Journal of Psychoeducational Assessment
1984, *2*, 29-36

VALIDITY OF THE EXPRESSIVE ONE-WORD PICTURE VOCABULARY TEST FOR LEARNING-DISABLED CHILDREN

Michael J. Furlong
Jeffrey F. Teuber
University of California—Santa Barbara

This study examined the concurrent and construct validity of the Expressive One-Word Picture Vocabulary Test (EOWPVT) vis-à-vis the Peabody Picture Vocabulary Test—Revised (PPVT-R), the Similarities and Vocabulary subtests of the Wechsler Intelligence Scale for Children—Revised (WISC-R), and the Auditory Reception and Auditory Association subtests of the Illinois Test of Psycholinguistic Abilities (ITPA). The subjects were 97 learning-disabled students (age range = 72-143 months) in the Honolulu area who were administered all tests as part of a multidisciplinary psychoeducational evaluation. The results indicated that the EOWPVT had acceptable concurrent validity, but the construct validity of the test was questionable. The EOWPVT scores were also found to be consistently higher than the PPVT-R scores. Implications for the assessment of learning-disabled children's expressive and receptive semantic skills are discussed.

Learning-disabled (LD) students often exhibit verbal comprehension delays with concomitant attenuation of scores on standard vocabulary tests (Wiig & Semel, 1975). For many years the Peabody Picture Vocabulary Test series (Dunn, 1959; Dunn & Dunn, 1981), which evaluates receptive vocabulary, was the primary diagnostic measure available to school psychologists assessing this important semantic skill of LD children. Although the processes involved in language comprehension are complex (Huttenlocher, 1974), a distinction is often made between receptive and expressive psycholinguistic abilities (Ellis, 1978; Kirk & Kirk, 1971). The limited number of expressive vocabulary tests available to examiners, however, created an assessment gap; Gardner (1979), in response to this need, developed the Expressive One-Word Picture Vocabulary Test (EOWPVT).

To date, few studies have compared performance on the EOWPVT and the Peabody Picture Vocabulary Test-Revised (PPVT-R). Data in the PPVT-R manual (Dunn & Dunn, 1981) report concurrent validity coefficients of .67, .70, and .77 between the PPVT-R and the EOWPVT for three samples of preschool-aged

We thank Kengo Takata, Kenneth Omura, Charlene McGraw, and Joan Geary for their assistance in completing this study.

This research was funded in part by an Academic Senate Grant from the University of California, Santa Barbara.

Address reprint requests to Dr. Michael J. Furlong, University of California, Graduate School of Education, Santa Barbara, CA 93106.

0734-2829/84/020029/$1.00/0

children. Goldstein, Allen, and Fleming (1982) report an EOWPVT-PPVT concurrent validity coefficient of .63 for a sample of 32 preschool children with cognitive deficits (IQs between 58 and 83). In summary, there have been few validity studies of the EOWPVT, and none has determined its concurrent validity with the PPVT-R for LD children in grade school, the population with which these tests are most commonly used. Consequently, although the reliability of the EOWPVT appears to be adequate (range = .91–.96, median = .94) for elementary-age children, its validity with LD children in this age range has yet to be established.

The purpose of this study was to examine the concurrent and construct validity of the EOWPVT for LD children in elementary grades. In completing this analysis, three questions were addressed: (1) What are the concurrent validity coefficients between the EOWPVT and the PPVT-R? (2) Are the scores obtained by LD students on both tests comparable; i.e., are the students' mean scores similar? (3) What is the relationship among the EOWPVT, PPVT-R, and other tests with a vocabulary emphasis; i.e., do receptive and expressive vocabulary tests measure unique response dimensions (construct validity)? This last objective was accomplished by correlating the EOWPVT and PPVT-R with subtests from the Wechsler Intelligence Scale for Children—Revised (WISC-R) (Wechsler, 1974) and the Illinois Test of Psycholinguistic Abilities (ITPA) (Kirk, McCarthy, & Kirk, 1968).

