

**A COMPARISON OF THE PPVT AND THE PPVT-R SCORES  
OF SEVENTH AND EIGHTH GRADE CHILDREN**

---

**KATHERINE LAMBERT GORMAN**

A COMPARISON OF THE PPVT AND THE PPVT-R SCORES  
OF SEVENTH AND EIGHTH GRADE CHILDREN

---

An Abstract  
Presented to  
the Graduate Council of  
Austin Peay State University

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Arts

---

by  
Katherine Lambert Gorman

December, 1984



## ABSTRACT

The purpose of this study was to compare scores derived from the original edition of the Peabody Picture Vocabulary Test with scores obtained from the revised edition. The subjects were thirty children in school placement at the seventh and eighth grade levels. The experiment was conducted in two testing phases to counterbalance for practice effect. During phase one the thirty subjects were split into two groups of fifteen each. The first group, labeled Group A, was tested with the original edition of the Peabody Picture Vocabulary Test. The second group, Group B, was tested with the revised edition of the Peabody Picture Vocabulary Test. No more than two weeks later the second phase of testing began. The subjects were administered the form with which they had not been tested.

The results were analyzed using the two-tailed t-test analysis for correlated measures. A significant difference was found between the scores obtained from the original edition and the revised edition of the Peabody Picture Vocabulary Test.

A COMPARISON OF THE PPVT AND THE PPVT-R SCORES  
OF SEVENTH AND EIGHTH GRADE CHILDREN

---

A Thesis  
Presented to  
the Graduate Council of  
Austin Peay State University

---

In Partial Fulfillment of  
the Requirements for the Degree  
Master of Arts

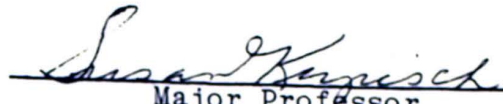
---

by  
Katherine Lambert Gorman


December, 1984

To the Graduate Council:

I am submitting herewith a Thesis written by Katherine Lambert Gorman entitled "A Comparison of the PPVT and the PPVT-R scores of Seventh and Eighth Grade Children." I recommend that it be accepted in partial fulfillment of the requirement for the degree of Master of Arts, with a major in Psychology.

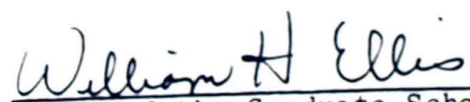
  
Major Professor

We have read this thesis  
and recommend its acceptance:

  
Second Committee Member

  
Third Committee Member

Accepted for the  
Graduate Council:

  
Dean of the Graduate School

## ACKNOWLEDGEMENTS

My sincere appreciation is extended to Dr. Susan Kupisch, Assistant Professor of Psychology, Austin Peay State University, who suggested the research and whose encouragement greatly aided the author. Appreciation is also extended to Dr. Linda Rudolph and Dr. John D. Martin for their assistance.

The author also wishes to express gratitude to the students who volunteered to serve as subjects in this study and to Tommy Bogard.

# TABLE OF CONTENTS

	PAGE
LIST OF TABLES . . . . .	vii
CHAPTER	
1. INTRODUCTION TO THE PROBLEM . . . . .	1
Literature Review . . . . .	2
2. METHOD . . . . .	9
The Subjects . . . . .	9
Description of the Instruments . . . . .	9
Administration and Scoring . . . . .	10
3. RESULTS . . . . .	11
4. DISCUSSION . . . . .	16
5. SUMMARY . . . . .	18
REFERENCES . . . . .	21

# LIST OF TABLES

TABLE	PAGE
1. Comparison of the Mean Age Equivalents Obtained on the PPVT and PPVT-R . . . . .	11
2. Comparison of the Mean Standard Scores Obtained on the PPVT and PPVT-R . . . . .	12
3. Mean PPVT and PPVT-R Chronological Age and Mental Age Scores by Years of Age . . . . .	13
4. Comparison of the Mean Age Equivalents Obtained on the PPVT and the PPVT-R with Mean Chronological Age . . . . .	14



## Chapter 1

### INTRODUCTION TO THE PROBLEM

The Peabody Picture Vocabulary Test was introduced in 1959 as a quick test measuring school age children's receptive vocabulary. Many factors resulted in its widespread use. It was a convenient test, often administered in approximately fifteen minutes. It was relatively simple to administer as the score sheet contained the correct answer, and scoring often could be completed during the testing phase.

