5-13-1980

Comparison of Two Receptive Language Tests (PPVT and TACL) Used with the Developmentally Delayed

Paul H. Lamb
Portland State University

AN ABSTRACT OF THE THESIS OF Paul H. Lamb for the Master of Science in Speech Communication: Emphasis Speech Pathology and Audiology presented May 13, 1980.

Title: Comparison of Two Receptive Language Tests (PPVT and TACL) Used with the Developmentally Delayed.

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:

[redacted]

Robert L. Casteel, Chairman

[redacted]

Robert H. English

[redacted]

Mary E. Gordon

[redacted]

Ronald E. Smith

The purpose of this investigation was to compare the results as recorded in mental ages of the PPVT and TACL when used with developmentally delayed children. One aspect was to observe how well the mean mental age from each test would compare with the mean mental age obtained from psychometric testing (WISC-R or SB-LM results). Another aspect was to determine how well the data from the

PPVT and from the TACL would correlate with the psychometric testing results.

The subjects were seventeen children whose IQ's were above 50 but below 85 as determined by previously administered SB-LM or WISC-R testing. A current mental age was extrapolated for each subject from the IQ or scaled-score results of this testing. The examiner screened for hearing and administered the PPVT and TACL to each subject. The subjects' current mental ages were over 3 years-0 months, but not over 7 years-0 months when tested with the PPVT and TACL. The mean mental age or age equivalent score for each test was compared. Correlations between the results of each test were calculated.

Group mean mental ages for the psychometric evaluations was 64.8 months, for the TACL 62.2 months, and for the PPVT 69.4 months. There was no significant difference between mental age mean obtained by the psychometric testing and the mean's of either language test as compared by the Wilcoxon Matched-Pairs Signed-Ranks Test. The mental ages of the PPVT were significantly higher than those of the TACL.

Correlations were obtained between test results. The correlation between the TACL and PPVT results was high. Coefficients of .870 between raw score results and .849 between mental age scores of these tests were obtained. The results of the psychometric assessments and the TACL testing

showed a moderate correlation and a substantial relationship. The psychometric mental ages had a coefficient of correlation of .633 with both the raw score and the mental age results of the TACL. The correlation between the psychometric testing and the PPVT was moderately high when the comparisons were made of raw scores or mental ages but low when the comparisons were made between IQ scores. The raw score of the PPVT had a coefficient correlation of .611 and the PPVT scores had a coefficient correlation of .576 when compared to the current mental ages of the psychometric testing. The correlation between the psychometric IQ's and the PPVT IQ's was .211.

The TACL appears to have several advantages over the PPVT as a receptive language test. It gives more detailed information regarding specific linguistic skills than does the PPVT. The average age equivalent score for a group of developmentally delayed children compared well with their average mental age obtained by psychometric testing.

Since the PPVT correlates well with the TACL, it must test a receptive language function as the TACL was designed solely to test understanding of language. Because the IQ scores as obtained by the PPVT had a low correlation to the psychometric IQ scores, the results of the PPVT should probably not be recorded as IQ's when used with developmentally delayed children.

COMPARISON OF TWO RECEPTIVE LANGUAGE TESTS (PPVT AND TACL) USED WITH THE DEVELOPMENTALLY DELAYED

by

PAUL H. LAMB

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
in
SPEECH COMMUNICATION

EMPHASIS: SPEECH PATHOLOGY/AUDIOLOGY

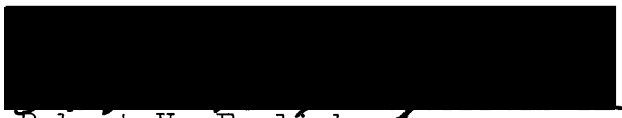
Portland State University

1980

TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the thesis of Paul H. Lamb presented May 13, 1980.

Robert L. Casteel, Chairman



Robert H. English

Mary E. Gordon

Ronald E. Smith

APPROVED:

Frank L. Roberts, Acting Head, Department of Speech Communication

Stanley E. Rauch, Dean of Graduate Studies and Research

ACKNOWLEDGMENTS

The author owes many personal debts of gratitude for assistance so freely and graciously extended. To Dr. Bob Casteel for his sound advice and ever kind and gentle words of encouragement. To Mary Gordon for her quiet, efficient, and effective answering of questions. To Dr. Bob English for his unabated enthusiasm for the completion of this task. To Dr. Jack Hegrenes for his cheerful direction in statistical matters. To all those people who helped me obtain my sample, especially Ernest Cristler, Billie Van Horn, and Barbara Cox. To each and every one I offer my heartfelt thanks.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

CHAPTER I

INTRODUCTION

Children with problems are often tested utilizing standardized instruments to assist in determining the underlying cause, which is the first step in the formulation of a management plan. Often a child with a problem will be evaluated by individuals representing several disciplines comprising a diagnostic team. This occurs in a school setting with a few individuals representing one or more specialties or in a medical setting with many disciplines available for consultation. In either setting, the speech pathologist's role is that of assessing the child's speech, hearing, and language competence and reporting these to the rest of the team in such a manner that the total of all evaluations may be more useful than the sum of the parts.

Developmentally delayed children represent a group who are often evaluated by multiple disciplinary teams. Determining if any given child's subnormal performance is due to a general retardation or a specific deficit is often a crucial issue. Assessment instruments that can help answer this question are used, and evaluation of a child's comprehension of language is necessary in these circumstances.

Two tests that may be used in testing comprehension of language are the Peabody Picture Vocabulary Test (PPVT) and the Test for Auditory Comprehension of Language (TACL). Both tests are nonverbal, multiple choice tests consisting of numbered plates with several line drawings to each plate from which the child is asked to choose one.

Since either test may be used, the clinician must decide whether to utilize both tests or only one of them. If only one, which one? To help in this decision, it would appear wise for the clinician to review what information he really desires or needs to obtain.

One desirable aspect of a test is that it measure what needs to be measured. The PPVT was designed to be an intelligence test, whereas the TACL was developed to be a receptive language test. While the PPVT contains only vocabulary items, primarily singular and plural nouns and some gerunds, the TACL has sections involving morphology and syntax as well as vocabulary. Additionally its vocabulary section covers several parts of speech.

Another desirable aspect of a test is that it can be reported in terms that are meaningful to both the examiner and other participating professionals who will deal with the report. Results of the PPVT are often to be reported as a mental age. It appears that with mentally retarded populations, the PPVT tends to yield higher intelligence

quotients and mental age scores than do the more comprehensive intelligence tests (Sattler, 1974). Results from the TACL can be reported in an age equivalency. The report, however, also can cover specific linguistic skills giving some definition to the linguistic areas needing attention and suggesting a starting point for intervention.

Likewise other factors are important to test selection, such as: familiarity of other professionals with the test, time required to administer and score, and test reliability and validity. The PPVT tends to be recognized by teachers, medical doctors, and other professionals who might be dealing with the child. The TACL, on the other hand, appears to be less well known. The PPVT has well documented reliability, both with normal and developmentally delayed populations. Only a few published studies have attested to either the reliability or the validity of the TACL. While the PPVT has many general studies concerning its validity, investigations involved with the developmentally delayed leaves its validity with this specific population somewhat in question.

STATEMENT OF PURPOSE

The purpose of this investigation was to compare the responses of a group of developmentally delayed children to the TACL and PPVT with the results from a comprehensive intelligence test.

And, to determine the answers to the following two questions:

1. With a sample of developmentally delayed children, does the mean of the mental ages obtained from the psychometric assessments differ significantly from the mean of the mental ages obtained through the TACL results and/or from the mean obtained through the PPVT results?
2. With a sample of developmentally delayed children, do the results obtained from psychometric assessment correlate more closely to those obtained through the TACL than those obtained through the PPVT?

CHAPTER II

REVIEW OF THE LITERATURE

Statements by the authors of PPVT and TACL made in the respective manuals were reviewed as to (1) the design and purpose of the tests, (2) validity, and (3) reliability. Literature concerning test validity and reliability with developmentally delayed children were reviewed.

All published articles reviewed by this investigator concerning the TACL are reported. The literature concerning the PPVT is voluminous. Dunn (1971) does an excellent job of reviewing articles published prior to 1965 in the general information section of the PPVT manual. The articles covered here were published after 1965 and were concerned with the use of the PPVT with developmentally delayed and/or educationally handicapped children. One exception to this is an article by Osicka (1976) who was concerned with two methods of scoring when multiple basals are present. Special emphasis is given to how the PPVT mental ages compared to and correlated with the Stanford-Binet and Wechsler Scales mental age results.

LITERATURE CONCERNING THE PPVT

Purpose for Using the Test

Dunn (1971) has stated the PPVT was designed ". . . to provide an estimate of a subject's verbal intelligence . . . ." He reviewed studies published between 1959 and 1965, and concluded the test was valid for obtaining estimates of verbal intelligence.

Both Berry (1969) and Kleffner (1973) have acknowledged the PPVT was designed to measure intelligence, but disagreed that this should be its primary use. Berry (1969) has stated the PPVT ". . . probably is neither an intelligence nor a language test, as it is called in some speech centers. It is a good measure of a child's comprehension of word meanings." Kleffner (1973) added "although proposed as a test of intelligence, its utility is limited to its measure of receptive vocabulary age."

