICD is only meaningful if the subtests are equal in difficulty. Therefore, a secondary goal of this study was to answer the question:

4. Can children be expected to achieve similar scores in receptive and expressive language on the PLS? on the SICD?

The author will use many abbreviations for test names and test scores throughout this thesis. In order to aid the reader, a list of abbreviations was prepared. It appears as Appendix A.

Limitations of the Study

First, it should be mentioned that the SICD, PLS, and PPVT-R yield other scores and have other uses than those examined here. The PPVT-R yields derived scores other than the age score and standard score. Besides being used as a test of language or vocabulary, the PPVT-R has been used to measure intelligence, academic achievement, and scholastic aptitude (Dunn and Dunn, 1981). The PLS and SICD inventory several types of language skills which permit the examiner to draw up a profile of the child's strong/weak areas in language learning. Any of the three tests (SICD, PLS, or PPVT-R) may be used to help plan a therapy program for a child, or evaluate the effectiveness of a program (pre- and post-testing). During the initial

evaluation of a child, any of the tests may serve to indicate areas where more in-depth testing is needed. These applications were not a focus of the present study.

Secondly, this is an inter-test reliability study, not a validity study. In other words, test scores from one test were compared with scores from another test, but actual language ages of the children were not established by any other means. Therefore, as a result of this study, one may not conclude that any one test does a "better" job of determining language ages. One may, however, conclude that one test will give a higher language score than another. It is hoped that speech clinicians will examine the inter-test comparisons and use this information when choosing tests, or when interpreting age scores from any one of these tests.

Context for Use of Language Tests

Speech pathologists who work at elementary schools screen kindergarten and primary grade children to determine if any of them have troubles with speech or language. Some preschools also offer developmental screening. For such screening, the PPVT-R and PLS may be used; the SICD is too long for this purpose. School-age children who are identifed by the screening program to be "at-risk" for speech or language problems are then evaluated at length, by the school speech pathologist. Preschoolers thought to have language problems are often referred to speech clinics for evaluation. A comprehensive speech/hearing/language evaluation, whether at a school or in a clinic, commonly involves some formal tests as well as informal testing

procedures. The SICD, PLS, and/or the PPVT-R may be used as part of the formal test battery. The clinician has a limited time in which to evaluate the child. It is therefore important to choose appropriate tests, beginning with tests that establish a general level of functioning, and following with tests that zero in on problem areas (Sanders, 1979).

Some Issues in the Choice of Language Tests

I. Test Philosophy

There are many schools of thought regarding the acquisition and learning of language, and even more approaches to testing for language problems. The speech/language clinician needs to realize that in choosing a formal test, he/she is in effect "buying the author's definition of language" (Siegel, 1976, as cited in Hutchinson et al., 1979). Informal assessment instruments also vary in emphasis and scope. A clinician may focus on the syntactic aspects of an informal language sample, for example, leaving out considerations of pragmatics, semantics or phonology.

II. Test Length

Of course, it is impossible to observe all of a child's language behaviors. In developmental testing, test authors (and clinicians) are challenged to select a few items at several developmental levels, that will accurately differentiate between more and less mature language users. Tests must be short enough so that fatigue does not depress a young child's performance.

III. Formal vs. Informal Testing

The PLS, SICD, and PPVT-R are formal tests; each test is published, distributed widely, and designed to be administered and interpreted according to specific instructions given in the test manual. In each case, test authors have conducted research which aids in interpreting test results.

Informal data include: 1) samples of the child's language while engaged in a particular activity; 2) specific language tasks administered apart from formal tests; 3) case history information; 4) referral information; and 5) observations of abilities related to language (motor, cognitive, auditory, speech, etc.).

Formal tests (or specific observational data) are useful because they briefly summarize a child's performance at one point in time. Consumers and legislators demand that speech/language clinicians be able to document the reasons for including or not including a child in in a program of language remediation. For children who are included, an IEP (Individual Educational Program) is written, specifying the results of the initial evaluation, therapy objectives, therapy procedures, and (later) post-testing results. Formal test results are easy to obtain, report, and compare for this purpose. Clinicians find it easy to begin with summary data like age scores, when describing the results of an evaluation to parents.

In addition to helping a clinician be more accountable, formal tests yield genuinely useful data for initial assessment of the

child's abilities, therapy planning, and therapy evaluation. This study focussed on the ability of the SICD, PLS, or PPVT-R to give a first impression of the child's level of language development during an initial assessment.

Informal testing has some advantages over formal testing. Informal procedures may be more flexible, enabling a clinician to try several ways of presenting a task. In collecting an informal language sample, the clinician presents a variety of toys, and tries several kinds of interaction to (hopefully) bring out the best possible language performance. Another informal procedure is to observe a parent and child playing together, or accomplishing a task together. The clinician can gather valuable information about the child's play behavior and language as well as the parent's communication techniques. Most of all, informal language samples or observations tend to yield language that is more natural and spontaneous. In contrast, formal test items are often short answers, with a rapid change of topic and no "real" context.

The main disadvantage of informal procedures is that the child's performance may be influenced as he/she interacts with the examiner. The examiner's initial prejudices regarding a child's language abilities may guide the examiner's behavior during collection of an informal language sample; this will bias the results. During observational procedures, examiner prejudice may influence the kinds of behaviors that are noticed.

Most speech/language evaluations include both formal and informal measures, so that the advantages of both are present.

IV. Subjectivity

The examiner must decide whether or not a child has answered test items (or performed informal tasks) correctly. Some test items demand a specific answer, and little examiner judgment is needed to score the answer correct. Receptive vocabulary items on the PPVT-R are an example of fairly objective test items. Either the child indicates the correct picture or he doesn't. Subjective items require more examiner judgment. If a child is asked to define "bird", he may answer:

"It flies. It has a nest."

"An animal."

"A parakeet and an ostrich."

Any of these might be correct, depending on the child's age and criteria set by the test for each developmental age level.

Formal tests tend to have more objective items than informal measures. The more objective items are easier to score, but they have a drawback, mentioned earlier, in that they are less like natural language. The SICD and PLS include some more "subjective" items; the test manuals provide guidelines for judging answers.

Even in the most "objective" tests, examiner judgment is needed to determine when the child is tired, when he is teasing, etc. In formal testing, the clinician's goal is to obtain the best possible language performance while following the test procedures faithfully.

Most clinicians use a combination of "art" and "science" to make a final judgment of a child's language skills (Allen et al., 1981). In other words, the decision to schedule a child for language therapy

is based on measures that are formal and informal, objective and subjective; in addition, a clinician will form impressions of a child's abilities by comparing him with other children the clinician has seen (Siegel, 1975).

Description of Tests and Literature Review

I. Preschool Language Sample (PLS)

The PLS is a screening and diagnostic instrument for children ages 1-7. It establishes expressive (verbal ability) and receptive (auditory comprehension) language age scores in 1 1/2 month intervals through age five, and in 3-month intervals from age five to seven.

The manual for the PLS states that (Zimmerman et al., 1979):

> unlike standard tests, skills in the PLS are positioned at age levels that represent the point at which most children have achieved such competency.... if a child scores below his age level on the Preschool Language Scale, he should be considered 'at risk' for language problems (pp. 4-5).

IA. PLS Test Administration. The PLS is divided into two subtests: Auditory Comprehension and Verbal Ability. The child performs both verbal and motor tasks. Many items involve a question-answer format. Some objects (blocks, sandpaper, coins) are used, however most comprehension items involve the use of a picture book accompanying the test. Each subtest is divided into sections representing six-month intervals in language development. For children over age five, one section represents a year's language development. Each section contains four test items. The examiner must estimate

the appropriate section (age level) at which to begin. All four items in an early section must be passed in order to work forward on the test. This is the basal. The examiner then works forward, item by item, through several sections (age levels). The test is discontinued when a child fails all four items in an advanced section. Each correct test item (up to the five-year level) represents 1 1/2 months of language development, for the purpose of obtaining an age score.

Many test items are repeated exactly, from one age level to the next. The examiner gives these items only once, and records the child's response. Criteria for passing the item at each age level differ. For example, a child who demonstrates an understanding of 2 of 5 prepositions will pass this item at the 2 year, 6 month level but not at the 3 year, 6 month level where a score of 4 out of 5 is required.

An articulation test using a word repetition format is included in the test. A child's score on specific consonant sounds contributes to his overall verbal ability score.

Once age scores for auditory comprehension (AC) and verbal ability (VA) are obtained, they can be converted into quotients called the Auditory Comprehension Quotient (ACQ) and the Verbal Ability Quotient (VAQ). Each quotient represents the ratio of a child's Auditory Comprehension score (or Verbal Ability score) to his chronological age, multiplied by 100, as follows:

ACQ = AC Age ÷ CA X 100

VAQ = VA Age ÷ CA X 100, where CA is the chronological age

(Zimmerman et al., 1979, p. 21).

The PLS is easy for a speech clinician to administer after several practice sessions, and it should take from 30-45 minutes to give. A profile is provided which designates the types of skills which each test item uses, including sensory discrimination, logical thinking, grammar and vocabulary, memory and attention span, temporal/spatial relations, and self-image. This is one of the few language tests covering the span of ages from one to seven.

IB. PLS Test Design and Norms. The PLS was first published in 1969, with a revised edition appearing in 1979. It was designed not only to yield age scores and quotients, but also to pinpoint areas of deficiency and strength, first in receptive and then in expressive language (Darley, 1979). The test authors call the PLS "an evaluation instrument, still in experimental form" (Zimmerman et al., 1979, p. 2). No normative data is given in the test manual. The authors of the test invite others to conduct normative studies. In reviewing the PLS, Proger (1971) cautioned examiners to use care in interpreting the age scores, since no breakdown is available of the percentage of three year olds (for example) who actually score at the 3-0 age level. Proger (1971) adds that, since the PLS test items were derived from several other assessment instruments, standardization would have been done with different normative populations. In addition, it could be argued that individual test items should carry different amounts of weight in determining a child's language age. However, the test weighs items equally--each item up to the five-year level is

counted as 1 1/2 months of language development.

IC. PLS Validity. The test manual reports several unpublished concurrent validity studies. In one, (Higbee, 1974, as cited in Zimmerman et al., 1979), the PLS and Illinois Test of Psycholinguistic Abilities (ITPA) were administered to 15 cerebral palsied children. The ITPA mean language quotient correlated highly (r=.97) with the PLS mean language quotient. Scott (1973, as cited in Zimmerman et al., 1979) compared the performance of 32 preschoolers on the Utah Test of Language Development and the PLS. The children scored much higher on the PLS (adding an average 19 points to the language quotients). The correlation between the two scales was r=.70.