METHODS

Procedure

The EOWPVT, PPVT-R, WISC-R, and ITPA were administered as part of a standard multidisciplinary assessment of students referred for suspected learning disabilities during the 1981–82 school year. The urban school district in which the study was completed had an average daily attendance of 37,000 and an ethnically diverse population. The subtest scores were gleaned from the reports written by school psychologists and speech and language specialists on five multidisciplinary assessment teams. Standard scores and mental ages, where appropriate, were collected for the EOWPVT, the PPVT-R, the WISC-R Vocabulary (WISC-R-Voc), the WISC-R Similarities (WISC-R—Sim), the ITPA Auditory Association (ITPA-AA), and the ITPA Auditory Reception (ITPA-AR) subtests. Eighty-five students were administered Form L and 12 students Form M of the PPVT-R. All of the scores were standardized to have a mean of 100 and a standard deviation of 15 to facilitate comparisons. It was felt that the tests could be reasonably categorized a priori as being primarily expressive (EOWPVT, ITPA-AA, WISC-R-Voc, WISC-R-Sim) or receptive (PPVT-R, ITPA-AR) in response format. The construct validity of the EOWPVT would be enhanced if the correlations within response formats were higher than the correlations between response formats. Since the norms of the four tests overlap in the 72–143 months range, data collection was limited to this age group. In addition, since the ITPA was normed only through 123 months, it was necessary to complete the analyses on two subsamples, age 72–123 months and age 124–143 months.

Subjects

Data were available for 97 LD students, 66 in the younger age group (72–123 months) and 31 in the older age group (124–143 months). The mean age of the

younger children was 100.0 months (SD = 14.9) compared to 133.0 months (SD = 6.6) for the older children. Consistent with district standards, the WISC-R Full Scale IQs (FS) were in the average range for both the younger ($\overline{X}$ = 94.7, SD = 11.9) and the older ($\overline{X}$ = 90.3, SD = 7.4) students. The Performance IQs (PIQ) were equivalent ($t(95) = 0.05$, NS), but the younger students obtained higher Verbal IQs (VIQ) than the older students ($t(95) = 2.96$, $p = .004$). The sample consisted of 70 boys and 27 girls and reflected the multiethnic composition of the school district: Asian-American (20.0%), Spanish Surname (11.1%), Pacific Islander (28.8%), Caucasian (15.6%), and other (24.4%). English was the primary language of all the students.

Students were determined to be LD during multidisciplinary case conferences (a school social worker, regular education teacher, and diagnostic-prescriptive teacher also participated in the evaluation). The district's criteria for LD certification included (a) normal intelligence, (b) an ability-achievement discrepancy in any of the seven areas outlined in PL 94-142 (1.5 years for ages 72–95 months, 2.0 years for ages 96–131 months, and 2.5 years for ages 132–179 months), (c) attentional deficits, (d) perceptual processing deficits (auditory, visual, or visual motor), and (e) expressive, receptive, or integrative language delays. To be certified LD, a child had to meet all five criteria, and the deficits could not be attributable primarily to cultural, language, emotional, cognitive, or sensory factors.

RESULTS

To answer the questions posed previously, the following analyses were completed: (a) Level analyses were performed to compare test means both between and within age groups; (b) correlations among the vocabulary tests were computed for each age group to derive validity coefficients; and (c) the correlation matrix for the younger age group was subjected to a maximum likelihood factor analysis to determine the construct validity of the EOWPVT.

Level Analysis

Since a total of 28 planned t tests comparing means both across and within the two age groups were completed, the experimentwise alpha level ($p = .05$) was partitioned using Dunn's procedure to ensure that the probability of a Type I error was not greater than 5%. Thus an alpha of .0018 was required for a mean difference to be significant.

The results of the level analyses (Table 1) across the two age groups indicated that the older LD students scored lower on all of the tests, but the only comparison that was significantly different was that for the WISC-R-Voc subtest ($t(95) = 3.77$, $p < .0001$).

Within the younger age group, one-sample t tests indicated that the students obtained significantly higher scores on the EOWPVT than on the PPVT-R ($t(65) = 8.10$, $p < .0001$) and the ITPA-AA ($t(65) = 3.95$, $p < .001$). The students' scores on the PPVT-R were found to be significantly lower than on the WISC-R-Sim ($t(65) = -6.52$, $p < .0001$), the WISC-R-Voc ($t(65) = 4.96$, $p < .0001$), the WISC-R-VIQ ($t(65) = -4.77$, $p < .0001$), and the WISC-R-PIQ ($t(65) = -6.91$, $p < .0001$).