Validity

Dunn (1971) cited two types of evidence for validity, i.e., "rational" and "statistical." He stated that the PPVT two common types of "rational" validity, "content" and "construct," were available, and that "content validity was built into the test . . . [by a] search in Webster's New Collegiate Dictionary (G. and C. Merriam, 1953) for all the words whose meaning could be depicted by a picture" (Dunn 1971). He supported construct validity by indicating almost

all recognized intelligence tests utilize vocabulary items as part of their construct, and that these items often are the best single indicator of the full scale score in the battery of intelligence subtests.

Osicka (1976) reported a study in which the Peabody, Form A, was administered to 4,414 children of average intelligence. The children had all been born of white mothers at the Kaiser-Oakland Hospital and their mothers were all enrolled in the Kaiser Foundation Health Plan. Of the subjects, 1,885 5-year-olds were tested within one month of their fifth birthday, and 876 9-year-olds, and 816 11-year-olds were tested within two weeks of their birthdays. The author stated ". . . the group is deficient only in the two extremes, the very affluent and the very indigent portions of the total population." He also stated ". . . the children tested are representative of a California school population." One result reported was that the average percentile for the group was 71.5. From this evidence, the author commented "Because children of average intelligence are doing so well on the Peabody, it is time to consider establishing new norms."

Osicka (1976) pointed out "Two methods of scoring the PPVT exist. One deducts credit for all errors above the lowest basal. The second and correct method deducts credit for errors above the highest basal . . . ." He indicated his group and their advisors from the Psychology Department

of the University of California at Berkeley had originally determined in 1969 that all errors should be deducted regardless of their relation to any basal. Osicka (1976) further stated:

> A revised Peabody manual was issued in 1971 (Dunn, 1971) with the only change from the 1965 manual being the addition of five paragraphs (Dunn, 1971, pp. 8-10) which deal specifically with how to score protocols on which the subject had two or more ceilings and/or basals. The unexpected instructions were that the lowest ceiling was to be used with all correct responses above it ignored, and that the highest basal (only basals below the accepted ceiling are to be considered) was to be used with all incorrect responses below it ignored.

Of 4,414 children all of whom were of average intelligence, none had multiple ceilings but 45 percent of them had two or more basals, causing their scores to differ from 1 to 19 points by method of grading used.

Validity with a Developmentally Delayed Population

Sattler (1974) stated the PPVT tends to yield higher intelligence quotients for mentally retarded groups than those obtained from the Stanford-Binet Intelligence Scale form L-M, (SB-LM). Fitzgerald, Pasewark, and Gloeckler (1970) further reported the PPVT overestimated the full scale IQ of the Wechsler Intelligence Scale for Children (WISC) in all but two of the twenty-two studies he reviewed comparing the WISC and PPVT with an educationally handicapped population. Supporting this view, Brown and Rice (1967), Burnett (1965), Covin (1977), Gensemer, Walker, and Cadman (1976),

Pasewark, Fitzgerald, and Gloeckler (1971), Sattler and Anderson (1973), Silberberg and Feldt (1966), Throne, Kaspar, and Schulman (1965), Wells and Pedrini (1967), and Zunich and Tolley (1968) reported the PPVT IQ or mental age to be significantly higher than the similar WISC or Stanford-Binet scores. Additionally, Allen, Haupt, and Jones (1964) reported the PPVT significantly overestimated WISC results for a population of retarded children with severe impairment in visual perceptual development, but was not significantly higher for retarded children whose visual perceptual development was appropriate for their mental ages.

Conversely, numerous studies reviewed showed no significant difference between PPVT IQ's and mental ages and those resulting from the WISC or Stanford-Binet. Such results were reported by Anderson and Flax (1968), Cochran (1970), Cochran and Pedrini (1969), Coyle, Dans, and Cork (1968), Ernhart (1970), Gage and Naumann (1965), Groden, Branson, and Mann (1976), Hammill and Irwin (1965), Kaufman and Ivanoff (1968), Kicklighter (1966), McArthur and Wakefield (1968), Pikulski (1973), Pilley, Harris, Miller, and Rice (1975), Richmond and Long (1977), Ritter, Duffey, and Fischman (1974), and Wells and Pedrini (1971).

Johnson and Johnson (1971), working with a population of 5-year-old headstart children, reported the mean PPVT IQ to be 74.76 while the Stanford-Binet IQ mean was

85.76. These PPVT results were significantly lower than the Stanford-Binet results. Rosenberg and Stroud (1966) tested 28 kindergarten age children from a poverty area. They found the PPVT ". . . overestimated seriously the prevalence of retardation." Their PPVT results gave a significantly lower score than the Stanford-Binet IQ's. Hatch and Covin (1977) compared the PPVT with the WISC on three groups of children. One group was 15 black inner-city kindergarten children, assumed to be comparable to the children in a study by Rosenberg and Stroud (1966), 37 black inner-city children enrolled in a headstart program, and 15 middle socio-economic level children in a University child study center. In all three groups the PPVT resulted in lower scores than the WISC, but the differences were not significant. Koh and Madow (1967) reported "PPVT MA's underestimated Stanford-Binet MA's below the five year level, overestimated above the nine and one-half MA level, and were more comparable between these two MA levels."

The findings of several studies reviewed did not show a significant difference between the PPVT results and those from the Wechsler or Stanford-Binet. Some of the authors of these studies believed the PPVT provided a good estimate of a child's intelligence and other authors believed the PPVT did not give a good estimate of a child's intelligence. Richmond and Lang (1977) in their study involving

39 children who were eligible for special education placement stated "This study suggests that the WISC-R intelligence score is significantly and positively correlated to the PPVT intelligence score." McArthur and Wakefield (1968) used 123 subjects who had a mean SB-LM IQ of 73.44. They reported "The generality of the findings in this study is indicated by the very consistent correlation between PPVT-A (IQ) vs. SB (IQ), PPVT-A (IQ) vs. WISC full scale, and PPVT-A (IQ) vs. WISC Verbal . . . ." Kicklighter (1966) completed a study comparing the PPVT with the SB-LM administered to a sample of 66 children who had been referred by their local schools as being mentally retarded. He stated his study " . . . would indicate that the PPVT is a valid instrument for use in screening children who are referred as mental retardates." Hammill and Irvin (1965) tested the relationship of the PPVT and the Stanford-Binet for 242 mentally subnormal children. Their results indicated "The PPVT is apparently a valid test for measuring the intelligence of both educable and trainable subjects." Shotwell, O'Connor, Gabet, and Dingman (1969) reported "Correlation analysis indicates that the relation between PPVT and the S-B LM is relatively strong." Their study involved 60 institutionalized mentally retarded children. Wells and Pedrini (1971) conducted a study comparing several tests administered to 150 institutionalized adult retardates. They stated their results showed the "PPVT IQ tended to fall

between the WAIS Verbal and Performance scale IQ's but correlated better with the Verbal scale than with the Performance scale."

Kaufman and Ivanoff (1968) completed a study in which 51 young adults, all with prior school diagnoses of mental retardation, were administered the PPVT, WAIS, and Wide Range Achievement Test. They stated:

> In comparing the PPVT with the WAIS and the WRAT, the investigation concluded that although the PPVT may provide an adequate screening instrument with some populations, in working with the mentally retarded, the reading section on the WRAT may more nearly measure functional ability comparable to the WAIS full scale IQ score. However, where the PPVT is used with the mentally retarded, it is suggested that PPVT mental ages be substituted for PPVT IQ scores.

Pilley et al. (1975) have noted, "While the obtained mean Peabody IQ closely approximated the mean Wechsler Full Scale, the Wechsler Verbal, and the Wechsler Performance IQ, were relatively low." The sample for their study consisted of 159 black junior high school age students in a school for students who exhibited academic difficulties. Mean scale scores of the WISC were between 5.7 and 6.6 except for a mean of 4.9 on information and 7.2 on coding.

Mueller (1968) conducted a study using six tests, two of which were the PPVT and the S-B LM. The purpose of the study was to determine the "predictive validity" of these tests. The results of the tests were compared to tested achievement of the subjects two years after the initial

testing. The subjects were 89 educable retarded children whose S-B IQ scores were between 50 and 80. The mean mental age for the PPVT was 5 years 5 months and for the S-B LM 5 years 11 months. He concluded "The data did not indicate adequate validity for the PPVT . . . ."

Weiner (1971) conducted a study of the reliability and validity of the Arthur Adaptation of the Leiter International Performance Scale (AALIPS) and the PPVT when used with preschool, language delayed children. The PPVT and AALIPS were administered to 27 children. Six months later these tests were readministered. Two years after the first testing, the PPVT, AALIPS, and WISC were administered to 21 of the original subjects. Results indicated "There were no statistically significant differences between the mean PPVT IQ and the mean WISC IQ's." Weiner (1971) further states:

> The AALIPS seems to be a highly useful test for determining the adequacy of nonverbal intellectual functioning on preschool, language delayed children . . . . [however] The PPVT . . . seems to have distinct limitations as a measure of adequacy of verbal functioning. While results on the test are acceptably stable, they do not predict later functioning on a broader test of ability [WISC] to solve verbal problems.

Shaw, Matthews, and Klove (1966) conducted a study to determine the relationship between the PPVT and the WISC. The subjects were 83 children who were underachieving in school, suspected of having some kind of organic brain damage, and were referred to a neuropsychology laboratory.