The following six studies involve comparisons of the PLS to the Peabody Picture Vocabulary Test (PPVT). The reader should be cautioned about a common error made in comparing PLS quotients to PPVT standard scores. A direct comparison of the mean PLS-ACQ to the mean PPVT Standard Score Equivalent (for a group of children) is not justified statistically. This is because PLS quotients are a different type of measure than PPVT standard scores. Standard scores from the PPVT

> express in standard deviation units the extent to which the subject's score exceeds, or falls below, the mean score of persons of the same age upon whom the test was standardized.... the PPVT-R, like many other standardized tests, uses a mean of 100 and a standard deviation of 15 (Dunn and Dunn, 1981, p. 91).

Quotients from the PLS do not compare a subject's performance to any normative group. Therefore, when looking at PLS quotients, one cannot assume a mean value of 100 or standard deviation of 15.

Comparing PLS-ACQ's directly to PPVT-R standard scores therefore amounts to comparing "apples and oranges". However, correlation values may be obtained between two types of scores; therefore, correlation values will be reported below.

Four comparisons of the PLS with the Peabody Picture Vocabulary Test (original edition) are cited in the PLS manual (Roston, 1977; Zimmerman and Steiner, 1971 and 1972; and Ward, 1970; all as cited in Zimmerman et al., 1979). PLS Auditory Comprehension Quotients (ACQ's) were compared to the Peabody Picture Vocabulary Test standard scores (called "Intelligence Quotients" by the original edition of the PPVT). Correlation coefficients ranged from a low of .26 in one study to a high of .66 in another.

Zimmerman and Steiner (1970), two of the PLS authors, presented a paper on the validity and reliability of the PLS. They administered the PLS, Peabody Picture Vocabulary Test (original edition), and another test to two successive classes of Head Start children. A total of 174 children were tested. Means and standard deviations were given for the PLS quotients and the Peabody Picture Vocabulary Test standard scores. The results from this study are reproduced as Table 5 (page 71). Zimmerman and Steiner (1970) noted that the PLS measures a wider range of language abilities than does the Peabody Picture Vocabulary Test.

In another study involving Headstart children, Zimmerman and Steiner (1971) obtained correlation values between .16 and .73 when comparing Peabody Picture Vocabulary Test standard scores to PLS-Auditory Comprehension Quotients. The results are reproduced as

Table 6 (page 72).

Lass and Golden (1975) investigated whether or not the PPVT could be considered as a test of general receptive language ability, by comparing it to the Token Test and to the PLS-AC. They (1975) administered the PPVT and the PLS to 24 children with speech and language disorders, ages 2-8 to 8-8. The correlation between the tests was .72. Study of their scatterdiagram reveals that three of the four children with lower scores (20-40 months) achieved much higher scores on the PLS-AC than on the PPVT. (Children with language ages higher than 60 months placed about the same on the two tests.)

In another study, Dodge (1980) used the PLS and the language subtest of the Denver Developmental Screening Test (DDST) to screen 486 preschool children. Children failing the screening were given a more extensive language assessment. The author found that the PLS and DDST were roughly equivalent in correctly identifying children who actually needed remediation (positive hits 75-80%). Both tests had a number of false positives, children failing the screening, but found later to have adequate language skills. Since more training is needed to administer the PLS than the DDST, Dodge recommended that the DDST be used alone (without the necessity of using the PLS) for the purpose of screening preschoolers.

Berryman (1983) administered the PLS to 672 preschool children, ages 3-8 to 5-4. Berryman was checking to see if items on each subtest were arranged in order of difficulty, and to see if the AC (receptive) and VA (expressive) items at each age level were of the same difficulty. Overall, Berryman found misplacement of five

individual items on the PLS. In addition, on the verbal ability scale, one 4-item section appeared to have been misplaced; more children passed the 4-item block of questions for 4- to 4 1/2-year-olds, than passed the (supposedly easier) four-item block for 3 1/2- to 4-year-olds. Berryman also found moderately high correlations ($r=.72$, $p<.001$) between the AC and VA subtests, with no significant differences between subtest means.

In Berryman's study, the preschoolers as a group achieved mean scores of 4 years, 10 1/2 months on both the VA and AC subtests. These scores are slightly higher than the mean chronological age (4 years, 4 months). Many children (40% on the AC subtest, 59% on the VA subtest) were administered all of the test items, because they kept passing at least 1 item in a four-item block. Berryman concluded that the PLS may be used with older preschoolers to compare expressive versus receptive language abilities for a child, but "the use of the PLS to classify children as language-delayed or to assign language ages is, as yet, a highly questionable procedure because normative data for the scale are not available" (Berryman, 1983, p. 84).

In summary, PLS age scres are not as yet validated by normative studies. Only a few inter-test comparisons have been done, but none have involved a comparison of age scores; and none have involved the SICD or the newly-revised PPVT-R.

ID. PLS Reliability. Split-half reliability was assessed by the PLS authors on two consecutive classes of Headstart children (Zimmerman and Steiner, 1970). Reliability coefficients (Spearman-

Brown formula) were .75 and .92. No short term test-retest reliability data were given in the PLS test manual.

II. Peabody Picture Vocabulary Test-Revised (PPVT-R).

This is a test of receptive vocabulary for ages 2 1/2 to 40. It yields raw scores that can be converted to age equivalent scores and standard scores. The PPVT-R is used by speech and language clinicians to help determine a general level of receptive language functioning. It is also used as a more in-depth check for receptive vocabulary skills.

IIA. PPVT-R Test Administration. The child (or adult) is shown a set of four black and white line drawings, printed on one page of a test booklet. The examiner says, "Show me (bird)", and the child's task is to point to, or otherwise indicate, the correct picture. The examiner begins testing at the child's estimated vocabulary level (usually according to chronological age), and must work either forward or backwards to get a basal of eight consecutive correct answers. Vocabulary items become more difficult as the test progresses. When the child fails 6 out of 8 consecutive items, the test is discontinued. Time required to administer the test is 20-30 minutes. Two alternate forms (called L and M) of the test are available. The test is very easy to administer.

IIB. PPVT-R Test Design and Norms. The original PPVT was published in 1959. It has been extensively studied by other researchers. The revised edition appeared in 1981. The original PPVT was standardized on a sample of white children from around

Nashville, Tennessee (Dunn and Dunn, 1981, p. 47). The revised edition (PPVT-R) was standardized on a much broader sample, in which (for children) age, sex, geographic area, parent's occupation, community size, and race were representative of the population of the United States according to the 1970 census (Naglieri, 1981). In addition, many pictures were changed in the PPVT-R, to eliminate cultural, racial or sex bias. Twenty-five plates were added to the test length, to provide a more sensitive measure (Dunn and Dunn, 1981, pp. 1-2).

IIC. PPVT-R Validity. The original PPVT was used as a measure of general intelligence, though it directly measures only one aspect of scholastic aptitude, vocabulary (Dunn and Dunn, 1981). According to the test authors (p. 59):

> many studies investigating the measurement of intelligence have shown that vocabulary is the best single type of test for predicting school success.

The original PPVT called its results "mental age" and "intelligence quotient", whereas the revised version calls these same scores "age equivalent" and "standard score equivalent."

Used as an intelligence test, the PPVT has been compared with the Stanford-Binet Intelligence Scale, the Wechsler Intelligence Scale for Children, the Wechsler Preschool and Primary Scale of Intelligence, and others. The reader is referred to the PPVT-R test manual for correlation values (Dunn and Dunn, 1981, p. 63). The test authors judge the PPVT to correlate "moderately well" with other intelligence tests (Dunn and Dunn, 1981, p. 67).

Used as a scholastic achievement test, the PPVT has been compared to the Peabody Individual Achievement Test (PIAT), Wide Range Achievement Test (WRAT), California Achievement Tests (CAT) and Metropolitan Achievement Tests (MAT). The PPVT-R test manual summarizes that the PPVT:

> correlates to a reasonable degree with measures of school achievement administered concurrently, but does less well as a predictive measure of school success (Dunn and Dunn, 1981, p. 68).

Used as a test of vocabulary, the PPVT correlates strongly with many other measures. A tabulation of 55 studies resulted in a median correlation value of .71. A summary of these correlations is given in the test manual (Dunn and Dunn, 1981, p. 62).

Many speech and language clinicians use the PPVT-R in language testing to provide an estimate of a child's receptive language age. However, the PPVT-R taps only one aspect of receptive language, knowledge of vocabulary. Other receptive abilities, such as the ability to follow verbal instructions, discriminate consonant sounds, and understand grammatical constructions, are not directly tested. In spite of this, there is some evidence (Teasdale, 1969) that the PPVT score gives a good estimate of general receptive language ability.

Teasdale (1969) administered the PPVT (original edition) and the Illinois Test of Psycholinguistic Abilities (ITPA) to 81 lower socio-economic status first grade children. The ITPA includes 12 subtests which purport to measure many skills underlying language. Among the ITPA subtests, the two highest correlations with the PPVT were found on the auditory-vocal association (.66) and auditory-

vocal automatic (.58) parts. This is not surprising, since the comprehension of auditory stimuli is required for both these subtests, in common with the PPVT. However, a more interesting finding is that the PPVT correlated most highly with the ITPA total score (.74), not with any individual ITPA subtest. Teasdale concluded that the PPVT is a reasonable indicator of general language abilities.

Sommers et al. (1978) reviewed six concurrent validity studies using a variety of language tests. In most instances, tests of receptive language abilities correlated well with each other, and tests of expressive language abilities correlated well with each other. Tests of just one receptive ability (such as comprehension of syntax) correlated well with broader tests of receptive abilities.

Sommers et al. (1978) also reported on their own investigation of 122 preschoolers with learning disabilities related to minimal brain dysfunction. They compared the childrens' performance on the PPVT to their performance on the Test for Auditory Comprehension of Language (TACL, a test covering a broad range of receptive language abilities), and with the Northwestern Syntax Screening Test (a test of both receptive and expressive syntax). The investigators found that the PPVT correlated well with both other receptive-language tests (TACL, r=.739, NSST-receptive portion, r=.704) and less well with the expressive language portion of the NSST (r=.518).

In a review of the original PPVT, Darley (1979, pp. 49-52) stated that this test may be used safely as a measure of receptive vocabulary. However, according to Darley, there has not been enough research comparing the PPVT to more general measures of language

ability. In addition, little is known about the predictive value of the PPVT, when it is used to assess a child's general level of functioning. Therefore, it is risky to use the PPVT as a basis for recommending remediation.

Costello and Ali (1971) administered the PPVT and ITPA to 67 low-income preschool children. They also collected teacher ratings of verbal abilities and psychiatrist ratings of quality of verbal exchange. Statistically significant but relatively small correlations were found with the teacher and psychiatrist ratings (.39 and .43, respectively). Correlations with two subtests of the ITPA were also statistically significant but relatively small (vocal encoding, .28; auditory-vocal association, .39). Comparisons with the Stanford-Binet were also listed. The investigators concluded that the PPVT shouldn't be used in isolation for the purposes of language or intellectual assessment.