Although the Peabody Picture Vocabulary Test had many characteristics which made it a popular test, it became the focus of criticism. In response to the criticism, a revision of the original test was published in 1979.

The revised edition of the Peabody Picture Vocabulary Test, the PPVT-R, was reportedly better standardized and updated. None of the factors which made the original PPVT convenient were altered, thus, the Revised Peabody Picture Vocabulary Test, like its predecessor, gained wide acceptance.

The revision of the Peabody Picture Vocabulary Test appears to be more up to date and reliable. The question arises if these improvements are sufficient enough to warrant the expense of purchase and administration. The problem also arises that the Peabody Picture Vocabulary Test - Revised may produce receptive vocabulary scores significantly different from the original version.

These are questions of interest to administrators who must decide whether to invest in the new copy. It is important to expand the research in this area since school systems are currently using different editions of the PPVT. Some school systems continue to utilize the PPVT and others have transferred to the PPVT-R. If, as the author suspects, these two editions are producing significantly different scores, a student's reading placement may be contingent upon the school system he/she attends. Due to these assumptions, a closer examination of the PPVT and the PPVT-R is warranted.

### Literature Review

The Peabody Picture Vocabulary Test was published in 1959. It is an untimed individual test that usually can be administered in 15 minutes. The test consists of 150 test plates, each corresponding with four numbered pictures. There are two forms of the test, differing in the stimulus word used and corresponding answer sheets. The answer sheets contain the stimulus word, the correct response number, and space for recording the subject's response. On the back of the answer sheet is space for recording observations and basic information about the examinee. The test itself is very simple to administer. The examiner reads a stimulus word and the examinee identifies the plate that corresponds most closely with the word.

The ease with which the Peabody Picture Vocabulary Test (PPVT) can be administered and scored has resulted in its being

one of the most widely used tests for receptive vocabulary. The test also has been the focus of a great deal of criticism.

The standardization group for the PPVT was too small and homogeneous (Jongsma, 1982). The standardization sample consisted of 4,012 children residing in or around Nashville, Tennessee. Critics pointed out that this group was not a representative sample of the population.

Another criticism of the Peabody Picture Vocabulary Test focused on the stimulus pictures utilized. The stimulus pictures were criticized for supporting traditionally biased racial and sex-role stereotypes (Pedriana, 1982). Females represented in the stimulus pictures often are seen performing domestic and subordinate tasks. Only one black person appears in the stimulus pictures and he is depicted as a train station porter. The stimulus pictures also are criticized for being outdated. The mode of dress in the stimulus pictures is thought to be outdated and foreign to today's children.

The reliability of the PPVT has been challenged, adding to further criticisms of the test (Bochner, 1978). The PPVT was found to be most reliable with white middle class children. When dealing with minorities, the PPVT became noticeably less reliable. The reliability between sexes also was questioned. Two studies report sex differences in PPVT reliability at the preschool level (Milgram, 1971; Payne, Hallaham, Ball, & Obenauf, 1972). In both studies, boys' scores were more reliable than those of girls; overall, the



median reliability for boys was 0.60 and for girls was 0.29. The PPVT also appears to be less reliable for older children than for preschool children (Bochner, 1978).

Critics of the PPVT also feel that it exhibits inadequate scaling of the norms (Lyman, 1965), and that the standardization group utilized is now outdated. Children today are believed to have a longer, more advanced vocabulary than standardization subjects used in 1959.

Another major criticism of the PPVT concerns the fact that scores tend to be inconsistent with the scores produced by many other popular tests used with children. The mean PPVT standard score often has been significantly different from the mean Stanford-Binet and Weschler Adult Intelligence Scale Quotients (Naglieri, 1981).

The PPVT has demonstrated a tendency to overestimate I.Q. scores when compared to the Weschler Intelligence Scale for Children-Revised. The PPVT was found to produce scores that were 11 or more I.Q. points higher than those produced by the WISC-R (Mize, 1979). Findings by Vance (1983) support this discrepancy. When comparing the PPVT to the WISC-R, it was found that the PPVT overestimated scores by 6 to 10 I.Q. points.