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR VOCABULARY MEASURES BY AGE

Vocabulary Measures	Age 72–123 Months (N = 66)		Age 124–143 Months (N = 31)		
	$\overline{X}$	SD	$\overline{X}$	SD	t
PPVT-R	83.0	15.3	75.9	12.4	2.26
EOWPVT	92.6	14.7	91.5	11.6	0.36
ITPA-AA	85.7	18.0	—	—	—
ITPA-AR	86.1	19.6	—	—	—
WISC-R-Sim	97.9	13.2	92.7	10.6	1.90
WISC-R-Voc	92.3	11.7	83.1	10.0	3.77*
PPVT-R-MA	83.9	17.9	97.5	17.1	—
EOWPVT-MA	91.9	19.7	119.4	17.1	—
WISC-R-VIQ	91.8	11.8	84.7	9.1	2.96
WISC-R-PIQ	99.1	14.1	98.9	11.0	0.05

Note. All values are standard scores ($\overline{X}$ = 100, SD = 15) with the exception of the PPVT-MA and EOWPVT-MA values, which are mental age scores expressed in months.

*$p < .0001$, partitioned to maintain experimentwise alpha level of .05. The remaining comparisons are nonsignificant using Dunn's procedure.

For the older group of LD students, the EOWPVT was significantly higher than the PPVT-R ($t(30) = 7.92$, $p < .0001$), the WISC-R-Voc ($t(30) = 4.78$, $p < .0001$), and the WISC-R-VIQ ($t(30) = 3.55$, $p < .001$). In comparison, the PPVT-R was significantly lower than the WISC-R-Voc ($t(30) = -4.81$, $p < .0001$), the WISC-R-Sim ($t(30) = -8.03$, $p < .0001$), the WISC-R-VIQ ($t(30) = -4.63$, $p < .0001$), and the WISC-R-PIQ ($t(30) = -8.38$, $p < .0001$).

In summary, the results of the level analyses indicated that the LD students scored higher on the EOWPVT than on the PPVT-R. This was further demonstrated by obtaining difference scores (EOWPVT minus PPVT-R) for each age group. This analysis indicated that both the younger ($\overline{X} = 9.6$, $SD = 9.6$) and the older ($\overline{X} = 15.6$, $SD = 11.0$) groups of LD children tended to obtain higher scores on the EOWPVT than the PPVT-R. In fact, for 24.3% and 45.2% of the younger and older pupils, respectively, the EOWPVT score was more than one standard deviation (15 points) above the PPVT-R score. Expressed in terms of the mental age equivalent index, the EOWPVT scores were also higher for the younger ($\overline{X} = 8.6$ months, $SD = 11.6$) and older ($\overline{X} = 21.9$ months, $SD = 14.4$) pupils. In brief, the performance of the LD children on the EOWPVT and PPVT-R was not equivalent, the older students revealing the greatest discrepancy between the two vocabulary tests. Additionally, the EOWPVT better approximated the WISC-R Verbal and Performance IQs than the PPVT-R, the latter underestimating these IQ scores for both the younger and older LD students.

Correlational and Factor Analyses

The concurrent validity coefficients between the EOWPVT and the PPVT-R were .80 and .58 for the younger and older LD students, respectively. The correlations among the vocabulary measures are presented in Table 2. Since the ITPA subtests were not available for the older LD students, it was possible to obtain only

TABLE 2
CORRELATIONS AMONG VOCABULARY MEASURES FOR LD STUDENTS

Vocabulary Measures	1	2	3	4	5	6	7	8
1. PPVT-R	—	.80*	.52	.65	.16	.39	.41	.18
2. EOWPVT	.58*	—	.59	.65	.27	.49	.53	.25
3. ITPA-AR	NA	NA	—	.52	.20	.37	.47	.34
4. ITPA-AA	NA	NA	NA	—	.21	.35	.42	.19
5. WISC-R-Sim	.49	.52	NA	NA	—	.48	.75	.20
6. WISC-R-Voc	.74	.60	NA	NA	.69	—	.82	.35
7. WISC-R-VIQ	.55	.49	NA	NA	.78	.82	—	.47
8. WISC-R-PIQ	.15	.12	NA	NA	.00	.21	.10	—

Note. Pearson correlations for students aged 72–123 months above the diagonal and aged 124–143 months below the diagonal. NA = not applicable.
*Concurrent validity coefficients.

intraexpressive (correlations among the EOWPVT, WISC-R-Voc, and WISC-R-Sim) and extraexpressive (correlations between the EOWPVT, WISC-R-Voc, WISC-R-Sim, and the PPVT-R) construct validity coefficients. An examination of these correlations indicates that the mean intraexpressive validity coefficient was .57 compared to a mean extraexpressive validity coefficient of .60. Thus, for the older students, the correlations among the EOWPVT, PPVT-R, and WISC-R subtests were all in the moderate range, with no evidence that the LD students' performance varied according to the response format of the tests.