They indicated that because the PPVT had good reliability, it must be measuring "something" fairly well, but because of its poor relationship with the WISC, this "something" was not intelligence. They further stated this relationship ". . . is least in the group of children where reliable and valid intelligence estimates are needed most, i.e., those with below average intelligence." They suggested this "something" might be a specific language deficit. Carr, Brown, and Rice (1967) responded to this suggestion by Shaw et al. and conducted a study comparing the PPVT to the ITPA with a population of 90 educable mentally retarded subjects attending special education classes. Auditory Decoding, Visual Decoding, Auditory-Vocal Association, Visual-Motor Association, Vocal Encoding, Motor Encoding, Auditory-Vocal Automatic, Auditory-Vocal Sequential, and Visual-Motor Sequential were the nine subtests of the ITPA that were compared with the PPVT. They stated "The present study, using the nine subtests of the Illinois Test of Psycholinguistic Abilities as predictor variables, failed to find the PPVT as an adequate measure of any of the nine specific language functions." Williams, Marks, and Bialer (1977) also conducted a study comparing the PPVT with the ITPA. The subjects were 48 mentally retarded children with mental ages from two to six years, and 48 children selected randomly from the

regular kindergarten through third grade of an elementary school. They reported:

> The PPVT and the auditory and visual subtests of the ITPA were administered to 48 normal and 48 retarded subjects. Results suggested that the PPVT is not an adequate measure of hearing vocabulary for mentally retarded subjects.

Reliability

Reliability coefficients, as reported in the PPVT manual, range from .67 for six-year-olds to .84 for seventeen- and eighteen-year-olds. Dunn (1965) pointed out ". . . wide ranges of ages among subjects tend to inflate correlations . . . [and] using only subjects falling within a narrow range of the intellectual continuum, as in the case of retardates, tends to reduce correlation appreciably." Lyman (1965) has observed another problem of accurately scoring the PPVT, especially with younger children. He writes:

> The publisher effects an economy in printing the manual by using 6-month (through 5 years) and 12-month chronological age classifications; however, the use of such gross intervals is reflected in big "jumps" in the IQ table. For example, for a raw score of 50, a child of 5-5 would receive an IQ of 101 while a child of 5-6 would receive an IQ of only 89. "Jumps" of as much as 20 IQ points can be found at the extremes for younger children. The Stanford-Binet, tabling CA's in one-month units, has much smaller "jumps" (rarely above 2 IQ points).

Bashaw and Ayers (1967) confirmed this. They stated that the mental age is the score that should be used and compared for young preschool children.

Blue (1969) did a temporal stability and alternate form reliability study with the trainable mentally retarded using 116 subjects ranging in age from 6-6 to 32-8 years. He retested one year after the first test and obtained reliability coefficients of .93 for temporal stability for mental ages and .92 for alternative form for mental ages. His conclusion was "High reliability was demonstrated in both alternate form testing and one year internal test retest . . . ."

Kahn (1966) conducted a long-term reliability study for the PPVT, Form A, with a sample of 141 young adult, mentally retarded subjects over a four year annual testing program. He obtained 372 paired scores resulting in correlation coefficients of .82 over a one year interval, .85 for a two year interval, and .80 for a three year interval. Kahn's conclusion was that ". . . long term coefficients were . . . approximately . . . those reported for short term reliability." Raskin and Fong (1970) tested the six month temporal stability of Form A with normal and educationally mentally retarded children. Correlations for mental age scores involving all subjects in the normal group was .87 and for all mentally retarded subjects .89. A subgroup of older (8-6 to 10-2 years) normal children (N = 23) resulted in a correlation coefficient of only .68. The authors felt this low correlation was due to a factor

of "learning" in the intervening time, even though the time period had spanned summer vacation. On the other hand, older educationally mentally retarded subjects' average mental age had decreased five months between tests. The authors felt this decline was due to the summer vacation break. Finally, Coyle et al. (1966), testing 45 children being seen for speech intervention, found a correlation of .81 between Form A and Form B of the PPVT indicating good alternate test reliability.

Time

Dunn (1971) has indicated the PPVT is an untimed test but should require only 10 to 15 minutes to administer. Silverstein and Hill (1967) administered the test to 100 institutionalized retarded children and reported an administration time required of 7.9 minutes on the average with a standard deviation of 3.2 minutes.

LITERATURE CONCERNING THE TACL

Purpose for Using the Test

Carrow (1973) stated the TACL was designed to serve two primary functions:

> The first function is to measure the auditory comprehension of language structure and, on the basis of the child's performance, permit assignment of the child to a developmental level of comprehension. The second function is diagnostic. Performance on specific items and groups of items allow the examiner to determine the areas of

linguistic difficulty. The child may have trouble in understanding prepositions or pronouns, etc. This information provides a basis for educational planning and intervention.

Validity

The test has been revised five times. Many of the studies reported in the manual under the section on Validity and Reliability were not published and the reviewer is not certain which study used which revision. The original revision (Carrow 1968) had 123 items. Bartel, Bryan, and Keehn (1973) reported the test they used consisted of 114 items. Carrow (1971, 1972) stated that the 1969 published edition of the test had 114 plates and that this was what she used in these studies. Marquardt and Saxman (1972) referred to the 1969 published edition so their study very likely used the 114 plate edition. The fifth edition (Carrow 1973) has 101 plates.

The manual claims validity in three different ways. First, the items followed previously demonstrable linguistic developmental patterns and an increase in age is paralleled by an increase in score which she claimed her study showed (Carrow 1968). Carrow's (1971) investigation of low economic Mexican-American children demonstrated that these children, although they were somewhat delayed compared to her standardization group, displayed linguistic development patterns similar to her standardization group. Additionally they displayed an increase in score with an increase in

age. Carrow's (1972) study with bilingual Mexican-American children also verified these two patterns.

The second claim made by Carrow (1973) in the manual was that the test successfully distinguishes between individuals who have known disorders and those with no disorders. For this assertion she referred to an unpublished study by herself and Lynch. She further indicated that Weiner's (1972) investigation and Marquardt and Saxman's (1972) study also supported this claim. Weiner (1972) conducted a longitudinal study of the language and language related behavior of dysphasic children in utilizing a battery of 13 tests related to auditory-perceptual, auditory-vocal, oral-motor, visual-perceptual, and visual-motor functioning. The TACL also was included as a measure of grammatic comprehension. Seven dysphasic children and seven normal control subjects were tested and then retested at one and two year intervals. The mean score of the TACL for the dysphasic children was 102.14 and for the control group of children was 111.14 on the second year testing. Weiner reported this difference was significant at the $p < .04$ level. All the children in the dysphasic group tested below the standardization group mean for their ages.

Marquardt and Saxman (1972) examined the relationship between language comprehension and auditory discrimination in 30 kindergarten children with numerous misarticulations

and 30 similarly aged children with proficient articulation. Both groups were administered the TACL and the Wepman Auditory Discrimination Test. Results showed the articulation error group performed significantly lower on both tests. For the TACL the articulation error group had 22.57 average errors with an SD or 6.34 errors and the articulation proficient group had 14.96 errors on the average with an SD of 4.62 errors.

The third claim for Validity is the TACL shows the change that occurs with improvement in the language of disordered children, as Weiner's (1972) study with dysphasic children demonstrated such growth.

Only one study utilizing the TACL was found by the reviewer that had been published since the last TACL revision and, hence, had not been reported by Carrow (1973). This was a study by Burrows and Neyland (1978) with 20 kindergarten children, ages 5 years 6 months to 6 years 5 months. They were administered the Gates-MacGintie Reading Tests (Readiness Skills), the Stanford Early School Achievement Test (level one), the Peabody Picture Vocabulary Test, the Test for Auditory Comprehension of Language, and the Goldman-Fristoe-Woodcock Diagnostic Auditory Discrimination Test. Resulting correlations were as follows (Burrows and Neyland 1978):

	Correlation
Reading Readiness and PPVT	0.67*
Reading Readiness and TACL	0.81*
Reading Readiness and Auditory Discrimination	0.81*
Achievement and PPVT	0.70*
Achievement and TACL	0.66*
Achievement and Auditory Discrimination	0.89*

*$p < .01$

Their conclusion was that the high correlations among each of the tests probably indicated they were all testing factors which reflected a child's linguistic competence.

Validity with a Developmentally Delayed Population

Bartel et al. (1973) conducted a study using the TACL with a classroom of trainable mentally retarded children. They used the 114 plate experimental revision. The researcher stated their purpose was to determine the test's appropriateness for use with low functioning, developmentally delayed children and to determine if linguistic comprehension developed in the same order with their sample as it did with the normal children reported by Carrow (1968). The results showed the correlation between the psychometric IQ and the raw score obtained by the TACL was +.80 ($p < .01$) and between mental age and raw score .70 ($p < .01$). The results also showed that the low functioning, developmentally delayed children appeared to acquire linguistic skills in about the same order that

normal children do but that ". . . even when equated on mental age, retarded children's use of grammatical categories was inferior to that of nonretarded children."

Reliability

Carrow (1973) reported an unpublished test-retest study administered by the Southwest Educational Developmental Laboratory. Both the Spanish and the English 1971 revisions of the TACL were used. Fifty-one students (mean age = 82.43 months) were administered the English version. Twenty-five of these children were Mexican-American. Thirty-two Mexican-American children were administered the Spanish version. A correlation coefficient for test-retest on the English version was .94 and .93 for the Spanish version. All the children were retested within a two week interval.