The revision of the PPVT is an attempt by the publishers to correct some of these problems. The revised edition, designed by Lloyd M. Dunn and Leata M. Dunn, was standardized in 1979 on a representative sample of 4,200 children, 2 years, 6 months to 18 years. The selection of the children participating in



this sample was based on six characteristics: race, age, sex, geographic location, occupation of major wage earner, and community size (Naglieri, 1981).

Twenty-five items were added to the PPVT-R in an attempt to increase the sensitivity of the test. One hundred fifty-six of the original stimulus words out of the original 300 were replaced with new words. All of the illustrations used were replaced by new ones to correct for sex and ethnic balance. Although the original PPVT utilized the same plates with both forms, the PPVT-R incorporated different plates for each form.

The examiner's manual has been expanded to include samples for the scoring of unique cases. Suggested starting places for each age group now are written on the test record. In the original edition, the examiner was required to refer to the test manual.

Although these improvements appear at face value to have produced a better test, studies comparing the PPVT-R with other widely used children's tests have produced varied results. Some studies show the PPVT-R, like its predecessor is inconsistent with scores produced by the WISC-R and the Stanford-Binet.

Seven studies cited by Bracken (1984) found significant differences, with the PPVT-R producing scores 5 to 15 points lower than the Stanford-Binet and WISC-R. The inconsistencies were not found in two studies comparing the PPVT-R with the McCarthy Scales of Children's Abilities (Bracken, 1984).

Breen (1981) found significantly positive relationships between the PPVT-R and WISC-R scores for children who had been referred for emotional disabilities. Worthing (1984) found in his study of 101 special needs students that mean performance on the WISC-R full scale (92.12) and on both forms of the PPVT-R (91.37) did not differ significantly.

Comparison of the PPVT-R with the Peabody Individual Achievement Test (PIAT) found that the PPVT-R was positively and significantly related to the subtests of the PIAT (Naglieri, 1981). The one exception found in this study was with the mathematics subtest on the PIAT. The mathematics subtest did not correlate with the PPVT-R.

In a study conducted by Vance (1983), it was found that the PPVT-R correlated positively and significantly with four subtests of the McCarthy Screening Test. The two remaining subtests, Numerical Memory and Leg Coordination, were positively but not significantly related to the PPVT-R.

These discrepant findings may reflect differences in sample characteristics such as chronological age and/or nature of the subjects used in the studies, but clearly it is indicated that more research is needed. It is also of interest to know how the PPVT-R compares with the original edition.

A basic trend has been established with children tested with both editions of the Peabody Picture Vocabulary Test. Children have a tendency to score lower on the revised edition of the PPVT. In testing 88 children, ages 2 years, 6 months to 5 years, 11 months, Naglieri (1981) found that the PPVT I.Q.

was significantly higher than the PPVT-R. The mental age of 4 years, 7 months is significantly higher than the mean PPVT-R age equivalent of 4 years.

Dunn's research (1981) also supported these findings. He states that a difference of 7 to 8 points can be found on the standard score equivalent with the PPVT showing the higher score. He found as much as a 17 point difference at the upper limits of the tests. Bracken (1984) studied 72 preschool children and found that the PPVT-R consistently produced lower scores than the PPVT.

In an experiment conducted by Choong (1983) the results also were similar. Eighty subjects ages 3 years, 6 months to 4 years, 6 months of age were tested with the PPVT and the PPVT-R. Results showed a difference that was statistically significant at the .001 level. The subjects showed a mean average age of 5 years on the PPVT. Their mean average age score was 4 years, 3 months on the PPVT-R. The PPVT produced scores that were significantly higher than those produced by the PPVT-R.

Pedriana (1982) found significant difference between scores produced by the PPVT and the PPVT-R. The subjects were 31 gifted children enrolled in two midwestern magnet schools. The PPVT produced consistently higher scores than the PPVT-R.

