

# Clinical Use of the Peabody Picture Vocabulary Test with Cleft Palate Pre-Schoolers

BETTY JANE McWILLIAMS, Ph.D.  
*Pittsburgh, Pennsylvania*

Thirty children with palatal clefts with a mean age of 4 years 10 months were given Forms A and B of the Peabody Picture Vocabulary Test; the Stanford-Binet Intelligence Scale, Form L-M; and the Illinois Test of Psycholinguistic Abilities. Comparisons among test scores indicated that the two forms of the PPVT agree with each other and with the S-B at the statistical levels reported in the standardization data. However, for individual children, neither form predicts reliably performance on the alternate form or on the S-B or the ITPA. For this reason, clinicians should be cautious in the clinical use of the PPVT.

The Peabody Picture Vocabulary Test (PPVT) has been used extensively as a device for screening vocabulary development, language development, and mental abilities in speech and hearing clinics. The manual for the PPVT (1965) indicates that an IQ of 100 was arbitrarily assigned to the mean raw score for subjects at each age level of the test. Further, the standard deviation was set at 15 IQ points.

The manual discusses the reliability for Forms A and B. Coefficients of equivalence for raw scores ranged from a low of .67 at the six-year-level to a high of .84 at the seventeen- and eighteen-year-levels. The median reliability coefficient was .77. Standard errors of measurement for IQ's were reported as ranging from 6.00 to 8.61 with the median falling at 7.20. The author of the test concluded, on the basis of his own experience and the work of other researchers, that coefficients of equivalence and temporal stability are satisfactory for both average and disabled children.

The author of the PPVT discusses several types of validity in relationship to the test, among them, congruent validity, defined as the extent to which the PPVT scores compare with scores on other vocabulary and intelligence tests. On the 1960 version of the Stanford-Binet, the mental age correlations are reported to have ranged from .82 to .86 with a median of .83.

The author's reporting on the PPVT appears to be straightforward and objective. However, clinicians may have used the test without reference to the data and to the test's limitations. Thus, they may have interpreted

---

Dr. McWilliams is Director, Cleft Palate Center, University of Pittsburgh.  
Supported in part by U. S. Public Health Service Grant DE-01697, National Institute of Dental Research.

individual test results more broadly than was originally intended or than statistical evidence would support.

### **Procedure**

In the clinical management of children with cleft palates, we began to notice that the two forms of the PPVT did not seem to be in close agreement with each other and that it was difficult to predict from either of the forms what a cleft child's performance would be on the Stanford-Binet (S-B) or the Illinois Test of Psycholinguistic Abilities (ITPA).

In order to investigate this problem systematically, we administered both forms of the PPVT to 30 children with palatal clefts ranging in age from 4 years 2 months to 7 years 8 months with a mean of 4 years 10 months. The children were also given the ITPA and the Stanford-Binet Intelligence Scale, Form L-M.

### **Results**

COMPARISON BETWEEN PEABODY FORMS A AND B. On Form A, these cleft children earned a mean IQ of 102.96. On Form B, the mean IQ was 100.3. The standard error of the difference was 9.04. A test for significance of difference between correlated means of 1.69 failed to reach significance. The correlation coefficient of .75 was well within the range reported in the manual. However, a comparison of the IQ's for each child on Forms A and B led to additional valuable information. It will be remembered that the standard deviation for IQ's is reported as 15 points. In this study, the differences between IQ's on the two forms ranged from 1 to 37 IQ points with a mean of 11.4. However, 11 children out of the group of 30 had differences of 15 IQ points or greater between the two test forms. While differences of this magnitude tended to be reflected in the higher score occurring on Form A, this was not invariably the case. In fact, the reverse was true for four children, one of whom achieved an IQ 37 points higher on Form B. These results are in essential agreement with those reported by Nicolosi and Kresheck (1972).

COMPARISON BETWEEN THE PEABODY AND THE STANFORD-BINET, FORM L-M. The Stanford-Binet Intelligence Test, Form L-M, yielded a mean IQ of 109 with a standard deviation of 17.86. The correlation coefficient between the Binet and PPVT A was .74, again within the general range reported in the test manual. It is interesting to note, however, that the differences between S-B and PPVT A IQ's ranged from 2 to 49 IQ points with a mean of 13.4. The S-B yielded a higher IQ in 19 cases. For those subjects, the range of differences was 2 to 49 with a mean of 15.1. For the 11 subjects whose S-B IQ was lower than the PPVT A, the differences ranged from 2 to 33 IQ points with a mean of 10.3.

The S-B mean IQ of 109 compared to a mean PPVT B IQ of 100. The correlation coefficient was .84, again within the general range of the correlations reported in the manual. However, individual differences between

the two tests ranged from 3 to 31 IQ points with a mean of 14. The S-B yielded a higher IQ for 22 of the 30 children. When the S-B was higher, the differences ranged from 1 to 24 IQ points with a mean of 15.9. When the PPVT was higher, differences ranged from 1 to 31 IQ points with a mean of 10.4.