To summarize, the PPVT-R is narrower in scope than the SICD or PLS (among other language tests), since only receptive vocabulary is tested. This author agrees with Darley (1979) that not enough research has been done comparing the PPVT to tests of broader receptive language abilities. Furthermore, there are currently no comparisons of the newly revised PPVT-R with other language tests.

IID. PPVT-R Reliability. In the PPVT-R normative study, correlation coefficients were broken down by age group. The split-half reliability coefficient for 2 to 4-year-olds ranged from .67 to .82, depending on the exact age group and on the alternate form (L or M) used (Dunn and Dunn, 1981, p. 54). Immediate retest (within

9 days) with an alternate form (L or M) resulted in a correlation of .82 or .83 for raw scores and .76 or .79 (depending on exact age group) for standard score equivalents. Delayed retest (9-31 days) with an alternate form resulted in a correlation of .52 or .78 for raw scores and .54 or .77 for standard score equivalents. The lower correlation is for ages 2-6 to 2-11; the higher correlation is for ages 3-0 to 3-11 (Dunn and Dunn, 1981, p. 56). The PPVT-R manual notes that test-retest reliability scores are lower for preschool children as a subgroup (Dunn and Dunn, 1981, p. 58).

In the study mentioned earlier, Costello and Ali (1971) administered the PPVT to 67 low-income preschool children. They found that the short-term (two-week) test-retest reliability of the PPVT could be increased from .77 to .86 by modifying the test presentation. Test items were arranged randomly, rather than in order of difficulty, and a reinforcement schedule was used.

III. Sequenced Inventory of Communication Development (SICD)

This is a test of general receptive and expressive communication abilities for children ages 4 months to 4 years. Its purpose is to inventory a wide range of communication abilities including language skills, in order to establish both a receptive communication age and an expressive communication age score (given in 4 month intervals). Items are placed on the SICD at the age level at which 75% of normal language-users would pass it. In the normative study for the SICD, most children "scored from -1 to +1 age level from their chronological age" (Hedrick et al., 1975, p. 10). In practice, this means that a

score of 3-4 is considered normal for a 3-8-year-old, but a score of 3-0 should raise some concern. No percentiles or standard scores are derived from this test. Because of its length, the SICD is used for longer evaluations but not for screening purposes, as mentioned earlier.

IIIA. SICD Test Administration. Verbal and motoric responses are required. Many of the items for children under age 2 are obtained through parent report. Tasks for older children include block stacking, imitation of sounds and words, question-answer, following directions, pointing to pictures, putting objects into a box, etc. During testing, the examiner pulls out one set of toys or pictures after another. For this reason, the SICD is a good choice for young children with short attention spans and a need for "hands-on" play. Because of the variety of types of response, the SICD may be used to pinpoint the child's strongest response modalities.

The scoring profile for the SICD enables the examiner to categorize test items by type of skill examined. For the receptive portion of the test, skills include awareness of sounds, awareness of speech, discrimination of sounds and speech, and understanding of words. The expressive portion of the test examines motor, vocal and verbal imitation skills; communication-initiating behaviors; vocal and verbal responsiveness; and quantity and quality (description) of verbal output.

The receptive portion of the test is given, followed by the expressive portion. Test items are generally arranged according to difficulty. A basal of 3 consecutive errors is required, and testing

is discontinued after 3 consecutive errors are made. This is a general guideline. Since some items are not in a strict order of difficulty on the SICD, examiners are asked to sample all items of the same level of difficulty at the estimated basal and ceiling (Hedrick et al., 1975, p. 28). For children above age 2, a 50-utterance language sample is taken at the conclusion of the expressive subtest, according to procedures outlined in Diagnostic Methods in Speech Pathology (Johnson, Darley, and Spreistersbach, 1963). Mean length of response is computed. The presence or absence of specific syntactical structures is also scored. These measures are included in the calculation of an expressive language age.

Finally, an articulation test may be given. However, this test is optional, and the results are not figured into the expressive language age score.

Test administration requires considerable practice and training (Lamberts, 1978). Quite a bit of dexterity is required to quickly locate the correct toys or pictures, administer the item, observe the child, write down the response, and produce the next toys.

IIIA(1). Detail on the administration of the language sample, SICD. Fifty utterances are collected. The examiner provides pictures or toys, and allows the child to select an item of interest. Open ended questions may be used to elicit language, such as "I wonder what you play with at home?" or "Tell me what they're doing here." (Johnson, Darley, and Spriestersbach, 1963, p. 165). The first 10

utterances are not used. Within the sampling period, the examiner may skip utterances (especially if hand-recording is used). Unintelligible utterances are not used. Utterances with single unintelligible words may be used, with a phonetic approximation of what was said. The examiner records his/her own eliciting behaviors as well as the child's words. Guidelines for delineating "one utterance" and for computing the MLU (Mean Length of Utterance) are taken from Johnson, Darley and Spriestersbach (1963, pp. 167-169).

Since these instructions are fairly broad, this author looked for additional guidelines for sampling with younger children. Bloom and Lahey (1978) state that younger children are best stimulated by toys, especially toys that involve construction activities. Bloom and Lahey also suggest that the examiner who uses a tape recording repeat the child's utterances during sampling, as long as the child is not distracted or overly interested in the repetitions. This makes transcription easier.

IIIB. SICD Test Design and Norms. The SICD was published in 1975. The test authors designed original test items and adopted others from existing tests (Hedrick et al., 1975). Some items examine pre-linguistic social and cognitive behaviors, as well as perceptual discrimination abilities. Others directly examine language use and understanding. No particular theory of language development was used to construct the test. Items were chosen for their ability to accurately pinpoint a child's communication age (Darley, 1979).

The test indicates general areas of deficit (such as cognitive deficits or perceptual deficits). However, its main diagnostic value is in helping to establish a receptive and expressive communication age (Lamberts, 1978).

An experimental version of the SICD was tested on 82 children to assess ease of administration, inter-observer reliability, and whether or not each test item was sensitive enough to pinpoint an age level. For example, items which were attained very slowly from the ages 4 months to 4 years were eliminated. Typically, included test items were acquired by 25-49 percent of the children at one age level, by 49-74% of the children at the next age level (4 months older), by 75-89% of the children at the next age level, and by 90+% at the next age level. Pages 10-15 of the test manual give a breakdown of each test item according to this format (Hedrick et al., 1975). The table indicates how slowly or quickly each item is acquired by what percent of children at each age level.

The normative sample for this initial edition was small. A total of 252 children were tested, with 21 children at each age level. The sample was balanced for socio-economic status (SES) but not for race, sex, or locality. All children were Caucasians living in the greater Seattle area. Children with ear or hearing difficulties, bilingual children, or children with known language difficulties were excluded (Hedrick et al., 1975).

IIIC. SICD Validity. The test authors compared SICD test scores to the PPVT (original edition), and obtained a correlation coefficient of .8097 for the SICD receptive score (SICD-RCA) and .7553 for the

SICD expressive score (SICD-ECA). The test authors concluded:

> The PPVT is expected to be grossly related to the RCA, a measure which encompasses more skills than single word vocabulary. The correlation of .8097 between RCA and PPVT reflects this gross agreement. It seems sufficiently low, however, to avoid the prediction of one score based on the knowledge of the other (Hedrick et al., 1975, p. 21).

Validity of the age scores is supported by the fact that many test items are adapted from other well-established instruments (Hedrick et al., 1975). In addition, the 252 subjects (normal language users) tested for the normative study placed at or very close to their chronological ages. Mean age scores and standard deviations for each age group are listed in Table 7 (page 73) (from Hedrick et al., 1975, p. 19).

A recent study of convergent validity was performed by Allen et al. (1981). They examined 182 3-0 to 3-11-year-old preschoolers, using the SICD, the Test for Auditory Comprehension of Language, TACL (Carrow, 1973) (a test of a broad range of receptive language abilities), and the Carrow Elicited Language Inventory, CELI (Carrow, 1974) (an expressive language test using sentence imitation). Examiners also made a clinical judgment of each child's language abilities, based on a checklist they designed for themselves. The list helped the clinicians to consider the child's behaviors during testing and to weigh the child's overall performance in light of his/her clinical experience. A judgment of "normal" or "language-impaired" was made for each child. For the purposes of their study, children placing at 12 months or below their chronological ages on the SICD, TACL, or

CELI were labeled language-impaired.

Results for this study are reporduced as Table 8 (page 74) (from Allen et al., 1981). The clinician's judgment was the milestone against which the other measures were compared. The SICD made 148 "hits" and 23 "misses". "Misses" can be broken down into 7 false positives and 16 false negatives. Twenty-six children who took the SICD were judged by the clinicians to be language-impaired. Of these, only 10 were correctly identified as impaired by the SICD. It should be remembered, however, that Allen and Bliss picked a 12-month delay as a cut-off point; the SICD manual does not specify a cut-off point, but its normative data suggest deviance when scores are 8 months below the chronological age.

A substantial number of misjudgments (false positives and false negatives) were identified for all three objective tests. The authors concluded that:

> the evaluation of impaired language behavior among three year olds is more art than science at the present time (Allen et al., 1981).

In summary, the SICD normative sample was relatively small (82 children), pointing to a need for further studies with a variety of samples of children. In addition, few inter-test comparisons have been conducted, and none of these has focussed on language age assignments.

IIID. SICD Reliability. Test-retest reliability was examined by the SICD test authors, with 10 subjects representing 6 age levels. Children were tested one week apart. "The mean percent of agreement across the ten subjects was 92.8 percent and the range was 88 to 98.6" (Hedrick et al., 1975, p. 20).

METHODS

Tests

The author gave all three tests (PLS, PPVT-R Form L, and SICD) to 30 preschool children, and compared each child's scores on one test with his/her scores on the other two tests.

Subjects

Subjects were 30 children from Linn and Benton County day care centers and preschools, ages 2-7 (2 years, 7 months) to 3-8. The three tests cover different age ranges, but they all included the age span from two to four years. This age range was restricted slightly when subjects were chosen, so that scores would be obtained for children performing a little above or below their chronological ages.

Day care center directors were contacted and asked to give written information (See Appendix B) regarding this study to parents of all children ages 2 1/2 to 3 1/2. Children with known hearing problems and children from bilingual homes were excluded. Four to six children were examined at each day care center. If more than six written consent forms were returned, the examiner randomly selected six children for testing.

Site of Testing and Order of Test Presentation

Each preschool provided its own testing room. Testing site did not vary for a given child; in other words, Child #3 was always tested

in the same room. Test order was counterbalanced among subjects. This was done to prevent influencing the test results because of a child's increasing familiarization with the testing procedures, examiner, etc.