At the present time research indicates that the PPVT and the PPVT-R produce significantly different scores. The bulk of these studies have focused on children below the fifth grade



level. More information is needed for teachers and school psychologists who are dealing with older age groups.

The present study was undertaken to expand the research to include children at the seventh and eighth grade levels. It is proposed that significant differences will be found in test scores derived from the PPVT and PPVT-R. Specifically, it is hypothesized that:

- 1) PPVT scores will be significantly different from PPVT-R scores.
- 2) PPVT-R scores will be significantly lower than PPVT scores.



## Chapter 2

### METHOD

#### The Subjects

The sample consisted of 30 students enrolled in seventh and eighth grade classrooms in Montgomery County, Tennessee, and Houston County, Tennessee. Sixteen male students and 14 female students participated in the study. The subjects ranged in age from 12 years, 6 months to 14 years, 2 months. The mean age was 13 years, 5 months.

#### Description of the Instruments

The Peabody Picture Vocabulary Test is an untimed individual test that usually can be administered in 15 minutes. There are two forms of the test, differing in the stimulus word used. Form A was utilized for this study. The answer sheet contains the stimulus word, the correct response number, and space for recording the subject's response. On the back of the answer sheet is space for recording observations and basic information about the examinee. The test itself is very simple to administer. The examiner reads a stimulus word and the examinee identifies the plate that corresponds most closely with the word.

The Peabody Picture Vocabulary Test-Revised was designed by Lloyd M. Dunn and Leata M. Dunn and was standardized in 1979. One hundred fifty-six of the original 300 stimulus words were replaced with new words. All of the illustrations were

replaced by new ones to correct for sex and ethnic balance. The examiner's manual was expanded to include samples for the scoring of unique cases. The administration manual was redesigned to a more convenient flip style which is free standing.

### Administration and Scoring

Appointments for administration of the PPVT and PPVT-R were made with the first 15 students who volunteered to participate. In Montgomery County five students were tested per day over a three day period. The individual testing took approximately 15 minutes per volunteer. Two weeks later the PPVT-R was administered using the same procedure. In Houston County the second group of 15 students were individually administered the PPVT-R first. The administration period was completed in one day. Two weeks later this group was administered the PPVT in the same manner.

The researcher hand-scored all tests. All information was organized so that the following variables were available for each subject: sex, race, age, PPVT standard score, PPVT-R standard score, age equivalent and percentile rank. The data were analyzed using the two-tailed t-test analysis for correlated measures.

Confidentiality was maintained by assigning a case number to each subject. Only the author had access to the original name list and corresponding number. Both the subject and guardian were required to sign a consent form before participating in the research.

## Chapter 3

### RESULTS

The two-tailed t-test analysis for correlated measures was used to compute the differences between scores attained on the original and revised test. Table 1 summarizes the differences indicated through use of age equivalent values.

Standard score differences are summarized in Table 2. Differences indicated between the mean chronological age and mean age equivalent scores are summarized in Tables 3 and 4.