Weiner (1972) reported "A vital question in the extent to which earlier results of the dysphasics on any of these measures predict later performance. Statistically significant rank correlation (Edwards, 1954) were obtained . . . on the Grammar Comprehension Test [a name for one of the TACL's revisions] ($r=0.79$; $p=0.02$)."

Marquardt and Saxman (1972), as a part of their study reported above, readministered the TACL to 10 subjects, (5 with poor articulation and 5 with good articulation) after the initial testing. They reported "The rank order

correlation between error scores for the first and second test administrations for the 10 subjects was 0.92." Weiner (1972), in the longitudinal study reported above, stated that the second year's testing with the TACL with the dysphasic children resulted in "Exactly the same distribution of scores [as] was obtained during the first year of study." This was three children scoring between the mean for their age and one standard deviation below the mean, two children scoring more than one but less than two standard deviations below the mean and two children scoring more than two standard deviations below the mean for their age.

Time

The TACL is an untimed test but Carrow (1973) indicated it should require about 20 minutes to administer. Bartel et al. (1973) reported the range of time to administer the Experimental Test of Linguistic Comprehension to the group of low performing, developmentally delayed children in her study took from 1-1/2 to 3-1/2 hours per child. This was with the experimental 114 plate instrument rather than the present 101 plate TACL. Prior to the present study this investigator had administered the TACL to three retarded children. It required 20 to 25 minutes to administer each of these tests. The experimenter phoned and asked Ms. Bartel (1977) why it took so long to administer the test

to her group. She stated these children had previously been determined by the school district as not being able to benefit from a classroom situation, but with a change of the state law, they were suddenly in such an environment. Attending behaviors were, for the most part, lacking. She indicated the experience of her study would probably not represent what would be expected in the average educationally mentally retarded or trainable mentally retarded classroom.

Summary of the Review

The literature reviewed would indicate the PPVT might yield a higher mental age score. Bartel et al. (1972) was the only study involving the TACL that addressed this subject. Their results would indicate that the TACL probably yields lower age equivalency for receptive language than mental ages with developmentally delayed children.

Since both the psychometric tests and the PPVT purports to measure intelligence while the TACL was designed to measure receptive language skills, logic would have indicated that the PPVT would correlate quite closely to the psychometric results while the TACL would correlate less well with these results. The literature concerning the correlation of the PPVT with the Stanford-Binet and Wechsler tests was variable and inconclusive.

CHAPTER III

METHODS AND PROCEDURES

SUBJECTS

General

Seventeen developmentally delayed children, 5 girls and 12 boys, comprised the subjects for this study. Their ages ranged from 6 years 3 months to 10 years 11 months, with a mean age of 8 years 3 months. Their IQ's ranged from 52 to 81, with a mean of 63. Their calculated current mental ages at the time data were collected for this study ranged from 3 years 8 months to 7 years 0 months with a mean age of 5 years 5 months. Six children were enrolled in Educationally Mentally Retarded classrooms in Umatilla County, Oregon, and 11 children in similar classes in the Portland metropolitan area.

Criteria for Selection

In order to be included in the investigation, each subject met the following criteria:

(1) possessed a mental age of between 3 and 7 years at the time data were collected for this study;

(2) received a psychometric assessment within the previous 36 months with one of the following: SB-LM or WISC-R;

(3) performed more than one standard deviation but less than four standard deviations below the mean for their chronological age group on the psychometric assessment;

(4) have available to the investigator results of the psychometric evaluation in such a manner that a current mental age might be calculated; and

(5) passed a hearing screening test of 25db or better for the pure tone frequencies of 500, 1000, 2000, and 4000Hz in the better ear.

INSTRUMENTATION

Peabody Picture Vocabulary Test

The PPVT (Dunn 1971) is a nonverbal, multiple-choice test designed to evaluate children between the ages of 2-1/2 and 18 years. It has 150 plates of 4 pictures each. The examiner reads the stimulus word, and the subject responds by pointing to or otherwise indicating the picture that best illustrates the word. The test is untimed, but normally takes 10 to 15 minutes to administer. There are two forms that differ only in that a different stimulus word is used for each plate. The manual gives suggested starting points for each age. From this starting point, a basal of 8 consecutive correct answers is determined by first working forward and then backward, if necessary.

Once the basal is determined, the test proceeds forward through the plates until a "ceiling" is established. This is the plate number of the last item presented in which there are six errors in any eight consecutive presentations. Once a basal and ceiling are established, the test is terminated and a raw score is obtained by taking the plate number of the ceiling and subtracting the numbers of errors between that and the basal. This procedure assures the child does not have to contend with a lot of choices obviously below or above his capacity.

According to Sattler (1974) the PPVT was standardized on a sample of 4,012 white subjects, residing in or near Nashville, Tennessee, with an age range of 2 years 6 months to 18 years. The manual (Dunn 1971) states that "By use of the tables, . . . the raw score can be converted to three types of . . . scores: 1) an age equivalent (mental age); 2) a standard score equivalent (intelligence quotient); and 3) a percentile equivalent (percentile)." Age norms extend from 1 year 9 months, through 18 years and give an index of the level of mental development, for a given subject. For instance, a child with a raw score of 75 on Form A is said to possess a mental age of 10 years since his ability to score on the PPVT is similar to that of the average 10-year-old.

A standard score equivalent is obtained by use of the norm tables from the raw score. This score provides a

comparison of an individual with other individuals of the same age.

Test for Auditory Comprehension of Language

The TACL (Carrow 1973) is a nonverbal, multiple-choice test designed to evaluate the receptive language of children between the ages of 3 years 0 months, and 7 years 0 months. It has 101 plates of 3 pictures each. A typical plate contains a picture representing a given concept to be tested, a picture representing a contrast to that of the test concept and a decoy. The examiner reads the stimulus word, phrase, or sentence and the subject responds by pointing to or otherwise indicating the picture best illustrating the stimulus. The test is untimed, but requires approximately 20 minutes to administer. The TACL covers a wide range of linguistic skills. On vocabulary items the TACL tests adjectives, adverbs, and prepositions, as well as nouns and verbs. Morphological constructs of "er" and "ist" added to free morphs are tested. Additionally, categories involving contrasts of case, gender, number, tense, status, voice, and mood are covered. Syntactic structure of prediction, modification, and complementation also are included. Results are tabulated on the back of the response form. By use of norm tables, the overall score can be reported as an equivalency age, a percentile rank by chronological age, and derived score showing

deviation from the chronological age means. The standard score norm tables are graduated by six month intervals. An analysis section is provided on the back of the test for grouping items by linguistic category. These subtest scores might be used for indicating linguistic areas of difficulty and for measuring change at a later time due to intervention or maturation. The response form also provides the age at which 75 and 90 percent of the children pass each item.

DATA COLLECTION

Physical Setting

Subjects 2, 7, 8, 10, 16, and 17, all residents of Umatilla County, were tested in a portable trailer containing a table and chairs. Each child went with the experimenter from the classroom to the trailer to be tested. Subjects 4, 5, 6, 9, 11, 12, 13, 14, and 15, all residents of the Portland metropolitan area, were tested within the room at their school where they were accustomed to being tested. Subjects 1 and 3 were tested at the kitchen table in their respective homes in Portland. In all cases, the examiner and child were the only occupants in a given room, and all rooms were quiet and well lighted.

Hearing Screening

Hearing screening was conducted by the examiner with a portable audiometer at 25db for the pure tone frequencies

of 500, 1000, 2000, and 4000Hz. Subjects 1, 3, 5, 6, 9, 11, 13, and 15 were tested with a Beltone model 10C. Subjects 4, 14, 16, and 17 were tested with a Beltone model 10D. Subjects 2, 7, and 10 were tested with a Macio model 2B. Subject 8 was tested with Macio model 16. In every testing situation the examiner tested the audiometer in the same room the subjects were to be tested in. When subjects 9, 11, and 12 were tested, the examiner could not hear a 1000Hz tone in his right and better ear at 10db but could hear it at 15db. At the times when all other subjects were tested, the examiner could hear a 1000Hz tone at 10db in his right ear but could not hear a 5db tone. Both earphones were checked with the examiner's right ear. All subjects except 1, 3, and 17 were screened for hearing one to several hours prior to being tested with the PPVT or TACL. Subjects 1, 3, and 17 were tested in the same session in which they were language tested, but the examiner took a 10-minute "break" between the hearing screening and the testing situation. The children remained with the examiner, but were allowed to talk about anything they desired.

Testing

Both the PPVT and TACL were usually administered in one session. The break between the tests varied as to the desire and needs of the child, but in no case was it less

than 5 minutes. Administration of the TACL to one child was interrupted by a break, but this is allowed by the manual instructions. The tests were alternated in order of presentation. The TACL was presented first with 8 subjects, and the PPVT was presented first with 9 subjects.

Each test was administered and scored according to the instruction manual. The examiner read to each child the directions recommended by the author. For the PPVT, the introduction was the one to be used with children below 8 yearsof age. With most children the instruction of "be sure to look carefully at all the pictures" was used at least once. Each response was recorded by number on the response form.