The examiner performed all of the testing according to standardized instructions given in each test manual. A 50-utterance language sample was required as part of the SICD. The examiner recorded this sample in two or three parts, during the initial 10-15 minutes of each testing session. This time was also used as a warm-up period to get the child accustomed to the testing room and examiner. The child was invited to play with the examiner using a variety of toys spread around the floor. The conversation was tape recorded. The examiner used sentence repetition or paraphrases often, both to aid in later transcription and (hopefully) to increase the child's verbal output. After about ten minutes of play, the examiner invited the child to sit at a table and begin one of the formal tests. Testing sessions varied from 45 minutes to one hour.

Most of the children completed the testing in three sessions. In five cases, where children were showing signs of fatigue after having completed the first half of the PLS (Auditory Comprehension subtest), testing was discontinued that day, and the second half (Verbal Ability subtest) was given on another day. All testing for a given child was completed within a two-week period.

Transcription of the language sample was made on the same day as the recording. The first ten utterances for each day were not used. Utterances were counted on the testing day (to see if 50 had

been collected), but not otherwise scored (MLU, etc.) until all testing was completed on a given child. Likewise, formal test responses were recorded item by item but not added or charted until all testing was completed on a child. This was done in order to reduce examiner bias regarding a given child's performance.

RESULTS

Description of Sample

Thirty children were tested; 21 males and 9 females. Twenty-eight children (93%) were Caucasians, with one child of Asian origin and one child part Native American. Seventeen children (57%) were urban and 13 (43%) lived in rural settings. Three single-parent families were represented (10%).

Most (51) of the 57 parents had completed some college work. Thirty-three parents were college graduates. All parents had completed at least part of high school. The educational breakdown of the parents is given in Table 1.

Table 1

Education Completed by Parents

	Number of Parents	Percent
part of high school	1	2
high school graduate	5	9
some college-level work	18	32
college graduate	12	21
post-college grad or professional training	21	37
Total	57	101

Occuaptions of the parents are listed by category in Table 2.

Table 2

Occupations of Parents

Professional, technical	40%
Managers, administrators	12%
Sales	2%
Clerical	11%
Crafsmen, foremen	5%
Operatives	2%
House persons	14%
Students	12%
Unknown	2%

More detail on parent occupations is provided in Appendix C.

In summary, the sample consisted of mostly-Caucasian children with highly-educated parents. White-collar occupations were over-represented. All children were involved in day care or parent education programs. Boys were over-represented in the sample. Characteristics of the sample and their bearing on the results are examined in the Discussion.

Test Results Re: Equivalence of Scores

The first two questions addressed by this study were:

1. Are these tests equivalent in their estimates of language ages?
2. Is any of the three tests so similar to another, that one could be used in place of the other?

According to the results of this study, the three tests yielded scores that were not equivalent, with the PLS giving the highest scores, PPVT-R giving the next-highest scores, and SICD giving the

lowest scores. This result is elaborated below. The following test comparisons were made:

Receptive Language Ages:

(1) PLS-Auditory Comprehension Age (PLS-AC)
versus
SICD-Receptive Communication Age (SICD-RCA)

(2) PPVT-R Age Equivalent (PPVT-R Age)
versus
SICD-Receptive Communication Age

(3) PLS-Auditory Comprehension Age
versus
PPVT-R Age Equivalent

Expressive Language Ages:

(4) PLS-Verbal Ability Age (PLS-VA)
versus
SICD-Expressive Communication Age (SICD-ECA)

Three methods were used to make the comparisons: histograms, a comparison of mean scores, and a t-test for paired score differences.

I. Histograms

The following histograms (Figures 1-4) compare each subject's scores on one test to his/her scores on another test. These figures demonstrate that in most cases, children may be expected to perform better on the PLS than on the PPVT-R, and better on the PPVT-R than on the SICD. For example, in Figure 1, comparing each child's PLS-AC scores to his SICD-RCA scores, the bar marked zero shows that none of the children scored the same (within four months)[1] on the PLS-AC and the SICD-RCA. Everything to the right of the zero-bar represents the percentage of children achieving higher scores on the PLS-AC, in this

[1] Within 4 months of zero counted as zero. A child scoring 3-3 on one test and 3-1 on the other would be counted in the zero-bar.

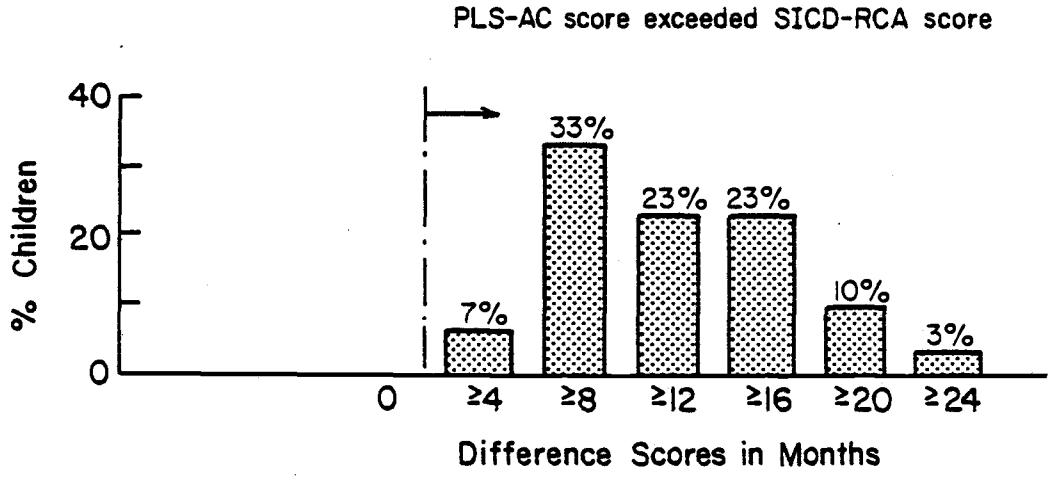

Figure 1. Distribution of PLS-AC minus SICD-RCA difference scores, in months. All children received higher age scores on the PLS-AC.

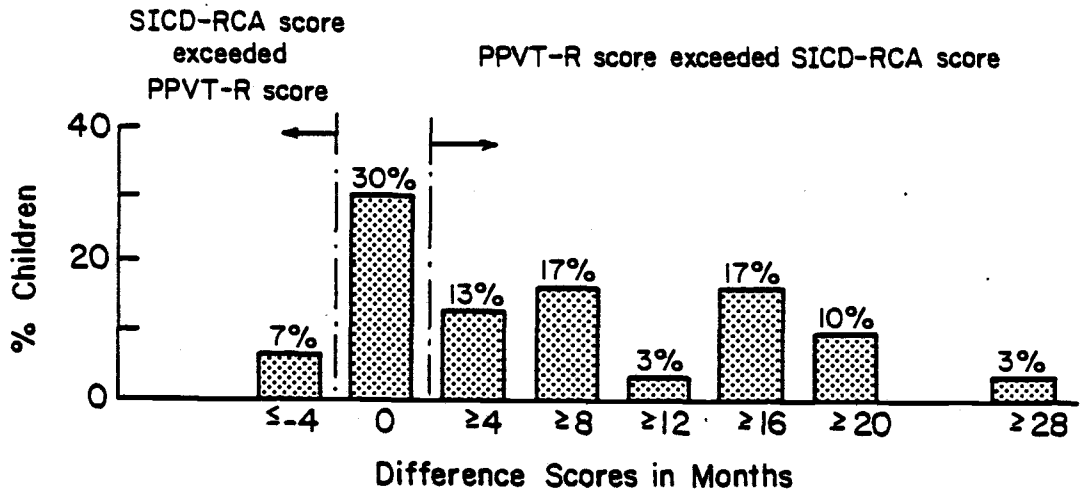

Figure 2. Distribution of PPVT-R minus SICD-RCA difference scores, in months. Sixty-three percent of the children achieved higher age scores on the PPVT-R (higher by at least 4 mo.)

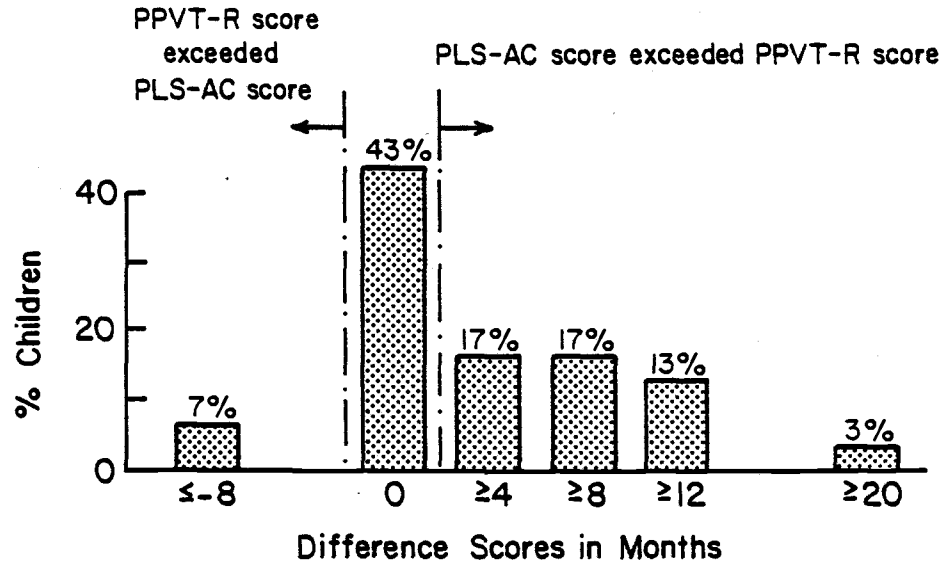

Figure 3. Distribution of PLS-AC minus PPVT-R difference scores, in months. Half of the children received higher age scores on the PLS-AC. Forty-three percent of the children achieved scores that were similar (within 4 mo.) on the two tests.

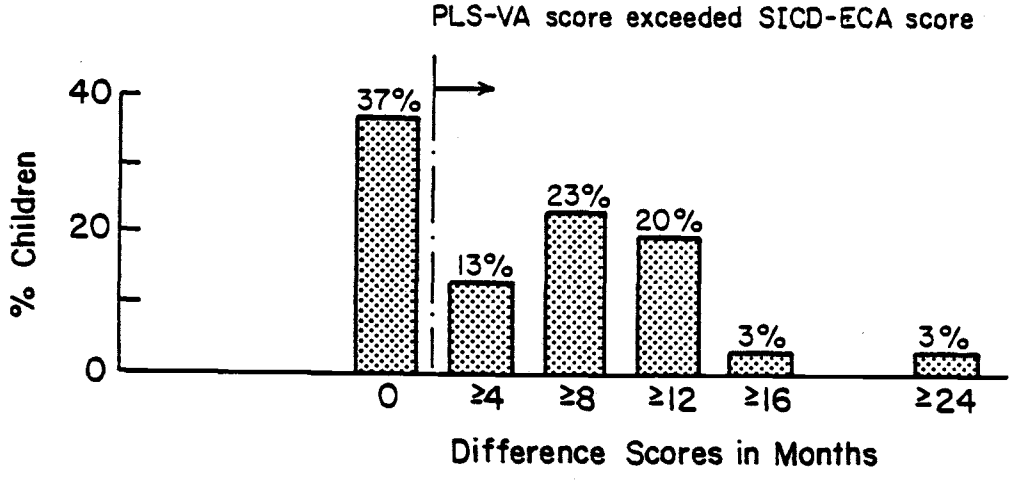

Figure 4. Distribution of PLS-VA minus SICD-ECA difference scores, in months. Sixty-three percent of the children achieved higher age scores on the PLS-VA (higher by at least 4 mo.)

case 100%. Everything to the left of the zero-bar would represent higher scores on the SICD-RCA. None of the children performed better on the SICD-RCA than on the PLS-AC.