Table 1  
COMPARISON OF THE MEAN AGE  
EQUIVALENTS OBTAINED ON THE PPVT AND PPVT-R

TEST	MEAN AGE EQUIVALENT IN YEARS AND MONTHS	t	SIGNIFICANCE LEVEL
PPVT-R	14 - 11	3.440	.01
PPVT	15 - 11		

N = 30

Table 2  
COMPARISON OF THE MEAN STANDARD  
SCORES OBTAINED ON THE PPVT AND PPVT-R

TEST	MEAN STANDARD SCORE	t	SIGNIFICANCE LEVEL
PPVT-R	108.12	5.092	.001
PPVT	116.8		

N = 30



Table 3

MEAN PPVT AND PPVT-R CHRONOLOGICAL AGE AND  
MENTAL AGE SCORES BY YEARS OF AGE

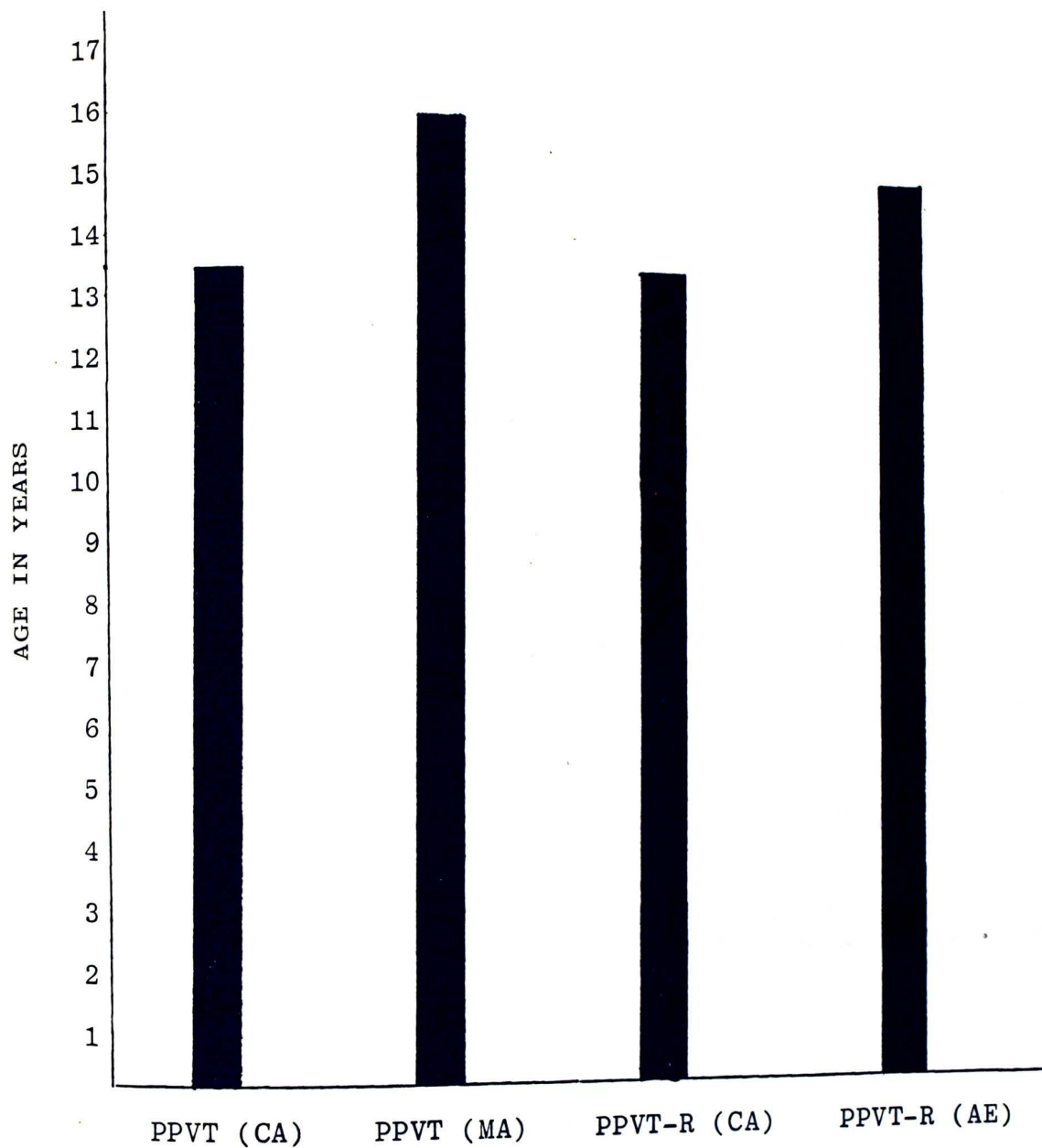


Table 4

COMPARISON OF THE MEAN AGE EQUIVALENTS OBTAINED  
ON THE PPVT AND THE PPVT-R WITH MEAN CHRONOLOGICAL AGE

TEST	MEAN AGE EQUIVALENT	SUBJECTS MEAN CHRONOLOGICAL AGE	DIFFERENCE
PPVT	15 - 11	13.44	2.476
PPVT-R	14 - 11		1.476

The results indicated that a significant difference existed between the mean age equivalent values of the PPVT and the PPVT-R (see Table 1). The results showed a significant difference existed between the mean standard scores of the PPVT and the PPVT-R.

