The use of the PPST and intelligence tests in teacher education programs

HOWARD CARVAJAL, JEFFREY KIXMILLER, MEGAN KNAPP,
JOSEPH VITT, and KENNETH A. WEAVER
Emporia State University, Emporia, Kansas

(Stephen F. Davis, *Sponsor*)

Thirty-one upper-division students from a midwestern university volunteered to take the Wechsler Adult Intelligence Scale-Revised, the Stanford-Binet Intelligence Scale: Fourth Edition, and the Peabody Picture Vocabulary Test-Revised. Correlations were calculated between the scores on the three intelligence tests and these students' reading, mathematics, and writing scores from the Pre-Professional Skills Tests. Correlations between PPST Reading and the three intelligence tests ranged from .57 to .68; for PPST Mathematics, .40 to .70; and for PPST Writing, .31 to .63. The moderate correlations between PPST and the intelligence tests suggest that the scales of the PPST do not measure unique skills and abilities. The PPVT-R appears to be the best screening and verification instrument.

The Pre-Professional Skills Tests (PPST) are designed to measure basic proficiency in three areas: reading, mathematics, and writing. The scores from the PPST are used with other information to determine selection and admission into upper-division programs, including teacher education, and as an exit exam in some cases. The 1988-1989 PPST Bulletin of Information lists 12 states among the "several current PPST users" [Educational Testing Service (ETS), 1988a, p. 31]. Thus, they are widely used but also considerably controversial.

The major criticism of the PPST centers on the failure of the authors to provide validity data on the proposed uses of the tests; Oppenheim (1985) cites the lack of correlational studies with other tests of basic skills or with teacher training curricula. Quellmalz (1985) attacks the tests, the fuzzy relationships between the PPST and the National Teacher Examination Program, and the lack of validity data. In summary, he states that "the appropriateness of the PPST for its proposed uses remains to be demonstrated" (p. 1189). ETS (1988a) states that "validity information must be collected and documented by each test user" (p. 12). Ordinarily, test authors are responsible for providing such information (American Educational Research Association et al., 1985). The publication of a disclaimer does not diminish that responsibility. Regardless of the controversy and criticism surrounding the PPST, it is a fact that they are being used nationwide for a variety of purposes.

In an effort to provide additional information about the PPST and what they measure, a study was designed to

determine the relationships between the PPST and three intelligence tests: the Wechsler Adult Intelligence Scale-Revised (WAIS-R), the Stanford-Binet Intelligence Scale: Fourth Edition (Binet IV), and the Peabody Picture Vocabulary Test-Revised (PPVT-R). In addition, the effectiveness of these intelligence-test scores for counseling and screening prospective students who must take the PPST for entry into certain programs, as well as for verifying scores of students who have taken the PPST but failed to make an appropriate score, was assessed.

The intelligence tests listed above were selected for two reasons. First, the WAIS-R ranked first in mention of the use of tests of all types in a 1982 survey of test usage by a variety of psychologists. The Stanford-Binet (1960 and 1972 versions) ranked fifteenth. Information about the Binet IV may be meaningful because of its newness. The PPVT-R was tied for seventh in frequency of use and is the most popular brief screening test of intelligence (Lubin, Larsen, & Matarazzo, 1984).

Second, the PPST are a type of competency tests; Anastasi (1988) stated "that intelligence tests correlate about as highly with achievement tests as different intelligence tests correlate with each other" (p. 414). The relationships between the PPST and the three intelligence tests may help to determine if the PPST measure a unique set of abilities and skills or if they measure constructs similar to those measured by intelligence tests.

METHOD

Subjects

Thirty-one volunteers (12 men, 19 women) were selected from two psychological testing classes at a midwestern university; 16 of the students were enrolled in the teacher education program. The mean age of the participants was 23 years, 2 months ($SD = 4$ years, 0 months); their ages ranged from 20 years, 2 months to 37 years, 1 month. Five students did not take the Binet IV, because their ages exceeded the norms

Correspondence should be addressed to Howard Carvajal, Division of Psychology, Emporia State University, 1200 Commercial, Emporia, KS 66801-5087.

reported for the test; 1 student did not take the WAIS-R because of illness. All took the PPVT-R and the PPST. Additional credit toward the class grade was given for participation.

Tests

The PPST comprise three tests. Reading purports to measure the "ability to understand, analyze, and evaluate written messages" (ETS, 1988a, p. 14). Mathematics assesses "mathematical competencies, including skills acquired from having studied mathematics from elementary through secondary school" (ETS, 1988a, p. 16). The Writing test is divided into two parts: a multiple choice section that requires usage and sentence correction and an essay that samples the student's writing skills. Scores are reported for each test; there are no partial scores or composite scores. The means and standard deviations of the scaled scores range from 176 to 178 and from 4.5 to 7.2, respectively (ETS, 1988b). The participants' scores were reported to the university after they took the tests at the local testing center.

The Full Scale Intelligence Quotient (FS IQ) of the WAIS-R has a mean of 100 and a standard deviation of 15. The General Purpose Abbreviated Battery (GPAB) of Binet IV was chosen in an effort to reduce testing time. Carvajal and Gerber (1987) reported that the average time required for the administration of the total battery was slightly less than 2 h, whereas the GPAB took approximately 52 min. The correlation between scores on the total battery and the GPAB was .94; therefore, little accuracy was lost by using the GPAB. The mean and standard deviation of the Binet IV composite standard age scores (SAS) are 100 and 16, respectively. Form L of the PPVT-R has a mean standard score equivalent (SSE) of 100 and a standard deviation of 15.