One problem encountered when making these comparisons was that the PPVT-R gives age scores in one-month increments, the PLS in 1 1/2 month increments, and the SICD in four month increments. The author decided to portray difference scores in four month increments. A child scoring 3-0 on one test and 3-7 on another would be counted as having a difference score of four months (in one direction or the other), whereas a child scoring 3-0 on one test and 3-8 on the other would be counted as having a difference score of eight months. In other words, difference scores were rounded down to the nearest four-month multiple.

Figures 1-4 are graphic presentations which permit the reader to draw his/her own conclusions about how dramatic the differences were. The following summaries emphasize how different the tests were:

(1) Figure 1 (already described) shows that all of the children achieved higher scores on the PLS-AC than on the SICD-RCA, by at least four months.

(2) Figure 2 shows that for 93% of the children, PPVT-R Age scores were equal to or greater than SICD-RCA scores. Sixty-three percent of the children performed better on the PPVT-R by at least

four months.

(3) Figure 3 shows that for 93% of the children, the PLS-AC scores were equal to or higher than the PPVT-R Age scores. Fifty percent of the children performed better on the PLS-AC by at least four months.

(4) Figure 4 shows that the PLS-VA yielded scores greater than or equal to the SICD-ECA for all of the children. Sixty-three percent of the children performed better on the PLS-VA by at least four months.

More conservative summaries follow, in which scores up to eight months apart (for one subject) are considered "roughly equivalent".

(1) Figure 1 shows that only 7% of the children achieved roughly similar scores on the PLS-AC and SICD-RCA. In all other cases (93%), children scored at least eight months higher on the PLS-AC.

(2) Figure 2 shows scores on the PPVT-R and SICD-RCA to be roughly equivalent for 50% of the children. The other 50% of the children all achieved higher scores (by at least eight months) on the PPVT-R.

(3) Figure 3 shows scores on the PLS-AC and PPVT-R to be roughly equivalent (within eight months of each other) for 60% of the children. Thirty-three percent of the children scored at least eight months higher on the PLS, but two children (7%) scored higher on the PPVT-R.

(4) Figure 4 shows scores on the PLS-VA and SICD-ECA to be roughly equivalent (within eight months) for 50% of the children. The other 50% of the children scored higher on the PLS-VA by at least eight months.

II. Comparison of Mean Scores

Mean scores provide only a rough comparison of the tests, since each mean score lumps all 30 children together. In other words, the mean scores do not compare Subject John Doe's scores against his other scores. Mean scores are given in Table 3.

Table 3

Mean Scores

Test	Mean Age Score
PLS-AC (Receptive)	4-5*
PLS-VA (Expressive)	4-2
PPVT-R Age (Receptive)	4-0
SICD-ECA (Expressive)	3-7
SICD-RCA (Receptive)	3-3
Mean Chronological Age	3-2

Test	Mean Quotient
PLS-ACQ (Receptive)	139.24

Test	Mean Standard Score
PPVT-R SSE (Receptive)	114.73

*4 years, 5 months

There are two striking results. One is that, taken as a group, the children scored above their chronological ages on all tests. Possible reasons for this are sample bias, ideal testing conditions, or test design (See Discussion). The overall high scores are of interest and may be kept in mind; however, the inter-test comparisons are the core of this study. The other result of note is the position of each subtest in relation to the other subtests. Again, it is shown that the highest scores were obtained on the two subtests of the PLS. The PPVT-R yielded the next highest scores. The SICD subtests yielded the lowest scores (relative to the PLS and PPVT-R).

III. t-test for Paired Score Differences

The paired t-test (also called the dependent t-test or within-subjects t-test) is the appropriate choice for a within-subjects research design such as this one. The first step of the paired t-test is to compare Subject A's performance on one test with Subject A's performance on the second test, Subject B with Subject B, Subject C with Subject C, etc. In other words, information about one individual's relative performance is not lost. The t-test is applied to the difference scores thus obtained between two tests. A two-tailed t-test was used in this study.

Inter-test differences were significant ($p<.01$) for every comparison:

1. Children scored higher on the PLS-AC than on the PPVT-R.
2. Children scored higher on the PLS-AC than on the SICD-RCA.
3. Children scored higher on the PPVT-R than on the SICD-RCA.

4. Children scored higher on the PLS-VA than on the SICD-ECA.

Prediction of One Test Score from Another Test Score

The third question addressed by this study was:

3. If the three tests are not equivalent, can the examiner use a correction factor or some other means to predict a child's score on one test, from his/her score on another test?

In brief, the answer is no. Pearson product-moment correlation coefficients were calculated for each comparison and are presented in Table 4 (page 43). Spearman's rho's were also calculated to verify the approximately normal distributions of the data. All r's were significant, indicating that subjects achieving high scores on one test tended to achieve high scores on the other tests (Weinberg and Schumaker, 1962).

Regression equations were calculated for each comparison (See Table 4), in hopes that a clinician could add a number of months to a child's test score in order to predict another score. For an inter-test reliability study, the situation is analogous to that of comparing Forms A and B of a single test. In order to apply a formula to make predictions for an individual child, a very high degree of correlation ($r>.90$) is required (Cronbach, 1984). In other words, one would have to be sure that the highest scorer on the PLS was the highest scorer on the SICD, etc., in order to apply a formula. This condition was not met; all r's were less than .90. Therefore, while a clinician may anticipate a higher score on the PLS than on the SICD based on this study, estimates of actual

Table 4

Mean Differences, Correlation Coefficients and Regression Equations

	difference bet. means	signif. level (paired t)	Pearson r	Spearman's rho	regression equations
PLS-AC vs. SICD-RCA	13.73 mo.	p<.01	.67	.61	PLS AC = 1.06(SICD RCA) + 11.34
PPVTR-Age vs. SICD-RCA	8.83 mo.	p<.01	.46	.40	PPVTR Age = .98(SICD RCA) + 9.62
PLS-AC vs. PPVTR-Age	4.90 mo.	p<.01	.69	.67	PPVTR Age = .92(PLS AC) - .69
PLS-VA vs. SICD-ECA	7.48 mo.	p<.01	.82	.83	PLS VA = 1.36(SICD ECA) - 7.95
PLS-ACQ vs. PPVTR-SSE			.40	.41	
PLS-AC vs. PLS-VA	2.65 mo.	p<.05	.80	.79	PLS VA = 1.10(PLS AC) - 7.77
SICD-RCA vs. SICD-ECA	-3.60 mo.	p<.01	.51	.43	SICD ECA = .68(SICD RCA) + 16.20

scores are not warranted.

Intra-test Comparisons

The fourth question addressed by this study was:

4. Can children be expected to achieve similar scores in receptive and expressive language on the PLS? on the SICD?

According to this study, the receptive language subtest of the PLS may yield higher scores than the expressive language subtest. Figure 5 (page 45) shows that 47% of the children scored higher on the PLS-AC than on the PLS-VA by at least four months. Thirty-three percent achieved subtest scores within four months of each other, and 20% scored higher on the verbal ability subtest. The mean score for the PLS-AC was 4 years, 5 months, while the mean score for the PLS-VA was 4 years, 2 months. The paired t-test upheld this difference at $p<.05$.

A more conservative reading of Figure 5 shows that 79% of the children achieved scores on the PLS-AC that were within eight months of their scores on the PLS-VA. The significance of these findings is explored in the Discussion.

The opposite pattern is found for the SICD: here, the expressive communication subtest appeared to yield higher scores than the receptive communication subtest. Figure 6 (page 46) shows that 63% of the children scored higher on the SICD-ECA. Twenty percent achieved similar RCA and ECA scores (within four months); and 16% achieved higher scores on the SICD-RCA. The mean score for the SICD-RCA was 3 years, 3 months, while the mean score for the SICD-ECA

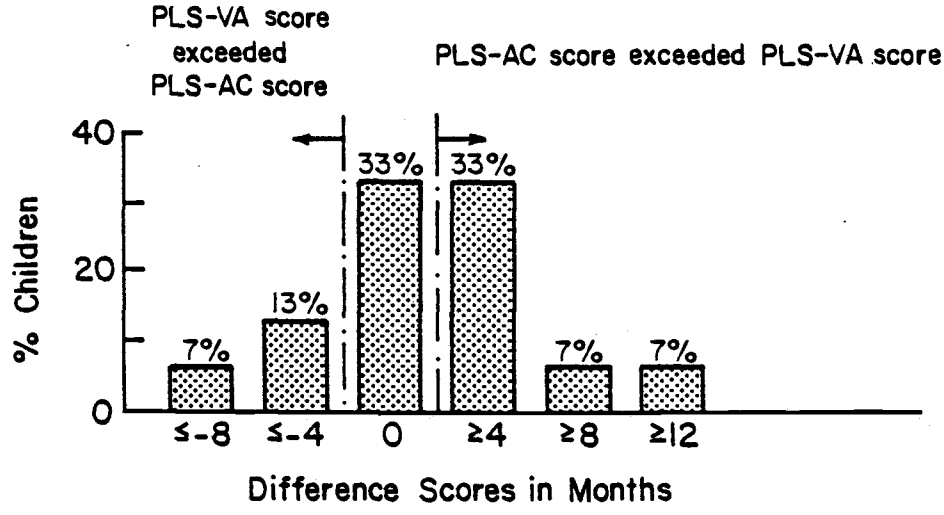

Figure 5. Distribution of PLS-AC minus PLS-VA difference scores, in months. Children tended to achieve slightly higher age scores on the PLS-AC (receptive subtest).