For 17 of the 30 subjects, both Forms A and B of the PPVT were lower than the S-B, while for only 6 subjects were both forms of the PPVT higher than the S-B. For the remaining 7 subjects, one PPVT form was higher and the other lower than the S-B.

Even for those 19 subjects whose PPVT A and B scores differed by no more than 15 IQ points, S-B IQ's were higher than either of the Peabody IQ's for 13 subjects. Only 4 subjects had lower Binet IQ's than were shown by either form of the Peabody.

For the 30 cleft subjects participating in this study, the PPVT, both Forms A and B, tended to underestimate performance on the Stanford-Binet. Thus, using the Peabody for predictive purposes relative to intelligence—at least, intelligence as measured by the S-B—would have been clinically unsound.

COMPARISON BETWEEN THE PEABODY AND THE ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES. When PPVT A was compared with the ITPA by means of the *t* test for significance of difference, the *t* of .78 between PPVT A mean IQ of 105 and mean PLQ (psycholinguistic quotient) of 104 was not significant. However, differences in these scores ranged from 0 to 41 points with a mean of 12.9 points. Nine children, nearly one-third of the population, showed differences of more than 15 points between the two tests. The PPVT A was higher than the ITPA for 5 of these 9 children. However, the ITPA yielded higher scores for the remaining 4 children; these differences ranged from 17 to 41 points, differences of sufficient magnitude to be misleading to the clinician using only a PPVT A.

Comparison between the PPVT B and the ITPA yielded similar results. The *t* of 1.56 between means, PPVT B 102 and ITPA PLQ 104, was not significant. However, differences between individual IQ's and PLQ's ranged from 0 to 30 points with a mean of 12.6. Ten of these differences were greater than 15 IQ points. The mean difference was 25 points for 6 of the 16 subjects whose ITPA scores were greater and 21.5 points for the 4 subjects with higher Peabody scores. COMPARISON BETWEEN STANFORD-BINET AND ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES.

When the S-B mean IQ of 109 and the ITPA mean of 104 were compared, it was found that the mean of the differences between paired scores was 10.7. The S-B yielded a *higher* IQ for 20 of the 29 children on whom both scores were available. Differences ranged from one to 30 IQ points with a mean of 13. For the 9 children whose S-B's were lower than their ITPA's, the differences were of lesser magnitude. They ranged from 2 to 17 IQ points with a mean of 6. Again, the S-B tended to elicit better overall performances from these children with cleft palate.

### Conclusions

These data suggest that the PPVT, while a useful instrument in association with other test procedures, may be a dangerous instrument for screening use with cleft palate children, particularly if the clinician interprets test results to be indicative of overall mental functioning, of linguistic skills, or even of performance on the alternate form of the PPVT itself. This is not intended to suggest that the reliability or the validity of the PPVT has not been well reported. It is meant to suggest only that clinicians may not be justified in making clinical decisions on the basis of the administration of one form of the Peabody. Children may be seriously misevaluated and falsely labeled if this is done. Instead, the PPVT may be a valuable tool when used in conjunction with other instruments as part of the process of differential diagnosis. A similar situation may exist for other children who are handicapped in ways that affect speech and linguistic functioning. It would appear that these children are entitled to a battery of psychological tests in order that they may be fairly evaluated. The S-B yields the most hopeful results for the child followed by the ITPA and finally the Peabody.

These findings are not unreasonable when it is remembered that the S-B with its broader scope permits a child to demonstrate different kinds of abilities including both performance and verbal items. The ITPA, on the other hand, taps verbal abilities, but it does so from a broader base than does the PPVT, which assesses vocabulary as defined by the terms of the test. Children suffering from communications deficits often have developmental irregularities which lead to somewhat erratic and uneven performance on intelligence and language tests. Since the PPVT tests a single ability, it does not permit the clinician to evaluate abilities that may be stronger or weaker than receptive vocabulary. Thus, it may unduly penalize a child whose vocabulary development is poor in comparison with his other abilities or it may overestimate the capacities of a youngster who knows many words but who suffers from other deficits which may influence both his ability to improve his basic linguistic and speech skills and to progress normally in school.

Speech clinicians who are either not knowledgeable about psychological testing or who are reluctant to seek consultation in this area run the risk on the one hand of underestimating the children with whom they work or of overestimating them on the other. Either error can be devastating to the child and may condemn him to therapeutic turmoil.

### References

- DUNN, LLOYD M., *Expanded Manual for the Peabody Picture Vocabulary Test*. Circle Pines, Minnesota: American Guidance Service, Inc. (1965).  
NICOLosi, LUCILLE and KRESHECK, JANET, Variability in Test Scores between Form A and Form B on the Peabody Picture Vocabulary Test, *Language, Speech, and Hearing Services in Schools*. Vol. III, No. 2, April, 1972.