One error in the administration protocol of the PPVT did occur. The examiner either had or could have had the information to determine a calculated current mental age of each child prior to giving the tests. This would have allowed starting the test with the "plate in keeping with the best estimate of their mental ages (Dunn 1965)." The examiner arbitrarily chose the plate with which to start after a short talk with each child. The implications of this variance from the manual of instructions are discussed in Chapters IV and V.

In scoring the PPVT the manual of instructions was followed, including the use of the basal closest to the established ceiling in the event of two or more basals.

The results and implications of this factor also are discussed in later chapters.

Recording Time of Administration

The time each test was started and completed was written on the response form. Time elapsed was obtained by subtraction and then recorded on the front page. The time recorded for the start of the test was when the examiner started to read the instructions to the child. The finish time was the time when the child made the last response. "Time for administering the test" did not include time spent in familiarizing the child with the examiner and test situation, the time required to arrange the material and child before administering a test, nor break time if used.

SCORING

Scoring of the PPVT and TACL

The PPVT and TACL were scored according to their respective manual of instructions. Results were recorded as age equivalents. All tests were scored in one afternoon. Average time required to score was obtained by scoring all of one set of tests at one time and dividing the total time taken by 17, the number of tests scored.

Calculations of "Current" Mental Ages

No child had been tested with the SB or WISC-R in the same month the examiner tested that child with the PPVT and TACL. This necessitated a calculation of an "extrapolated" or "current" mental age.

The procedure for determining current mental age for the WISC-R was as follows. The scaled scores for each subtest were obtained from the test. The current chronological age was calculated. The "Scale Equivalents of Raw Score" table for that chronological age was referred to and the "raw score" that corresponded to each "scaled score" was obtained. Procedures outlined in the WISC-R manual (Wechsler 1974) were then followed to obtain a current mental age as determined from Appendix D of the manual. If a child had all subtests that scored in the age levels in the table, then a "mean test age" was obtained. If the child had subtest scores that were "below" the 6 year 2 month level and a "mean test age" could not be calculated then a "median test age" was calculated. If a "median test age" could not be obtained then the formula MA = IQ x CA ÷ 100 was used. One child in the sample had her mental age obtained by a "mean test age," and one child had his mental age calculated by the "median test age" method, and one child had his mental age calculated by the formula method.

The SB-LM has two sets of standardization tables; one set was published in 1960 with the introduction of the L-M edition. The second set was published in 1972 and entitled Stanford-Binet Intelligence Scale, Form L-M, 1972 norm tables (Thorndike 1973).

The first step of the procedure for extrapolating a current mental age for children who had been tested by the SB-LM was to decide which set of norms to use. If the experimenter had available to him both the mental age and IQ of the test results, it was a simple matter to verify which set of norms the examiner had originally used. In such cases, the same set of norms were used to update a current mental age. If, however, the experimenter only had an IQ score with no way to determine which set of norms had been used to obtain that score, then the experimenter used the 1972 norms for determining a current mental age.

The procedure for obtaining the current mental age was as follows. The child's current chronological age was determined. The proper current age was then located in the vertical ordinate in the norm tables. Within the table the IQ reported for that child was located and then the mental age that corresponded to that current age and IQ was read from the horizontal ordinate column.

DATA ANALYSIS

The two questions asked in the statement of purpose required two different analyses of the data. The first question was concerned with comparing the average mental age obtained from each of the tests and the second question was concerned with how the data from one test correlated with the data from another of the tests.

The Wilcoxon Matched-Pairs Signed-Ranks Test (Siegel 1956) was used to compare the average mental ages obtained from each test and to compare mean IQ obtained by psychometric testing to mean IQ obtained by the PPVT results.

The Pearson Product Moment Coefficient of Correlation was obtained by use of Texas Instrument MBA Calculator. Correlations were made between the mental ages obtained from the psychometric testing and those mental ages obtained using the TACL and PPVT. Correlations were made between the mental ages obtained from the psychometric testing and the raw scores obtained from the TACL, the PPVT, and the morphology plus syntax subsections of the TACL. Coefficient of correlation between the IQ results of the PPVT and psychometric testing were also calculated.

CHAPTER IV

RESULTS AND DISCUSSION

RESULTS

The first specific question to be answered was: With a sample of developmentally delayed children, does the mean of the mental ages obtained from the psychometric assessments differ significantly from the mean of the mental age obtained through the TACL results and/or from the mean obtained through the PPVT results? The mean mental age as obtained by the psychometric testing was 64.8 months, by the TACL was 62.2 months, and by the PPVT, 69.4 months. Table I outlines the range, mean, and median for these mental ages. See Appendix A for the raw data.

TABLE I

COMPARISON, IN MONTHS, OF MENTAL OR EQUIVALENT AGES OF SUBJECTS AS OBTAINED BY VARIOUS TESTING

	Range	Mean	Median
Mental age by			
Psychometric testing	44 to 84	64.8	62
PPVT testing	36 to 105	69.4	73
Age equivalent by			
TACL testing	37 to 82	62.2	62

There was no significant difference between the mental age mean of the TACL and the extrapolated current mental age mean of the psychometric testing. Neither was there any significant difference between the mental age mean of the PPVT and the psychometric testing. Therefore the answer to the first question is no, there was no significant difference between the psychometric mental ages and those of the PPVT and the TACL. The results of the PPVT were significantly higher than the equivalent age mean of the TACL results (see Table II and Appendixes B, C, and D).

TABLE II

COMPARISON OF MENTAL AGE MEANS OF VARIOUS TESTS BY WILCOXON MATCHED-PAIRS SIGNED-RANKS TEST

Test Comparison	N	$\bar{X}$ Comparison	T Score
Psychometric to TACL	17	64.8 to 62.2	T = 55
Psychometric to PPVT	17	64.8 to 69.4	T = 53
PPVT to TACL	16**	69.8 to 61.8	T = 15.5*

*Significant at $p < .01$

**In compliance with the Wilcoxon test, subject 6 was not counted in comparing PPVT and TACL results as the mental age for these two tests for this subject was the same.

The question addressed dealt with results in mental ages or age equivalents. The psychometric results, however, were originally obtained as Intelligence Quotients (IQ). The PPVT results can be recorded in IQ's. The PPVT IQ's ranged from 49 to 107, with a mean of 77.35 and a median of 76. The psychometric IQ's ranged from 52 to 84 with a mean

of 64.1 and a median of 60 (see Appendix E). As analyzed by the Wilcoxon Matched-Pairs Signed-Ranks Test the PPVT IQ's were significantly higher than the psychometric IQ's (see Table III).

TABLE III

COMPARISON OF PPVT IQ'S AND PSYCHOMETRIC IQ'S BY WILCOXON MATCHED-PAIRS SIGNED-RANKS TEST

Tests Compared	N	$\overline{X}$ Comparison	T Score
PPVT IQ's to Psychometric IQ's	16	77.35 to 64.1	T = 15.5*

*Significant at .01 or smaller with two-tailed test.

The results of this study showed the mean mental age obtained by the PPVT to be higher and the mean mental age obtained by the TACL to be lower than the average mental age obtained by the psychometric testing. These differences were not statistically significant but do tend to agree with results expected by the literature review.

The second specific question to be answered was: With a sample of developmentally delayed children, do the results obtained from the psychometric assessment correlate more closely to those obtained through the TACL than those obtained through the PPVT? The results of the TACL and PPVT display a high correlation with each other but only a moderate correlation with psychometric assessment results (refer to Appendix F for raw data). The correlation

coefficient between the mental age results of the TACL and the PPVT are shown in Table IV.

TABLE IV

COEFFICIENTS OF CORRELATIONS USING DATA RECORDED IN MENTAL AGES FOR THE PPVT, TACL, AND PSYCHOMETRIC TESTING

	Mental Age by PPVT	Mental Age by TACL
Current Mental Age by Psychometric Testing	.576	.633
Mental Age by TACL	.849	-

Correlations using raw scores tended to be slightly higher than correlations between mental ages. The raw score of syntactical plus morphological subsections of the TACL was also correlated with the other results and showed high correlation with the PPVT raw score and a substantial relationship with the psychometric results.

TABLE V

COEFFICIENTS OF CORRELATIONS USING PPVT AND TACL RAW SCORES AND PSYCHOMETRIC MENTAL AGES

Raw Score	Raw Score PPVT	Raw Score TACL	Current Mental Age by Psychometric Testing
PPVT	-	.870	.611
TACL	.870	-	.633
Morphology and Syntax of TACL	.844	.970	.584

The results of psychometric testing were originally obtained in the form of a recorded Intelligence Quotient (IQ). An IQ was obtained from the PPVT raw score, also. The IQ from the psychometric testing and the IQ from the PPVT had a coefficient of correlation of only .221, which indicates a definite but small relationship.

Large differences between the PPVT and psychometric IQ's were noted with several children. With a sample of 17 children, 6 children had variances between IQ's of less than 10 points, 4 subjects had variances between 10 and 20 IQ points, and 7 children had IQ variances of 20 or more points. Of these 7 children, the PPVT underestimated 1 child by 21 points and overestimated 6 subjects by 20 to 43 IQ points (see Appendix F).

DISCUSSION

Mean Mental Age Results

The majority of the literature reviewed indicated that the PPVT mental age results tend to be higher than the same results obtained from the WISC or Stanford-Binet. Present data would tend to support such findings. In this investigation the mean mental age of the psychometric testing was 64.8 months and for the PPVT was 69.4 months. These results were not significantly different. The results as expressed in IQ's showed the psychometric test results with a mean IQ of 64.1 and PPVT results with a mean IQ of 77.35.