For the younger LD students, a maximum likelihood factor analysis was completed to determine if the correlations among the EOWPVT, PPVT-R, ITPA subtests, and WISC-R subtests would cluster according to their primary response format; that is, expressive versus receptive. The results of this analysis are presented in Table 3 and indicate that two factors were derived. The first factor was comprised of the EOWPVT, PPVT-R, the two ITPA subtests, and the WISC-R-Voc subtest. This factor accounted for 41.4% of the total response variance and is clearly a general vocabulary factor. The second factor accounted for only 18.2% of the response variance and is defined by high loadings on the two WISC-R subtests and can best be described as a test-specific factor. Thus, for the younger LD stu-

TABLE 3
MAXIMUM LIKELIHOOD FACTOR ANALYSIS WITH VARIMAX ROTATION FOR LD STUDENTS, AGE 72–123 MONTHS

Rotated Loadings*	Factor 1	Factor 2	h^2
PPVT-R	.88	—	.79
EOWPVT	.86	—	.83
ITPA-AR	.59	—	.41
ITPA-AA	.71	—	.53
WISC-R-Sim	—	.67	.46
WISC-R-Voc	.35	.67	.56
Variance Explained	41.4%	18.2%	

*Factor loadings of .35 or higher are listed.

dents, there was also no indication that their performance on these measures of semantic skills was related in a systematic fashion to the primary response format of the tests.

DISCUSSION

The objective of this study was to examine the concurrent and construct validity of the Expressive One-Word Picture Vocabulary Test for LD elementary school children. The correlations among the EOWPVT, PPVT-R, and WISC-R-VIQ were found to be in the moderate-to-high range (r = .41–.80), indicating that the EOWPVT has adequate concurrent validity for LD children. Since the PPVT-R is often the test of choice to assess vocabulary skills, it was also important to examine the construct validity of the EOWPVT to determine if it measures a unique expressive-response domain. These analyses showed that for both the younger and the older pupils there was no evidence that the EOWPVT measures a unique dimension of vocabulary abilities. Particularly for the younger (72–95 months) LD students, the EOWPVT, PPVT-R, ITPA-AR, and ITPA-AA shared a large amount of common response variance that can be labeled general vocabulary ability.

With respect to the level analysis, it was found that one quarter of the younger pupils and nearly one half of the older pupils had a discrepancy between the EOWPVT and the PPVT-R greater than one standard deviation, with the scores on the former being higher. Comparisons of the EOWPVT with the WISC-R-VIQ revealed that these scores were comparable for the younger LD students but were overestimated for the older LD students. Previous research with mildly retarded children (Prasse & Bracken, 1981) and LD children (Breen & Siewert, 1983) have reported that the PPVT-R underestimates the WISC-R-VIQ. These earlier results were replicated, adding confidence to the finding that the EOWPVT approximated the WISC-R-VIQ better than did the PPVT-R.

The combinative effect of these findings suggests that the EOWPVT adequately assesses the vocabulary ability of LD children. The lower concurrent validity coefficients and the overestimated WISC-R-VIQ for the older LD students suggests that the EOWPVT may be used most confidently with LD children in the 6–10-year range. Gardner (1983) has recently developed new norms for children ages 12 years, 0 months through 15 years, 11 months that may make the EOWPVT more useful with older LD children.

These results, however, pose a very interesting interpretive dilemma for evaluators who elect to administer both the EOWPVT and the PPVT-R. One would naturally expect a child to perform better on a receptive vocabulary test than an expressive vocabulary test because the receptive test relies primarily on word recognition, not word recall. The counterintuitive result found in this study may have been due to the restricted normative sample (children from the San Francisco area) used to develop the EOWPVT, a test characteristic that prompted Altepeter (1983) to warn that the norms should be used with "considerable caution." The fact that the EOWPVT better approximated the WISC-R-VIQ than the PPVT-R for this sample of LD children, however, suggests that this parsimonious explanation may not be exhaustive.