The difference between the mean age equivalents of the PPVT and the PPVT-R is twelve months, with the PPVT having the higher score. As seen in Table 1, the mean age equivalent for the PPVT was 15 years, 11 months, while the mean age equivalent for the PPVT-R was 14 years, 11 months. Results of the two-tailed t-test analysis for correlated measures established that this difference was statistically different beyond the .01 level of significance.

Results of the two-tailed t-test analysis for correlated measures established the mean PPVT standard score value was higher than the mean PPVT-R standard score, beyond the .001 level of statistical significance (see Table 2). The mean PPVT standard score was 116.8 being 8.68 points higher than the mean PPVT-R standard score of 108.12.

These results clearly indicate that the mean PPVT-R age equivalent value is closer to the chronological age of the subjects tested (see Table 4). The mean chronological age of the subjects was 13 years, 5 months; their mean PPVT-R age equivalent was 14 years, 11 months, a difference of 1.476. The mean PPVT age equivalent value of 15 years, 11 months is higher than the subjects' mean chronological age by 2.476.

DISCUSSION

The results of this research are consistent with previous studies conducted comparing PPVT and PPVT-R scores. Results of the two-tailed t-test analysis for correlated measures established a significant difference at the .001 level between standard score values derived from the PPVT and the PPVT-R. This is consistent with results found by Choong (1983).

The PPVT produced higher age equivalent values and standard scores than the PPVT-R. These data were consistent with previous research. Bracken (1984) found that the PPVT consistently produces higher scores than the PPVT-R.

The mean age equivalent of the PPVT-R was closer to the subject's chronological age than related scores on the PPVT. The PPVT produced age equivalents much higher than the subject's chronological age. These data also were consistent with data cited in previous research (Choong, 1983).

Utilization of these data spawns several assumptions. It is suggested that the use of the PPVT may result in an inflated age equivalent. Changes in the standardization may have brought the PPVT-R more into line with the present generation of children being tested.

The significant difference in scores attained by the subjects on the PPVT and PPVT-R indicate that these tests should not be used interchangeably. School systems should implement only one



edition since using both editions will yield scores that are not uniform. The PPVT-R yields age equivalent score values that are closer to that of the examinee's chronological age, suggesting it is more appropriate for use today. These data also suggest that purchase and implementation of the new edition may be warranted. This study may suggest that a need for further research into the PPVT-R exists, especially research utilizing older age groups.

## Chapter 5

### SUMMARY

The purpose of this study was to compare scores of seventh and eighth grade students derived from the original edition of the Peabody Picture Vocabulary Test with scores obtained from the revised edition.

The original Peabody Picture Vocabulary Test was published in 1959. Although widely utilized as an untimed individual test for receptive vocabulary, the PPVT has been the focus of much criticism.

The standardization group was considered to be too small and homogenous, thus not a representative sample. The norms were criticized for inadequate scaling and results of the PPVT were inconsistent with many other tests used for children. This inconsistency caused questions about the PPVT's reliability. Another criticism was that the PPVT is outdated. Critics feel that the PPVT is no longer adequate for today's children.

In an attempt to correct these problems, the PPVT was revised in 1979 by Lloyd M. Dunn and Leata M. Dunn. A more representative sample was chosen for the standardization group and items were added to increase sensitivity of the PPVT-R. All of the illustrations were redone to correct for sex and ethnic balance.

On the surface these changes appear to have created an

improved version of the PPVT. The question arises if these changes are sufficient to warrant the expense of purchasing and utilizing the revised edition.

Several studies have been conducted comparing the PPVT with the PPVT-R. The bulk of these studies have focused on children below the second grade level. Although these studies yield positive information about the PPVT-R, more research was needed utilizing older children as subjects.

This research compared the score derived from thirty seventh and eighth grade children on the PPVT and the PPVT-R. Utilizing a two-tailed t-test analysis of data, significant differences were found between PPVT and PPVT-R age equivalent scores. These differences were of sufficient magnitude to be significant at the .01 level. Significant differences were found between the PPVT and PPVT-R mean standard scores. This correlation was statistically significant at the .001 level. The PPVT-R mean age equivalent value was found to be more compatible with the subject's mean chronological age than the PPVT.