The PPVT results recorded as IQ's are significantly higher than the psychometric results. This relationship was encountered quite often in the literature.

The mean mental age of psychometric results was 64.8 months while the mean mental age of the TACL results was 62.2 months. There was no significant difference between these means but the results tended in the direction expected. Bartel et al. (1973) stated that the results of determining at what mental age each item of the experimental TACL were acquired by the developmentally delayed children in their investigation showed that these children's use of grammatical categories was inferior to nonretarded children even when equated on mental age. If this is true, a difference between the mean equivalent age obtained by the TACL could be expected to be lower than mean mental age obtained from psychometric testing.

The PPVT results were significantly higher than those of the TACL. Since the results of the TACL and PPVT were significantly different, one or the other or both must vary to some degree from the results of the psychometric tests, even though in this case the differences were not significant.

Correlation

The PPVT and TACL had a correlation coefficient of .849 between their mental age results and .870 between

their raw scores. This represents a high degree of correlation and displays a marked relationship between these instruments. The correlation between the raw scores of the PPVT and the syntactical plus morphological subtests of the TACL was .844. These high correlations indicate the PPVT and TACL test similar skills. Since the TACL was designed solely as a test for auditory comprehension of language, it appears logical that the PPVT must test some form of receptive language.

Since the PPVT had been designed as an "intelligence test," it could be expected to correlate better with the WISC-R or the SB-LM than does the TACL. The correlation of PPVT IQ's to psychometric IQ's in the current investigation was a very low .221. Of the 17 subjects, 7 had PPVT results that varied more than 20 IQ points from the WISC-R or SB-LM results. One child's PPVT IQ score was 43 points higher than his SB-LM score. Ritter et al. (1974) labeled this wide difference of IQ scores "variability." They compared the IQ's obtained by Draw-A-Person test, the PPVT, and SB-LM when given to 31 normal kindergarten children. They stated:

> The PPVT evidenced the greatest test score variability across the three IQ categories. The test was found to overestimate IQ, as determined by the SB with four of the overestimates exceeding twenty points. At the same time the PPVT underestimated IQ by as much as thirty points with seven underestimates exceeding twenty points. No significant differences in overall mean IQ were found between the PPVT and SB, perhaps because of the counterbalancing effects of the gross overestimates and underestimates.

This current study did find a significant difference between the PPVT and psychometric IQ means. Where Ritter et al. (1974) had a counter-balanced effect, this study did not. The results of either study would seriously challenge the validity of using the PPVT IQ as an estimate of intelligence. A possibility of having a 30 or 43 IQ point difference from what the S-B results would give is not satisfactory.

Multiple Basals for the PPVT

Of the 17 children, 7, or 41 percent of the sample, created multiple basals in responding to the PPVT. Osicka (1976) reported that 45 percent of the children in his larger study had multiple basals.

The plate number chosen to start the testing could have had an effect on the number of basals. A lower plate number chosen to start the test gives more opportunity for a multiple basal than if a higher plate number had been chosen. Dunn (1965-1971), in the manual of instructions, gives suggested starting plate numbers by chronological age. He states that for a subject suspected of subnormal learning ability, the best estimate of the child's mental age should be used to determine which plate is to be used as the starting point. The experimenter had available to him data that could have been used to calculate a "current mental age" from previously administered psychometric

tests (S-B LM or WISC-R). This information was not used. Instead the experimenter talked for a short while with the child and then subjectively made his own estimation of each child's abilities. It is this experimenter's opinion that a similar process is often, if not usually, used in most testing stiuations.

The plate number arbitrarily chosen by the examiner varied in 11 instances from the plate number that would have been chosen had the "current mental age" been calculated from the psychometric IQ's prior to testing with the PPVT. Two children were started above the "correct" plate number and 9 children were started below. For the 2 children that were started at a higher than recommended plate number, no apparent problems are observed. One child correctly responded to 16 plates prior to making her first error. The other child completed 6 plates before making an error, did 2 plates going backwards, and then finished the test in an appropriate, normal manner.

Nine children started below their recommended plate number. Three had no errors between the actual starting plate number and the recommended plate number. Two would have had to respond to extra plates going backwards because of an error close to their recommended starting point. Therefore no apparent problems existed for five of the children that started below their recommended starting plate number.

For the other 4 children, had the experimenter calculated a "current mental age" from the psychometric results and used that information to determine the starting plate number, starting at the higher plate number would have eliminated an error that created an "extra" basal (see Appendixes G and H).

Since no errors were used in the scoring that occurred below the basal closest to the ceiling, it is the investigator's opinion that the results of the PPVT were in no way affected by using other than "the best estimate" of the child's mental age to determine the starting plate number.

Usefulness of the TACL

The TACL appears to be a viable alternative to the PPVT as an instrument to test receptive language ability with the developmentally delayed child, providing the child's mental age lies between 3 and 7 years. It apparently can be administered to such a child in about 10 to 15 extra minutes, including scoring time, since 14-1/2 minutes were required to administer and score the PPVT, while 25-1/2 minutes were required for the TACL. Because of its design, additional information is available from the results of the TACL that is not available from the results of the PPVT. This is especially true in indicating language concepts and structures that need

attention for intervention and to indicate starting points for such intervention. This allows the speech-language pathologist that is working on a multi-disciplinary team to make a unique and valuable contribution to the overall evaluation. The TACL can perform a very useful language screening function.

The results of the TACL correlate moderately well with the results of the Stanford-Binet or Wechsler. Its design and content, however, is that of a receptive language test, not an intelligence test. Therefore, if the mental age obtained from the TACL is appreciably higher or lower than that obtained from a comprehensive intelligence test, the assumption can be made that the child's receptive language skills are correspondingly higher or lower than his overall intelligence level. Bartel et al. (1973) stated that the group of TMR's in their study were acquiring language skills at a slower pace than the normal group which was used for the test standardization. For instance, a retarded child with a mental age of 3-1/2 years would possess less linguistic skills than a normal child whose chronological age was 3-1/2. If this is true, it would not be surprising to see a developmentally delayed child with a younger mental age from the TACL than the mental age recorded by the SB-LM or WISC-R for the same child.

PPVT as a Test of Language

Carr et al. (1967) and Williams et al. (1977) indicated that in their estimation the PPVT was not a test of "language" because it failed to correlate well with certain subtests of the ITPA. In this study however, the PPVT's coefficient of correlation with the TACL, a test designed only to assess receptive language, was .849 for mental ages and .870 for raw scores. The raw score of the PPVT had a coefficient of correlation with data from the morphological plus syntactical subsections of the TACL of .844. This would indicate the PPVT does test a language function.

CHAPTER V

SUMMARY AND IMPLICATIONS

SUMMARY

The purpose of this investigation was to compare the results as recorded in mental ages of the PPVT and TACL when used with developmentally delayed children. One aspect was to observe how well the mean mental age from each test would compare with the mean mental age obtained from psychometric testing (WISC-R or SB-LM results). Another aspect was to determine how well the data from the PPVT and the data from the TACL would correlate with the psychometric testing results. The subjects, 17 children whose IQ's as determined by a previously administered SB-LM or WISC-R were below more than one but less than four standard deviations of the mean for their chronological age and whose "current" mental age was over 3 years but not over 7 years, were tested with both the PPVT and TACL. A "current" mental age was extrapolated from the IQ results of the psychometric testing. The results of the PPVT and the TACL were recorded in mental ages. Group Mean scores between tests were compared, and correlations between individual paired results for the different tests were obtained.

Group mean mental ages for the psychometric evaluations was 64.8 months, for the TACL 62.2 months, and for the PPVT 69.4 months. There was no significant difference as tested by the Wilcoxon Matched-Pairs Signed-Ranks Test between the mental age group mean obtained from the psychometric testing and those obtained from either the PPVT or the TACL. There was a significant difference, however, between the group mean of PPVT and the TACL. The PPVT results gave a significantly higher mental age score than did the TACL as determined by the Wilcoxon Matched-Pairs Signed-Ranks Test at $p < .01$.

Correlations were determined between test results. The best correlation was between the raw score of the TACL and the syntactical plus morphological subsection of this test. The correlation coefficient was .970 which demonstrates a very dependable relationship. The correlation between the TACL and PPVT results were high. Coefficients were .870 between raw score results and .849 between mental age score. The correlation between the PPVT raw score and the raw score of the morphology plus syntax subsections of the TACL was .844. The results of the psychometric assessments and the results of the PPVT and TACL when expressed as raw scores on mental ages showed a moderate correlation and a substantial relationship. The psychometric results had a coefficient of correlation of .633 with both the raw score and mental age results of

the TACL, .611 with the raw score of the PPVT, .576 with the mental age results of the PPVT, and .584 with the raw score of the morphology plus syntax subsections of the TACL. The IQ results of the psychometric testing had a low correlation with the IQ results of the PPVT with a coefficient of correlation of .211.

IMPLICATIONS

Research

As mentioned above, the author asked several professionals who use the PPVT how they scored the test when a multiple basal occurred. It is the author's opinion that less than 50 percent of the individuals using the PPVT score it according to the 1971 manual of instruction. If this is true, considerable variations as to reporting results exists. A study to systematically determine what percentage of the users are not aware of the 1971 manual instructions might be useful.