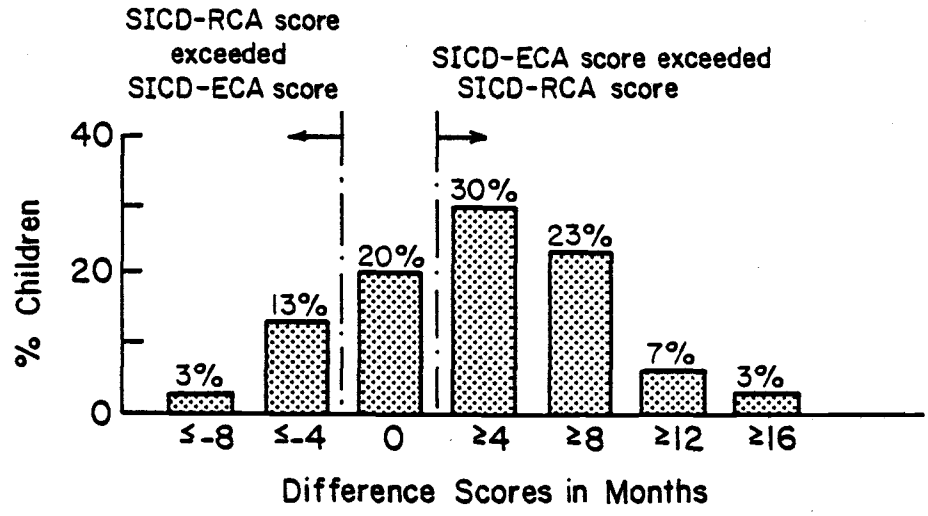

Figure 6. Distribution of SICD-ECA minus SICD-RCA difference scores, in months. Children tended to achieve slightly higher age scores on the SICD-ECA (expressive subtest).

was 3 years, 7 months. The paired t-test upheld this difference at $p<.01$.

A more conservative reading of Figure 6 shows that 63% of the children achieved scores on the SICD-RCA that were within eight months of their scores on the SICD-ECA. Still, 33% of the children scored higher on the ECA portion by at least eight months.

DISCUSSION

Three variables may have contributed to the overall-high test scores: sampling bias, ideal testing conditions, and test designs. Characteristics of the sample are examined first.

The investigator endeavored to obtain a sample that would represent this two-county area in the State of Oregon. She chose some day care centers generally reputed to serve upper and middle class children, as well as some serving middle and lower class children. In all, five day care centers/preschools were involved. In addition, the investigator involved one parent education class in which younger children enjoyed preschool-like activities with their parents, twice a week.

In most respects (race, rural/urban dwelling, occupation), the sample represents this two-county area fairly well. In terms of education, the two-county area from which this sample was drawn has a large proportion of highly educated persons (probably due to the presence of a university and community college). Occupations range widely, but there are many people employed in agriculture and trades (for example) who have college degrees. As it turned out, the sample included an even greater percentage of highly educated parents than does this geographic area.

The investigator also suggests that any study which selects children from day care centers rather than from the population of children at-large, is liable to include some bias. Children in

day care may receive greater language stimulation than other children. This was not a focus of the present study.

Finally, this study was not designed to test mostly boys (70% of the sample). Within each day care center, children who were the appropriate age for the study were selected randomly, as described earlier. As it happened, a majority of the children drawn "out of this hat" were boys.

To summarize, these were "typical preschoolers" for this area, except most were boys, and many children had highly-educated parents.

Overall-high test scores may also have been influenced by better than usual testing conditions. The examiner tested 25 children at their day care centers in rooms close to the regular classrooms. Five children who were enrolled in a parent education class (and play-time) for two-year-olds were examined at home. In a private communication, Elizabeth Prather, one of the authors of the SICD, wrote

> We [SICD authors] would not...expect that a representative sample of children would consistently score above their chronological ages. I would expect such a result to indicate either a class bias, or very possibly a 'testing within the preschool setting' where children are familiar and comfortable. Our normative sample all were brought to the Univ. of Wash. medical center for testing, and probably many thought they would be seeing the doctor shortly.

In the present study, children achieved scores closest to their chronological ages on the SICD. In this regard, it is interesting to note that the children in the present study resemble the children in the SICD's normative study in terms of geographic location, race, and (possibly) SES.

In summary, favorable testing conditions and sample biases may have been responsible for overall-high scores. However, it must be remembered that the focus of this study is on within-subject differences. The histograms compare Subject A's scores to Subject A's other scores. Each child experienced the same testing environment throughout the study. The order of tests was counterbalanced among subjects. Therefore, it should be of significance that most of the children scored higher on the PLS than on the SICD, etc.

How different are the tests, according to this study? Part of the answer depends on what we define as "disparate" scores for one child. The reader is referred back to Figures 1-4, in which each child's score on one test was subtracted from his own score on another test. For the histogram analysis, test scores were rounded down to the nearest four months, as described earlier. When scores four months apart were considered as disparate, then:

Comparison #1: All (100%) children scored higher on the PLS-AC than on the SICD-RCA. This is the most clearcut inter-test difference.

Comparison #2: Most (63%) children scored higher on the PPVT-R than on the SICD-RCA. Another 30% achieved "similar scores" (within four months) on the two tests.

Comparison #3: Half (50%) of the children scored higher on the PIS-AC than on the PPVT-R. Another 43% achieved similar scores on the two tests. This is the "weakest" demonstration of inter-test differences, as shown below.

Comparison #4: Most (63%) of the children scored higher on the PLS-VA than on the SICD-RCA. The other 37% all achieved similar

scores on the two tests.

Either the four-month or the eight-month difference score criterion could be justified. A consistent pattern of four-month or greater difference scores is a demonstrable inter-test difference. On the other hand, the eight-month difference score criterion may be argued for, on the basis of children's performances on language tests such as the SICD. For the SICD, "among the normative sample of children most scored from -1 to +1 age level from their chronological age" (Hedrick et al., 1975, p. 10). Since age levels are in four-month increments, a normal language-user could easily get a score four months above or below his/her chronological age (CA), but would be less likely to achieve a score eight months or more from the CA.

When scores need to be at least eight months higher (or lower) on one test than another in order to be considered "different", the effect is to widen the zero-bar of the histograms. Of course, the effect is to make the tests appear more similar. This picture emerges:

Comparison #1: Most (93%) of the children still scored higher on the PLS-AC than on the SICD-RCA. The other 7% achieved similar scores (within eight months) on the two tests.

Comparison #2: Half (50%) of the children scored higher on the PPVT-R than on the SICD-RCA. The other half (50%) achieved similar scores on the two tests. No children scored higher on the SICD-RCA. In this case, the trend is towards higher PPVT-R scores, but many children performed similarly (within eight months) on the

two tests.

Comparison #3: Just 33% of the children scored higher on the PLS-AC than on the PPVT-R, with most (60%) of the remaining children achieving similar scores (within eight months) on the two tests. In this case, we see that even though the scale tips towards higher scores on the PLS-AC, the majority of the children fall in the middle ground (where scores are within eight months on the two tests).

Comparison #4: Half (50%) of the children scored higher on the PLS-VA than on the SICD-ECA. The other half (50%) achieved similar scores on the two tests. No children scored higher on the SICD-ECA. The trend is toward higher PLS-VA scores, with no children achieving higher SICD scores, but in this case, too, a number of children performed similarly (within eight months) on the two tests.

To summarize, the histograms demonstrated the widest difference between the PLS-AC and SICD-RCA. In other inter-test comparisons, the gap between test scores was less dramatic, but in all cases one test tended to produce higher scores than another. In the paired t-test analysis, where scores were not rounded to four-month increments, all inter-test differences were significant at $p<.05$.

It is beyond the scope of this study to investigate why the inter-test differences were found. However, it is apparent that the tests differ in many respects: types of responses required, length of test, scope (the PPVT-R being a test of receptive vocabulary), etc. The overall purposes of each of the tests are different, though they all produce age scores and are all used as diagnostic

instruments by speech and language clinicians. It is also noteworthy that different samples of children were tested in the research studies underlying each test (See Introduction). Let it suffice to say that in terms of age scores achieved, the tests varied in difficulty.

The correlation coefficients listed in Table 4 (page 43) give some insight into how the tests varied in relative terms. Most of the correlations were between .4 and .8. These are similar to inter-test correlations obtained by Lass and Golden (1975) when comparing the PLS and PPVT; Hedrick et al. (1975) when comparing the SICD and PPVT; and Zimmerman and Steiner (1971) when comparing the PLS and PPVT. The above authors concluded that the tests they examined were measuring similar aspects of language behavior. However, correlations in this range (.4 to .8) are modest; they could be higher. This author believes that her modest correlations could be a result of any of the following three factors:

1) The tests were not measuring the same aspects of language behavior (to some extent).

2) Testing was not simultaneous; therefore the issue of test-retest reliability came into play. In this study, all children were tested within a two week period. Earlier, the author cited test-retest correlations of .83 and .82 (depending on exact age group) for children taking an alternate form of the PPVT-R within nine days of the first test (Dunn and Dunn, 1981, p. 56). The SICD authors obtained test-retest scores for 60 children, tested one week apart. The percent of agreement between tests was 90.41

for children two years and older. Scores from the second testing were higher, probably due to familiarity with the test and examiner, as well as teaching by the parent who observed the first testing session (Hedrick et al., 1975).

3) The narrow range of ages in this study makes it difficult to correlate scores between tests. This factor may have come into play when the two subtests of the SICD were compared (r=.43). The SICD yields scores in four-month increments. Scores ranged from 24 months to 48+ months (counted as 52 months for purposes of statistical analysis) on the ECA subtest, with eight data points to compare. The RCA subtest scores ranged from 28 to 48 months, with only six data points.

How do the inter-test results compare to those of other studies? The PLS manual (Zimmerman et al., 1979) described several studies comparing the PLS to the PPVT (original edition), however, comparisons were always in terms of quotients versus standard scores; no age score comparisons were available.

As reported earlier, Lass and Golden (1975) compared scores on the PLS to scores on the PPVT (original edition) for 24 children with speech and language disorders. Three of the four "youngest" children (scoring between 20-40 months in receptive language ages) performed better on the PLS than on the PPVT by at least 12 months, a finding which agrees with those of the present study.

Berryman (1983) studied 672 preschool children from a mid-western community, and found that they achieved PLS scores an average of 6 1/2 months above their chronological ages. This suggests that average

preschoolers may achieve high marks on the PLS. However, it is beyond the scope of the present study to assess the validity of the age scores, only their comparison to scores on other tests. The present study examined children who achieved relatively high scores on all three tests, and found significant discrepancies between them. It would be interesting to see if these differences held up for a) children suspected of having language problems or for b) children from different geographic regions or backgrounds.

This author found no studies comparing the SICD to the PLS, and only one study (Hedrick et al., 1975) comparing the SICD to the PPVT, original edition. As described earlier, the correlation was .81. No age scores were given.

This author wishes to point out that many inter-test studies report correlation coefficients without including mean, median, or difference scores. Therefore, the reader is unable to see if the actual range of scores (or means) is higher for one test than another. This author would recommend that future studies report scores in absolute as well as in relative terms.