The data have several implications. The PPVT-R yields age equivalent scores closer to the chronological age of the subjects than the PPVT. One implication may be that the PPVT is producing inflated scores. Another implication may be that the PPVT-R may be producing a truer indication of the subject's receptive vocabulary level.

Comparisons of the PPVT and the PPVT-R produce significantly different scores, indicating that both editions should not be

utilized in the same school system. Consistent utilization of only one edition should be practiced or the results may be uneven school placement due to fluctuating scores. One final implication may be that more research is needed on older age groups comparing the PPVT with the PPVT-R scores.



## REFERENCES

- Bochner, Sandra (1978). Reliability of the Peabody Picture Vocabulary Test: A Review of 32 Selected Research Studies Published Between 1965 and 1974, Psychology in the Schools, 15, 320-325.
- Bracken, Bruce A. & Prasse, David P. (1984). Concurrent Validity of the PPVT-R for "At Risk" Preschool Children, Psychology in the Schools, 20, 13-15.
- Bracken, Bruce A. & Prasse, David P. (1984). Peabody Picture Vocabulary Test-Revised: An Appraisal and Review, School Psychology Review, 13, 49-60.
- Breen, Michael J. (1981). Comparison of the Weschler Intelligence Scale for Children-Revised and the Peabody Picture Vocabulary Test-Revised for a referred population, Psychological Reports, (49)3, p. 717-718.
- Choong, Jennie & McMahon, Joan (1983). Comparison of Scores Obtained on the PPVT and the PPVT-R, Journal of Speech and Hearing Disorders, 48, 40-43.
- Dunn, L. M. & Dunn, L. M. (1981). Peabody Picture Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service.
- Jongsma, Eugene A. (1982). Peabody Picture Vocabulary Test-Revised (PPVT-R), Journal of Reading, 25, 360-364.
- Kipps, Debit & Hanson, Dave (1983). Test Review: The Revised PPVT, School Psychology Review, I, 112-113.

Lyman, Howard B. (1965). Mental Measurements Yearbook.

Highland Park, NJ: Gryphon Press, p. 820-821.

Milgram, N. A. (1971). I. Q. Constancy in Disadvantaged Negro Children, Psychological Reports, 29, 319-326.

Mize, John M. & Callaway, Byron (1979). Comparison of Reading Disabled Children's Scores on the WISC-R, Peabody Picture Vocabulary Test, and Slorson Intelligence Test, Psychology in the Schools, 16, 356-359.

Naglieri, Jack A. & Naglieri, Diane A. (1981). Comparison of the PPVT and the PPVT-R for Preschool Children: Implications for the Practitioner, Psychology in the Schools, 18, 434-436.

Naglieri, Jack A. (1981). Concurrent Validity of the Revised Peabody Picture Vocabulary Test, Psychology in the Schools, 18, 286-289.

Payne, J. S.; Hallahan, D. P.; Ball, D. W. & Obenauf, P. A. (1972). Sex Differences in Reliability and Congruent Validity of Stanford-Binet Short Form and Peabody Picture Vocabulary Test, Psychological Reports, 31, 934.

Pedriana, Anthony J. & Bracken, Bruce A. (1982). Performance of Gifted Children on the PPVT and the PPVT-R, Psychology in the Schools, 19, 183-185.

Sattler, Jerome, M. & Altes, Linda M. (1984). Performance of Bilingual and Monolingual Hispanic Children on the Peabody Picture Vocabulary Test-Revised and the McCarthy Perceptual Performance Scale, Psychology in the Schools, 21, 313-316.

Vance, Booney; Kitsen, Donald & Singer, Marc (1983). Comparison of the Peabody Picture Vocabulary Test-Revised and the McCarthy Screening Test, Psychology in the Schools, 20, 21-24.

Worthing, Ralph J.; Phye, Gary D. & Nunn, Gerald D. (1984). Equivalence and Concurrent Validity of PPVT-R Forms L and M for School-Age Children with Special Needs, Psychology in the Schools, 21, 296-299.

Wright, Dan (1983). Effectiveness of the PPVT-R for Screening Gifted Students, Psychology in the Schools, 20, 25-26.