The PPVT was designed as a test of intelligence and the TACL was designed as a test of receptive language. Carr et al. (1967), and Williams et al. (1977) indicated that in their opinion the PPVT did not test a language skill since it did not correlate well with any of the ITPA subtests with which they compared it. This author was surprised to see that the PPVT correlated better with the TACL than it did with the psychometric testing. Conversely, the syntax plus

morphology sections of the TACL correlated slightly better with the psychometric testing results than did the PPVT. Both of these facts, if true in a larger context, should have some impact on how the PPVT is used. Therefore, a similar study involving a sample of normal children would be illuminating.

Since the TACL correlated reasonably well with the psychometric results, an underlying factor must be common to both. This factor is undoubtedly an understanding of language. A study comparing the vocabulary section of the WISC-R with the TACL might be helpful.

Clinical

The extra minutes that a speech-language pathologist might invest in using the TACL instead of the PPVT might be a wise use of time. The additional information gained allows the speech-language pathologist to better describe the language competence of the child and to better develop a program for intervention.

BIBLIOGRAPHY

Allen, R. M., Haupt, T. D., and Jones, R. W. A suggested use and non-use for the Peabody Picture Vocabulary Test with the retarded child. Psychological Reports, 15, 421-422 (1964).

Anderson, D. E., and Flax, M. L. A comparison of the Peabody Picture Vocabulary Test with the Wechsler Intelligence Scale for Children. The Journal of Educational Research, 62, 114-116 (1968).

Bartel N. R. Personal phone conversation, January 24, 1977.

Bartel, N. R., Bryan, D., and Keehn, S. Language comprehension in the moderately retarded child. Exceptional Children, 39, 375-382 (1973).

Bashaw, W. L., and Ayers, J. B. An evaluation of the norms and reliability of the PPVT for preschool subjects. Educational and Psychological Measurement, 27, 1069-1975 (1967).

Berry, M. F. Language Disorders of Children, The Bases and Diagnosis. New York: Appleton-Century Crofts (1969).

Blue, C. M. PPVT temporal stability and alternate form reliability with the trainable mentally retarded. American Journal of Mental Deficiency, 73, 745-748 (1969).

Brown, L. F., and Rice, J. A. The Peabody Picture Vocabulary Test: validity for EMRs. American Journal of Mental Deficiency, 71, 901-903 (1967).

Burnett, A. Comparison of the PPVT, Wechsler-Bellevue, and Stanford-Binet on educable retardates. American Journal of Mental Deficiency, 69, 712-715 (1965).

Burrows, E. H., Neyland, D. Reading skills, auditory comprehension of language, and academic achievement. Journal of Speech and Hearing Disorders, 43, 467-472 (1978).

Carr, D. L., Brown, L. F., and Rice, J. A. The PPVT in the assessment of language deficits. American Journal of Mental Deficiency, 71, 937-940 (1967).

Carrow, E. Test for Auditory Comprehension of Language. 5th ed. Austin, Texas: Learning Concepts (1973).

_____. Auditory comprehension of English by monolingual and bilingual preschool children. Journal of Speech and Hearing Research, 15, 407-412 (1972).

_____. Comprehension of English and Spanish by preschool Mexican-American children. Modern Language Journal, 55, 299-306 (1971).

Carrow, M. A. The development of auditory comprehension of language structure in children. Journal of Speech and Hearing Disorders, 33, 99-111 (1968).

Cochran, M. L. A profile of psychological test scores for retarded adults. American Journal of Mental Deficiency, 74, 582-584 (1970).

_____, and Pedrini, D. T. The concurrent validity of the 1965 WRAT with adult retardates. American Journal of Mental Deficiency, 73, 654-656 (1969).

Condit, J. E., Lewandowski, D. G., and Saccuzzo, D. P. Efficiency of Peabody Picture Vocabulary Test in estimating WISC scores for delinquents. Psychological Reports, 38, 359-362 (1976).

Covin, T. M. II. Relationship of the SIT and PPVT to the WISC-R. Journal of School Psychology, 15, 259-260 (1977).

Coyle, F. A., Dans, C., and Cork, E. Form equivalence of the PPVT on children in speech therapy. Psychological Reports, 23, 1002 (1968).

Dunn, L. M. Expanded Manual for the Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service (1965); rev. ed. (1971).

_____. Manual for the Peabody Picture Vocabulary Test. Minneapolis: American Guidance Service (1959).

Ernhart, C. B. The correlation of Peabody Picture Vocabulary and Wechsler Adult Intelligence Scale scores for adult psychiatric patients. Journal of Clinical Psychology, 26, 470-471 (1970).

Fitzgerald, B. J., Pasewark, R. A., and Gloeckler, T. Use of the Peabody Picture Vocabulary Test with the educationally handicapped. Journal of School Psychology, 8, 296-300 (1970).

Gage, G. E., and Naumann, T. F. Correlation of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children. The Journal of Educational Research, 58, 466-468 (1965).

Gensemer, I. B., Walker, J. C., and Cadman, T. E. Using the Peabody Picture Vocabulary Test with children having difficulty learning. Journal of Learning Disabilities, 9, 179-181 (1976).

Groden, G., Branson, M., and Mann, L. Relationships between performances of young handicapped children on Peabody Picture Vocabulary Test and revised Stanford-Binet scale. Perceptual and Motor Skills, 42, 1227-1232 (1976).

Hammill, D., and Irwin, O. C. Peabody Picture Vocabulary Test as a measure of intelligence for mentally subnormal children. Training School Bulletin, 62, 126-131 (1965).

Hatch, G. L., and Covin, T. M. Comparability of WISC and Peabody IQ's of young children from three heterogeneous groups. Psychological Reports, 40, 1345-1346 (1977).

Johnson, D. L., and Johnson, C. A. Comparison of four intelligence tests used with culturally disadvantaged children. Psychological Reports, 28, 209-210 (1971).

Kahn, H. Evidence for long-term reliability of the PPVT with adolescent and young adult retardates. American Journal of Mental Deficiency, 70, 895-898 (1966).

Kaufman, H. I., and Ivanoff, J. M. Evaluating the mentally retarded with the Peabody Picture Vocabulary Test. American Journal of Mental Deficiency, 73, 396-398 (1968).

Kicklighter, R. H. Correlation of Peabody Picture Vocabulary Test scores and the Stanford-Binet Intelligence Scale, form L-M scores in an educable mentally retarded population. Journal of School Psychology, 5, 75-76 (1966).

Kimbrell, D. L. Comparison of PPVT, FRPVT, RS-B and academic achievement scores among institutionalized educable mental retardates. Perceptual and Motor Skills, 23, 1178 (1966).

Kleffner, F. R. Language Disorders in Children. New York: The Bobbs-Merrill Company, Inc. (1973).

Koh, T. H., and Madow, A. A. Relationship between PPVT and Stanford-Binet performance of institutionalized retardates. American Journal of Mental Deficiency, 72, 108-113 (1967).

Lyman. H. B. In The Sixth Mental Measurements Yearbook, by O. K. Buros. Highland Park, N. J.: The Gryphon Press (1965).

Marquardt, T. P., and Saxman, J. H. Language comprehension and auditory discrimination in articulation deficient kindergarten children. Journal of Speech and Hearing Research, 15, 382-387 (1972).

McArthur, C. R., and Wakefield, H. E. Validation of the PPVT with the Stanford-Binet L-M and the WISC on educable mental retardates. American Journal of Mental Deficiency, 73, 465-467 (1968).

McCarthy, K. A., and Steckler, J. F. Temporal stability and concurrent validity of the PPVT with typical first grade students. Educational and Psychological Measurement, 33, 469-472 (1973).

Milgram, N. A. A note on the PPVT in mental retardates. American Journal of Mental Deficiency, 72, 496-497 (1967).

Mueller, M. W. Validity of six tests of ability with educable mental retardates. Journal of School Psychology, 6, 136-146 (1968).

Osicka, C. J. Peabody Picture Vocabulary Test: comments on administration and two methods of scoring. Psychological Reports, 38, 1135-1144 (1976).

Pasewark, R. A., Fitzgerald, B. J., and Gloeckler, T. Relationship of Peabody Picture Vocabulary Test and Wechsler Intelligence Scale for Children in an educable retarded group: a cautionary note. Psychological Reports, 28, 405-406 (1971).

Pikulski, J. J. The validity of three brief measures of intelligence for disabled readers. Journal of Educational Research, 67, 67-68 (1973).

Pilley J., Harris, C., Miller, J., and Rice, D. Correlations of scores on the Wechsler Intelligence Scale for Children and Peabody Picture Vocabulary Test for adolescents in a special education program. Psychological Reports, 37, 139-144 (1975).

Raskin, L. M., and Fong, L. J. Temporal stability of the PPVT in normal and educable-retarded children. Psychological Reports, 26, 547-549 (1970).

Richmond, B. O., and Long, M. III. WISC-R and PPVT scores for black and white mentally retarded children. Journal of School Psychology, 15, no. 3, 261-263 (1977).

Ritter, D. R., Duffey, J. B., and Fischman, R. Comparisons of the intellectual estimates on the draw-a-person test, Peabody Picture Vocabulary test, and Stanford-Binet (L-M) for kindergarten children. Psychology in Schools, 11, 412-415 (1974).