Intra-test comparisons are discussed next. How different were the subtests of the PLS? of the SICD? As seen earlier in the inter-test comparisons, it depends on one's definition of "different". The "four-month criterion" highlights differences between tests, as follows:

Comparison #5: The largest percentage of children (47%) got higher scores on the PLS-AC than on the PLS-VA, with 33% of the

children achieving similar scores (within four months) on the two subtests.

Comparison #6: The majority of the children (63%) achieved higher scores on the SICD-ECA, with 20% of the children achieving similar scores on the ECA and RCA subtests.

But when an eight-month discrepancy between an individual's scores is required to define a score as "higher" or "lower", then the subtests appear more similar:

Comparison #5: The majority (79%) of the children scored similarly on the PLS-VA and PLS-AC subtests, and

Comparison #6: The majority (63%) of the children scored similarly on the SICD-ECA and SICD-RCA subtests, with 33% of the children scoring higher on the SICD-ECA, and 3% (1 child) scoring higher on the SICD-RCA.

Why might scores on the PLS-AC be higher than scores on the PLS-VA? The investigator noticed that during administration of the PLS, many of the children were able to get at least one item at each age level. According to the directions for administering the PLS, this means that the test must be continued. In practice, this meant that the test was lengthy (up to one hour). Many three-year-olds were tested through the 5 and 6-year-old levels (though they might miss 2 or 3 items at several levels). This investigator believes that, after working long and hard on the AC subtest, many children were tired for the VA subtest. This investigator would be interested in seeing a study in which the VA subtest was given first, compared with the standard procedure of giving the AC subtest first.

As mentioned earlier, Berryman (1983) gave the PLS to 672 preschoolers, ages 3-8 to 5-4. Berryman found equivalence between the two subtests, so her results differ from that of this study. This author believes that the two PLS subtests may, in fact, be equivalent. The discrepancy may be explained by the fact that Berryman's subjects were slightly older, so they were able to begin at higher levels of the test, possibly reducing test fatigue as a factor.

Why might children score higher on the ECA portion of the SICD than on the RCA portion? The discrepancy (though small) may be built into the test design. The SICD normative study indicated at most a very slight "ECA-advantage" for children ages 3-0 and 3-4, with practically identical RCA and ECA scores at ages 2-8, 3-8, and 4-0 (Hedrick et al., 1975; reproduced as Table 7. More recently, one of the SICD authors, Elizabeth Prather (personal communication) reported a large study by Allen and Bliss, which will appear in the revised SICD manual. Allen and Bliss's RCA ages for white children ages 2-8 and older averaged 1.4 months lower than those reported in the SICD manual, and ECA ages tended to be about 2 months higher. The present study seems to uphold this slight ECA-RCA gap among 3-year-old, (mostly) white preschoolers.

CONCLUSION

The first two goals of the study were to determine:

1. Are these tests equivalent in their estimates of language ages

2. Is any of the three tests so similar to another, that one could be used in the place of the other?

The three tests yielded scores that were not equivalent, with the PLS giving the highest scores, PPVT-R giving the next-highest scores, and SICD giving the lowest scores. For purposes of determining a child's general level of language functioning, one test cannot be used in the place of another. In the author's opinion, the tendency of one test to produce higher scores than another should be kept in mind during test selection and interpretation.

The third goal of the study was to determine:

3. If the three tests are not equivalent, can the examiner use a correction factor or some other means to predict a child's score on one test, from his/her score on another test?

The answer is no. While all inter-test differences were significant at $p<.01$ (paired t-test), correlation coefficients were not high enough to warrant prediction of an individual child's score based on his/her other scores.

The fourth goal of the study was to answer the following:

4. Can children be expected to achieve similar scores in receptive and expressive language on the PLS? on the SICD?

It appears that the SICD-ECA subtest may produce age scores slightly above the scores for the SICD-RCA subtest. This study also

indicated a slight difference between scores on the PLS-AC and PLS-VA for this group of children (age approximately three), with the PLS-AC yielding higher scores. It is the author's opinion that test fatigue should be considered as a factor which may slightly depress PLS-VA scores in comparison to PLS-AC scores, when younger children are tested.

It is the author's opinion that future inter-test comparisons should include measures of central tendency (descriptions of the age scores, quotients and/or standard scores received) as well as correlation coefficients. Correlation coefficients estimate to what extent the highest scorer on Test A would be the highest scorer on Test B, etc., but give no estimate of actual scores received. Measures of central tendency allow the reader to see if scores on one test were the same, higher, or much higher than the scores on another test.

Topics for future studies might include:

1) a study investigating test length and fatigue as factors influencing younger children's PLS scores.

2) studies in a variety of geographic areas, etc., involving a large number of typical preschoolers, so that more information is available for both the PLS and SICD re: age score performances of various subgroups.

3) more inter-test comparisons of the (relatively new) PPVT-R with other tests of vocabulary or language development.

The most important findings of this study were that:

1) higher age scores were obtained on the PLS than on the SICD

or PPVT-R.

2) the SICD produced lower age scores than those obtained on either the PPVT-R or the PLS. In addition, a subject's SICD-ECA score was on the average one age level (4 months) above his/her SICD-RCA score.

3) the PPVT-R usually yielded receptive language age score estimates that were lower than PLS-AC scores but higher than SICD-RCA scores.

BIBLIOGRAPHY

Allen, Doris V., Bliss, Lynn S., and Timmons, Jack. Language evaluation: Science or art? Jour. Sp. Hear. Dis., 1981, 46, 66-68.

Berryman, Joan D. Use of the Revised Preschool Language Scale with older preschool children. Lang. Sp. Hear. Serv. Schools, 1983, 14, 79-85.

Bloom, Lois and Lahey, Margaret. Language Development and Language Disorders. New York: John Wiley & Sons, 1978.

Carrow, Elizabeth. Test for Auditory Comprehension of Language (TACL) Austin, TX: Urban Research Group of Educational Concepts, 1973.

Carrow, Elizabeth. Carrow Elicited Language Inventory (CELI). Boston: Teaching Resources Corp., 1974.

Costello, Joan and Ali, Faizunisa. Reliability and validity of Peabody Picture Vocabulary Test scores of disadvantaged preschool children. Psych. Reports, 1971, 28, 755-760.

Cronbach, Lee J. Essentials of Psychological Testing (4th ed.). New York: Harper & Row, 1984.

Darley, Frederic L. (Ed.) Evaluation of Appraisal Techniques in Speech and Language Pathology. Reading, MA: Addison-Wesley, 1979.

Dodge, Gordon R. A comparison of language screening methods. Lang. Sp. Hear. Serv. Schools, 1980, 11, 214-217.

Dunn, Lloyd M. Peabody Picture Vocabulary Test (PPVT). Circle Pines, MN: American Guidance Service, 1959.

Dunn, Lloyd M. and Dunn, Leota M. Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service, 1981.

Frankenburg, W.K., Dodds, J.B., Fandal, A.W., Kazuk, E. and Cohrs, M. Denver Developmental Screening Test (DDST) (Rev. ed.). Denver: LADOCA Project and Publishing Foundation, 1975.

Hedrick, Dona Lea, Prather, Elizabeth M., and Tobin, Annette R. Sequenced Inventory of Communication Development. Seattle: Univ. of Washington Press, 1975.

Higbee, N.L. A comparative study of the Preschool Language Scale and the Illinois Test of Psycholinguistics when administered to a population of cerebral palsied children. Unpublished Master's thesis, Univ. of Southern California, 1974.

Hutchinson, Barbara B., Hanson, Marvin L., and Mecham, Merlin J. Diagnostic Handbook of Speech Pathology. Baltimore: Williams & Wilkins, 1979.

Johnson, Wendell, Darley, Frederic L., and Spriestersbach, D.C. Diagnostic Methods in Speech Pathology. New York: Harper and Row, 1963.

Kirk, S.A., McCarthy, J.J., and Kirk, W. The Illinois Test of Psycholinguistic Abilities (ITPA) (Rev. ed.). Urbana, IL: University of Illinois Press, 1968.

Lamberts, Frances. Review of a test for communicative behaviors. Jour. Learn. Disab., 1978, 11, 28-29.

Lass, N. and Golden, S.A. A comparative study of children's performance on three tests for receptive language abilities. J. Auditory Res., 1975, 15, 177-182.

Naglieri, Jack A. Concurrent validity of the Revised Peabody Picture Vocabulary Test. Psych. in the Schools, 1981, 18, 286-289.

Payne, James S., Ball, Donald W., and Stainback, William C. Note on reliability and congruent validity of the Peabody Picture Vocabulary Test with disadvantaged preschool children. Psych. Reports, 1972, 30, 22.

Prather, Elizabeth, personal communication, March 28, 1983.

Proger, Barton B. Test review #5: Three informal preschool evaluation scales. Jour. Special Ed., 1971, 5, 85-91.

Roston, A. Peabody Picture Vocabulary Test and Preschool Language Scale results for preschool and kindergarten children. Unpublished paper, 1977.

Sanders, Louis J. Procedure Guides for Evaluation of Speech and Language Disorders in Children (4th ed.). Danville, IL: Interstate Printers and Publishers, Inc., 1979.

Scott, G. H. A comparison of the Utah Test of Language Development and the Preschool Language Scale. Unpublished Master's thesis, Illinois State University, 1973.

Siegel, Gerald M. The use of language tests. Lang. Sp. Hear. Serv. Schools, 1975, 6, 211-217.

Siegel, G.M. and Broen, P. A. Language assessment. In Communication Assessment and Intervention, ed. L. L. Lloyd. Baltimore: Univ. Park Press, 1976.

Sommers, Ronald K., Erdige, Susan, and Peterson, M. Kathleen. How valid are children's language tests? Jour. Special Ed., 1978, 12, 393-407.

Teasdale, G. R. Validity of the PPVT as a test of language ability with lower SES children. Psych. Reports, 1969, 25, 746.

Ward, D. A comparison of the Preschool Language Scale and the Peabody Picture Vocabulary Test. Unpublished paper, San Diego State College, 1970.

Weinberg, George H. and Schumaker, John A. Statistics: An Intuitive Approach. Belmont, CA: Brooks/Cole Publishing Co., 1962, 1969.

Zimmerman, Irla L. and Steiner, Violette G. Validity and evaluation of the Preschool Language Scale. Paper presented at the Western Psychological Assoc. Convention (Los Angeles, CA, April 15-18, 1970). Available on microfiche: ED 076 711. 9 pp.

Zimmerman, Irla L., and Steiner, Violette G. A comparative study of the Preschool Language Scale and the Peabody Picture Vocabulary Test. Paper presented at the convention of the Calif. Assoc. of School Psychologists and Psychometrists/Nat'l Assoc. of School Psychologists, San Francisco, March 26, 1971. (NASP Abstracts)

Zimmerman, Irla L. and Steiner, Violette G. A comparison of the PLS with the Columbia Mental Maturities Scale and the PPVT for Head Start children. Unpublished paper, 1972.