A second factor that may have accounted for the EOWPVT/PPVT-R discrepancy can be found in an examination of the item content of these two tests. In the age range of the LD children who participated in this study, 72–143 months, the EOWPVT contains 69 items, Form L of the PPVT-R 50 items, and Form M of the PPVT-R 50 items. Essentially all of the EOWPVT words are simple or classification nouns (97%), whereas less emphasis is given in Form L (80%) and Form M (72%) of the PPVT-R to these word types. Both versions of the PPVT-R include more adjective and verb forms (Form L = 20%, Form M = 28%); LD students may obtain higher scores on the EOWPVT than on the PPVT-R because the former makes less complex demands of their semantic abilities. The fact that the discrepancy between the EOWPVT and the PPVT-R was greatest for the older LD students, the age range at which the PPVT-R includes many verbs and adjectives, lends support to this explanation. This certainly does not imply, however, that the EOWPVT is a more useful test with which to assess the vocabulary abilities of LD children. Since the item content of the PPVT-R presents more complex stimuli than the EOWPVT, it may have greater clinical diagnostic utility and certainly has more prescriptive utility. In any case, additional cross-validation research is needed before a thorough understanding of this differential performance is crystalized.

Gardner (1979) has touted the EOWPVT as a useful test with which to evaluate childrens' verbal cognitive abilities. While this contention has been questioned (Altepeter, 1983), the results of the present study indicate that the EOWPVT/WISC-R-VIQ correlation is comparable to the PPVT-R/WISC-R-VIQ correlation. This finding lends support to the use of the EOWPVT with LD children; however, because evidence substantiating its construct validity was not found, interpretation should proceed cautiously. If a measure of general vocabulary ability is desired, there is little to choose between the EOWPVT and the PPVT-R. However, the fact that the EOWPVT scores were more comparable to the WISC-R-VIQ than the PPVT-R is an attractive feature when a brief measure of verbal intelligence is needed (cf. Sattler, Bohanan, & Moore, 1980). In addition, the tendency for the LD children in this study to obtain higher scores on the EOWPVT than on the PPVT-R suggests that when both tests are administered, mental age scores should be reported cautiously, lest evaluators be put in the position of explaining widely discrepant scores. Evaluators should consider reporting standard scores or percentiles and complete an error analysis to ascertain the types of errors LD children make on each test. Most important, evaluators should not assume that the EOWPVT adequately assesses the expressive semantic skills of LD children.

References

Altepeter, T. A discussion of the Expressive One-Word Picture Vocabulary Test. *School Psychology Review,* 1983, *12,* 106–109.

Breen, M. J., & Siewert, J. C. Comparison of the Peabody Picture Vocabulary Test—Revised and the Wechsler Intelligence Scale for Children—Revised for learning-disabled and referred students. *Journal of Psychoeducational Assessment,* 1983, *1,* 95–99.

Dunn, L. M. *Peabody Picture Vocabulary Test.*

Circle Pines, Minn.: American Guidance Service, 1959.

Dunn, L. M., & Dunn, L. M. *Peabody Picture Vocabulary Test—Revised.* Circle Pines, Minn.: American Guidance Service, 1981.

Ellis, H. C. *Fundamentals of human learning, memory, and cognition* (2nd ed.). Dubuque, Ia.: Wm. C. Brown, 1978.

Gardner, M. F. *Expressive One-Word Picture Vocabulary Test.* Novato, Calif.: Academic Therapy Publications, 1979.

Gardner, M. F. *Expressive One-Word Picture Vocabulary Tests: Upper Extension.* Novato, Calif.: Academic Therapy Publications, 1983.

Goldstein, D. J., Allen, C. M., & Fleming, L. P. Relationship between the Expressive One-Word Picture Vocabulary Test and measures of intelligence, receptive vocabulary, and visual-motor coordination in borderline and mildly retarded children. *Psychology in the Schools,* 1982, *19,* 315-318.

Huttenlocher, J. The origins of language comprehension. In R. L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium.* Potomac, Md.: Lawrence Erlbaum, 1974.

Kirk, S. A., & Kirk, W. D. *Psycholinguistic learning disabilities: Diagnosis and remediation.* Urbana, Ill.: University of Illinois Press, 1971.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. *The Illinois Test of Psycholinguistic Abilities.* Urbana, Ill.: University of Illinois Press, 1968.

Prasse, D. P., & Bracken, B. A. Comparison of the PPVT and WISC-R with urban educable mentally retarded students. *Psychology in the Schools,* 1981, *18,* 174-177.

Sattler, J. M., Bohanan, A. L., & Moore, M. K. Relationship between PPVT and WISC-R in children with reading disabilities. *Psychology in the Schools,* 1980, *17,* 331-334.

Wechsler, D. *Manual for the Wechsler Intelligence Scale for Children—Revised.* New York: Psychological Corp., 1974.

Wiig, E. H., & Semel, E. M. Productive language abilities in learning-disabled adolescents. *Journal of Learning Disabilities,* 1975, *8,* 578-586.