Rosenberg, L. A., and Stroud, M. Limitations of brief intelligence testing with young children. Psychological Reports, 19, 721-722 (1966).

Sattler, J. M. Assessment of Children's Intelligence. Philadelphia: W. B. Saunders Co. (1974).

_____, and Anderson, N. E. The Peabody Picture Vocabulary Test, Stanford-Binet, and the modified Stanford-Binet with normal and cerebral palsied preschool children. The Journal of Special Education, 7, 119-123 (1973).

Shaw, D. J., Matthews, C. G., and Klove, H. The equivalence of WISC and PPVT IQ's. American Journal of Mental Deficiency, 70, 601-604 (1966).

Shotwell, A. M., O'Connor, G., Gabet, Y., and Dingman, H. F. Relation of the Peabody Picture Vocabulary Test IQ to the Stanford-Binet IQ. American Journal of Mental Deficiency, 74, 39-42 (1969).

Siegel, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill (1976).

Silberberg, N. E., and Feldt, L. S. The Peabody Picture Vocabulary Test as an IQ screening technique for primary grade referral cases. Journal of School Psychology, 5, 21-30 (1966).

Silverstein, A. B., and Hill, T. V. Comparability of three picture vocabulary tests with retarded school children. Training School Bulletin, 64, 58-61 (1967).

Thorndike, R. L. Stanford-Binet Intelligence Scale, Form L-M, 1972 norm tables. Boston: Houghton Mifflin (1973).

Throne, F. M., Kasper, J. C., and Schulman, J. L. The Peabody Picture Vocabulary Test in comparison with other intelligence tests and an achievement test in a group of mentally retarded boys. Educational and Psychological Measurement, 25, 589-595 (1965).

Wechsler, D. Intelligence Scale for Children - Revised. New York: The Psychological Corporation (1974).

Weiner, P. S. Stability and validity of two measures of intelligence used with children whose language development is delayed. Journal of Speech and Hearing Research, 14, 254-261 (1971).

_____. The perceptual level functioning of dysphasic children: a follow-up study. Journal of Speech and Hearing Research, 15, 423-438 (1972).

Wells, D. G., and Pedrini, D. T. Relationships between the S-B L-M, G-H, and PPVT with institutionalized retardates. American Journal of Mental Deficiency, 72, 412-415 (1967).

_____. Relationships among Wechsler adult intelligence scale, Goodenough-Harris, and Peabody Picture Vocabulary tests with institutionalized retarded adults. Perceptual and Motor Skills, 33, 227-232 (1971).

Williams, A. M., Marks, C. J., Bialer, I. Validity of the PPVT as a measure of hearing vocabulary in mentally retarded and normal children. Journal of Speech and Hearing Research, 20, 205-211 (1977).

Zunich, M., and Tolley, J. Performance on the Peabody Picture Vocabulary Test and Stanford-Binet by institutionalized mentally retarded children. Psychological Reports, 22, 1212 (1968).

APPENDIX A

TABLE VI

MENTAL AND/OR EQUIVALENT AGES OBTAINED FROM TESTING

Subject	Mental Ages		
	Psychometric	TACL	PPVT
1	44	37	36
2	55	66	71
3	56	50	47
4	60	48	49
5	60	46	59
6	60	62	62
7	61	71	75
8	62	58	73
9	62	82	87
10	62	76	99
11	63	49	51
12	67	58	75
13	69	62	78
14	76	66	56
15	77	73	78
16	84	74	78
17	84	80	105

APPENDIX B

TABLE VII

COMPARING MENTAL AND/OR EQUIVALENT AGES OBTAINED BY THE TACL TO THOSE OBTAINED BY PSYCHOMETRIC TESTING

Subj.	Mental or Equivalent Ages		Wilcoxon Matched-Pairs Signed-Ranks Test		
	Psychometric Testing	TACL	d	Rank of d	Rank of Less Frequent Sign
1	44	37	- 7	6.5	
2	55	66	+11	12	12
3	56	50	- 6	5	
4	60	48	-12	13	
5	60	46	-14	15	
6	60	62	+ 2	1	1
7	61	71	+10	10	10
8	62	58	- 4	3	
9	62	82	+20	17	17
10	62	76	+14	15	15
11	63	49	-14	15	
12	67	58	- 9	8	
13	69	62	- 7	6.5	
14	76	66	-10	10	
15	77	73	- 4	3	
16	84	74	-10	10	
17	84	80	- 4	3	
N = 17					T = 55

APPENDIX C

TABLE VIII

COMPARING MENTAL AND/OR EQUIVALENT AGES OBTAINED BY THE PPVT TO THOSE OBTAINED BY PSYCHOMETRIC TESTING

Subj.	Mental or Equivalent Ages		Wilcoxon Matched-Pairs Signed-Ranks Test		
	Psychometric Testing	PPVT	d	Rank of d	Rank of Less Frequent Sign
1	44	36	- 8	5.5	5.5
2	55	71	+16	13	
3	56	47	- 9	7.5	7.5
4	60	49	-11	9.6	9.5
5	60	59	- 1	1.5	1.5
6	60	62	+ 2	3	
7	61	75	+14	12	
8	62	73	+11	9.5	
9	62	87	+25	16	
10	62	99	+37	17	
11	63	51	-12	11	11
12	67	75	+ 8	5.5	
13	69	78	+ 9	7.5	
14	76	56	-20	14	14
15	77	78	+ 1	1.5	
16	84	78	- 6	4	4
17	84	105	+21	15	
N = 17					T = 53

APPENDIX D

TABLE IX

COMPARING MENTAL AND/OR EQUIVALENT AGES OBTAINED BY THE TACL TO THOSE OBTAINED BY THE PPVT

Subj.	Mental or Equivalent Ages		Wilcoxon Matched-Pairs Signed-Ranks Test		
	TACL	PPVT	d	Rank of d	Rank of Less Frequent Sign
1	37	36	- 1	- 1.5	1.5
2	66	71	+ 5	8	
3	50	47	- 3	- 4	4
4	48	49	+ 1	1.5	
5	46	59	+13	11	
6*	62	62	0		
7	71	75	+ 4	5.5	
8	58	73	+15	13	
9	82	87	+ 5	8	
10	76	99	+23	15	
11	49	51	+ 2	3	
12	58	75	+17	14	
13	62	78	+14	12	
14	66	56	-10	-10	10
15	73	78	+ 5	8	
16	74	78	+ 4	5.5	
17	80	105	+25	16	
N = 16					T = 15.5

*In compliance to test procedure subject #6 was not counted because d = 0.

APPENDIX E

TABLE X

COMPARISON OF PPVT IQ'S TO PSYCHOMETRIC IQ'S

Subj.	Psycho-metric IQ	PPVT IQ	Wilcoxon Matched-Pairs Signed-Ranks Test		Variance over 20 Points
			Differ-ence	Rank of Less Frequent Sign	
1	55	49	- 6	3.5	
2	60	89	+29		+29
3	60	62	+ 2		
4	60	66	+ 6		
5	58	71	+13		
6	72	91	+19		
7	64	93	+29		+29
8	52	70	+18		
9	52	79	+27		+27
10	59	102	+43		+43
11	73	68	- 5	2	
12	56	76	+20		+20
13	62	78	+16		
14*	73	73	0		
15	69	78	+ 9		
16	84	63	-21	10	-21
17	81	107	+26		+26
N = 16				T = 15.5	

*Subject #14 was not counted as difference equaled 0.

APPENDIX F

TABLE XI

COMPARISON OF PSYCHOMETRIC MENTAL AGE TO RAW SCORES OF PPVT, TACL, AND MORPHOLOGY SYNTAX SECTION OF THE TACL

Subj.	Psychometric Mental Age	Raw Score		
		PPVT	TACL	TACL Morph. + Synt. Only
1	44	29	49	25
2	55	55	77	43
3	56	41	68	36
4	60	43	66	39
5	60	49	64	37
6	60	51	73	41
7	61	57	81	46
8	62	56	69	35
9	62	62	93	54
10	62	67	85	49
11	63	44	67	37
12	67	57	71	40
13	69	58	73	47
14	76	47	77	41
15	77	58	82	46
16	84	58	83	46
17	84	70	89	53

APPENDIX G

TABLE XII

ANALYSIS OF MULTIPLE BASALS ON PPVT

Subj.	No. of Basals	No. of Errors Below "Highest" Basal	Mental Age if Highest Basal Used	Mental Age if All "Errors" Used	Diff. in Months of Mental Age
2	2	2	5-11	5- 7	4
3	2	3	3-11	3- 8	3
8	3	2	6- 1	5- 9	4
9	3	5	7- 3	6- 3	12
11	2	4	4- 3	3-10	5
12	2	1	6- 3	6- 1	2
13	2	2	6- 6	6- 1	5
	Total	19			35

APPENDIX H

TABLE XIII

SUBJECTS WITH MULTIPLE BASALS AND EFFECT OF STARTING AT TOO LOW A PLATE NUMBER

Subj.	Current Mental Age	Recommended Starting Plate	Actual Starting Plate	No. of Basals	Basals Eliminated if Started Correctly	Errors Eliminated if Started Correctly
2	44	15	25	2	0	0
3	55	25	15	2	1	3
8	62	25	15	3	1	1
9	62	25	25	3	0	0
11	63	25	25	2	0	0
12	67	40	25	2	1	1
13	69	40	25	2	1	2