Zimmerman, Irla L., Steiner, Violette G., and Evatt, Roberta L. Preschool Language Scale. Columbus, OH: Charles E. Merrill Publ. Co., 1969, 1979.

APPENDICES

Appendix A.

List of Abbreviations

AC	Auditory Comprehension, the receptive subtest of the Preschool Language Scale (Zimmerman *et al.*, 1979)
ACQ	Auditory Comprehension Quotient, a score from the Preschool Language Scale
CA	chronological age
CELI	Carrow Elicited Language Inventory (Carrow, 1974)
DDST	Denver Developmental Screening Test (Frankenburg *et al.*, 1975)
ECA	Expressive Communication Age, a score from the Sequenced Inventory of Communication Development (Hedrick *et al.*, 1975)
ITPA	Illinois Test of Psycholinguistic Abilities (Kirk *et al.*, 1968)
MLU	mean length of utterance
PLS	Preschool Language Scale (Zimmerman *et al.*, 1979)
PLS-AC or PLS-A	Preschool Language Scale-Auditory Comprehension subtest. The abbreviation PLS-AC is used in this thesis to label the *age score* obtained by a child on the Auditory Comprehension subtest.
PLS-ACQ	Preschool Language Scale Auditory Comprehension Quotient
PLS-VA or PLS-V	Preschool Language Scale-Verbal Ability subtest. The abbreviation PLS-VA is used in this thesis to label the *age score* obtained by a child on the Verbal Ability subtest.
PPVT	Peabody Picture Vocabulary Test (original edition) (Dunn, 1959)
PPVT-R	Peabody Picture Vocabulary Test-Revised (Dunn and Dunn, 1981)
RCA	Receptive Communication Age, a score from the Sequenced Inventory of Communication Development (Hedrick *et al.*, 1975)

Appendix A., List of Abbreviations, continued

SES	socio-economic status
SICD	Sequenced Inventory of Communication Development (Hedrick et al., 1975)
SICD-ECA	Sequenced Inventory of Communication Development Expressive Communication Age
SICD-RCA	Sequenced Inventory of Communication Development Receptive Communication Age
TACL	Test for Auditory Comprehension of Language (Carrow, 1973)
VA	Verbal Ability, expressive subtest of the Preschool Language Scale (Zimmerman et al., 1979)
VAQ	Verbal Ability Quotient, a score from the Preschool Language Scale

Appendix B.

Parent Consent and Confidential Information Forms

PRESCHOOL LANGUAGE STUDY

PARENT CONSENT FORM

Dear Parent:

I have received permission from Oregon State University to conduct a study investigating 3 different language tests for preschoolers. Each of the tests is designed to give estimates of a child's ability to understand language, and the child's ability to express himself or herself. I will be interested in seeing whether or not the three tests give a consistent picture of each child's skills. This information is important to speech therapists and to educators. Hopefully, with the results of this study, speech therapists will be able to 1) report language test results to parents more accurately, 2) use language tests more accurately to rule out language troubles for children who have trouble making many speech sounds, and 3) more accurately determine which preschoolers need help for language development.

I am interested in testing typical preschoolers, most of whom will be normal language-users. Each child will, over a 2-week period, be tested with all 3 tests. The tests consist of activities which are geared for preschoolers, such as pointing to pictures and toys, following simple directions, imitating movements and sounds, naming objects, answering simple questions, and talking as they play with toys. Most children find these activities enjoyable.

I will meet with each child 3 times, for ½ hour to 1 hour each time. Testing will be done at the child's school or day care center. I will use a code number system to identify each child, so that the test results and the identification of the children remain confidential. No fees are involved.

Because this research will be of interest to other speech therapists and educators, I will need to show that I am testing a typical group of children from this area. For this reason, I am asking for some biographical information, such as parent occupation, race of child, etc. When I write up this study, I will not report this kind of information for individual children, but I will include a general description of the entire group of children. For example, I will (hopefully) be able to say that the children came from families with a wide range of parent occupations, matching the typical occupations for adults in this county.

At each school, I will test 6-9 children. If more parents than this are interested in having their children tested, then children's names will be drawn out of a hat. This means that your child may not be tested, even if you give consent for him or her to participate. If your child is not included, but you are especially interested in speech or language testing, you may contact Oregon State University's speech clinic at 754-2461 for information.

If you have any questions regarding this letter and your child's participation in the study, please call me at 745-7166. Your consent for your child to participate may be withdrawn at any time. Please check the appropriate boxes on the attached form and return the form to _____ at _____ as soon as possible.

Parents requesting test scores will be given an appointment with the investigator to thoroughly explain the results. Test results will be shown to teachers or day care providers only with your written approval (check the

Appendix B., Parent Consent Forms, continued

appropriate box on the next page).

We feel that this study will be beneficial to many children in the future and appreciate your approval for your child to participate. Thank you.

Sincerely,

Larky Hansen
Graduate Student
Speech Pathology
Oregon State University

Appendix B., Parent Consent Forms, continued

CONFIDENTIAL INFORMATION

FORM TO BE RETURNED

Child's name_____ Today's date_____

Parent's name_____

I have received a full and satisfactory explanation of the research study to be conducted by Larky Hansen, Graduate Student, Oregon State University, and understand that at no time will the identity of the participants be revealed. I further understand that I may discontinue my child's participation at any time. Finally, I understand that any information collected may be part of a research publication at some future date, but again, the identity of participants and schools will not be revealed.

☐ Yes. I consent to my child participating in this study.

☐ No. I do not wish my child to participate.

If you would like your child to participate, would you like to have an appointment after your child is tested, to discuss the results with the investigator?

☐ I do request an appointment time. My telephone number is_____.

☐ I do not request an appointment time.

If you would like your child to participate, would you like your child's teacher to receive a report of the results?

☐ Yes, I would like my child's results to be reported to his/her teacher.

☐ No, I do not want test results to be reported to his/her teacher.

If you would like your child to participate, please supply the following information:

	1st parent	2nd parent (if child is in 2 parent home)
Parent occupation:	_____	_____

(PLEASE CONTINUE ON NEXT PAGE)

Appendix B., Parent Consent Forms, continued

CONFIDENTIAL INFORMATION

	1st parent	2nd parent (if child is in 2-parent home)
Parent education. For each parent, circle one:	completed less than 7 grades	completed less than 7 grades
	completed 7 grades	completed 7 grades
	completed part of high school	completed part of high school
	high school graduate or equivalent	high school graduate or equivalent
	completed some college-level work	completed some college-level work
	college graduate	college graduate
	post-college graduate or professional training	post-college graduate or professional training

Number of children at home: _____

What position is this child? (oldest, 2nd child, etc.)_____

Birthdate of child _____

Sex of child _____

Race of child_____

In which city or town do you live? _____ Do you live in the country or in town?_____

Thank you for completing this information.

Appendix C. Occupations of Parents

Professional, technical--40%
- teachers-6
- research scientist
- nurse
- engineer
- waste water treatment
- counselor-educator
- research technician
- horticulturist
- design supervisor
- biochemistry research assistant
- cost accountant
- engineer tech.
- photographer
- spectrographer
- minister
- technical editor
- plant pathologist
- financial analyst assistant

Managers, administrators--12%
- builder contractor
- construction manager
- engineering manager
- marketing management
- store manager
- office manager
- systems administrator

Sales--2%
- sales clerk

Clerical--11%
- postal worker
- secretaries-2
- clerk
- sec'y-clerk-seamstress (clerical and operative)
- dept. asst. clerical

Craftsmen, foremen--5%
- millwright
- builder
- carpenter

Operatives--2%
- welder

House persons--14%

Students--12%

Unknown--2%

Table 5

Preschool Language Scale and Peabody Picture Vocabulary Test (from Zimmerman and Steiner, 1970)

	TEST Fall 1968		RETEST Spring 1969		TEST Fall 1969	
	M	SD	M	SD	M	SD
BOYS - N 33					N 37	
PLS-Auditory	82.2	17.59	102.4	16.37	88.3	15.7
PLS-Verbal	79.0	20.27	97.1	20.37	82.3	18.1
PPVT	74.1	18.21	90.7	20.71	87.3	14.8
GIRLS - N 19					N 47	
PLS-Auditory	96.0	14.49	117.7	15.78	89.2	17.5
PLS-Verbal	96.6	18.63	115.2	12.52	86.2	16.4
PPVT	87.2	16.71	102.6	14.18	80.1	17.0

Table 6

Peabody Picture Vocabulary Test and Preschool Language Scale Results for Head Start Classes (from Zimmerman and Steiner, 1971)

	Test Fall 1968	Retest Spring 1969	Test Fall 1969	Retest Spring 1970
N 52	M SD	M SD	M SD	M SD
PPVT	78.9 ± 18.75	95.1 ± 19.44	83.5 ± 16.77	100.8 ± 14.49
PLS-A	87.2 ± 17.79	108.0 ± 17.76	88.4 ± 15.93	108.0 ± 13.91
PLS-V	85.4 ± 21.43	103.7 ± 19.96	85.6 ± 16.55	104.6 ± 12.53
r PPVT PLS-A	.72**	.50**	.73**	.16
r PPVT PLS-V	.65**	.52**	.47**	.64**

N 82	Test Fall 1970
PPVT	82.2 ± 21.07
PLS-A	87.3 ± 17.77
PLS-V	83.3 ± 19.88
r PPVT PLS-A	.60**
r PPVT PLS-V	.62**

** Significant at the .01 level

Table 7

SICD: Mean Receptive Communication Age and Expressive Communication Age Scores and Standard Deviations for the 21 Subjects in Each Age Group

	RCA	Sd	ECA	Sd
4 mo	4.76	1.41	5.14	1.85
8 mo	8.76	2.05	7.23	2.41
12 mo	11.43	2.29	11.81	3.46
16 mo	15.05	3.32	16.57	3.85
20 mo	20.95	3.32	20.57	1.43
24 mo	22.73	2.86	22.73	3.12
28 mo	29.14	4.41	28.76	6.15
32 mo	31.81	5.29	31.81	4.98
36 mo	34.29	5.87	35.24	6.53
40 mo	38.48	5.29	41.71	6.27
44 mo	43.05	3.77	42.29	4.99
48 mo	43.81	6.87	43.81	5.29

Table 8

Relationships Between Test Decision (including SICD) and Clinical Judgment (of Language Impairment) (Allen et al., 1981)

Test	N	Test Decision	Clinical Judgment: Normal-Speaking	Clinical Judgment: Language-Impaired	X^2
CELI	148	N	122	5	
		I	8	13	51.38**
TACL	149	N	124	17	
		I	4	4	6.14**
SICD	171	N	138	16	
		I	7	10	24.23**

$p < 0.01$