Procedure

Testing was conducted between October 22, 1987 and May 4, 1988. The order of administration of Binet IV and WAIS-R was counter-balanced for those participants who took both tests. The PPVT-R was the last test given in every case. A different examiner gave each test. The three examiners were second-year graduate students who had successfully completed all of the required testing courses in either the clinical or school psychology curriculum. Tests were administered during the day in rooms specifically designed and equipped for testing.

RESULTS AND DISCUSSION

Each subject's six raw scores were converted to *z* scores and analyzed with a 2 (sex: male or female) $\times$ 2 (major: teacher education or other) $\times$ 6 (tests: WAIS-R, Binet IV, PPVT-R, PPST Reading, PPST Mathematics, and PPST Writing) mixed-factor analysis of variance. The tests main effect [$F(5,105) = 3.68$] and the sex $\times$ tests interaction [$F(5,105) = 3.28$] were significant at the $p < .01$ level. The three PPST means were individually compared with the WAIS-R, Binet IV, and PPVT-R means with the use of the Bonferroni procedure. The WAIS-R and Binet IV means (.43 and .32, respectively) were significantly greater than the PPST Reading mean ($-.10$). Fisher's least significant difference test of multiple comparisons was performed on the means in the interaction. Females scored significantly greater on the WAIS-R ($M = .55$), PPST Writing ($M = .47$), and Binet IV ($M = .44$) than on the PPST Reading ($M = .04$). Males scored significantly greater on the PPST Mathematics ($M = .26$), WAIS-R ($M = .25$), and PPVT-R ($M = .24$) than on the PPST Reading ($M = -.27$) and PPST Writing ($M = -.56$). The PPST Writing mean for males was lower than all six female means, and the PPST Reading mean for males was lower than all female means except PPST Reading.

Table 1
Descriptive Statistics

	<i>M</i>	<i>SD</i>
PPST		
Reading	177.6	6.0
Mathematics	179.2	5.5
Writing	176.2	4.0
WAIS-R Full Scale IQ	106.4	12.4
Binet IV Composite SAS	105.5	9.9
PPVT-R SSE	103.5	10.6

The descriptive statistics are summarized in Table 1. The PPST means were within ± 1.4 points of the published norms; the means of the intelligence tests were from 3.5 to 6.4 points above the norms but well within one standard deviation. All of the standard deviations were smaller than those of the published norms, with the exception of PPST Reading, which was .1 point larger.

The Pearson correlation was used to estimate the relationships of the WAIS-R FS IQ, the Binet IV composite SAS, and the PPVT-R SSE with the scaled scores of the three PPST tests. With the exception of the WAIS-R and PPST Writing, all of the correlations were statistically significant ($p < .05$). The correlations are summarized in Table 2.

The highest correlation was between Binet IV and PPST Mathematics. Of the six subtests of the GPAB of Binet IV, only Quantitative deals with arithmetic concepts and has a format similar to PPST Mathematics. Both tests are composed of written word problems accompanied by figures and allow the use of pencil and paper. However, this high correlation was not surprising to the authors. In a previous study of college freshmen, the Binet IV composite SAS had the highest correlation of the same three intelligence tests with the American College Test (ACT) Mathematics score ($r = .76$; Carvajal, McKnab, Gerber, Hewes, & Smith, 1989). In addition, intelligence tests correlated with achievement tests as highly as with each other (see Anastasi, 1988).

Regardless of the controversy and criticism of the PPST, they are used in many ways, for example, to determine eligibility for entry into an academic program, such as the professional courses in a teacher-training program, or as an exiting standard for program completion or university graduation. The results (see Table 2) indicate that both the Binet IV and PPVT-R scores can be used a priori to counsel and/or screen future PPST-takers,

Table 2
Pearson Correlations

	WAIS-R	Binet IV	PPVT-R
PPST			
Reading	.57†	.57†	.68‡
Mathematics	.40*	.70‡	.49†
Writing	.31	.63‡	.62‡
WAIS-R		.74‡	.47*
Binet IV			.59†

* $p < .05$. † $p < .01$. ‡ $p < .001$.

as well as a posteriori to verify PPST scores, especially below-standard ones, in which cases counseling may be indicated. Given the increased expense and longer administration of the Binet IV, compared with the PPVT-R, the PPVT-R is recommended for the aforementioned purposes. In addition, the PPVT-R is easy to administer and score and does not require national testing dates, as does the PPST.

Because of the moderate relationships between the PPST and the three intelligence tests employed in this study, it is apparent that the PPST are not unique in the measurement of skills and abilities. As mentioned earlier, 8 of the 9 correlations were statistically significant, and 7 of the 9 were from .49 to .70. Thus, the strongest relationship found in this study accounted for 49% of the total variance. Until such time as validity data for the PPST from a wide variety of studies become available, one should exercise caution in making and/or accepting claims for the PPST and what they measure.

REFERENCES

- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION, & NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- ANASTASI, A. (1988). *Psychological testing*. (6th ed.). New York: Macmillan.
- CARVAJAL, H., & GERBER, J. (1987). 1986 Stanford-Binet Abbreviated Forms. *Psychological Reports*, *61*, 285-286.
- CARVAJAL, H., MCKNAB, P., GERBER, J., HEWES, P., & SMITH, P. (1989). Counseling college-bound students: Can ACT scores be predicted? *The School Counselor*, *36*, 186-191.
- EDUCATIONAL TESTING SERVICE. (1988a). *PPST 1988-1989 bulletin of information*. Princeton, NJ: Author.
- EDUCATIONAL TESTING SERVICE. (1988b). *Pre-Professional Skills Tests score interpretation guide*. Princeton, NJ: Author.
- LUBIN, B., LARSEN, R., & MATARAZZO, J. (1984). Patterns of psychological test usage in the United States: 1935-1982. *American Psychologist*, *39*, 451-454.
- OPPENHEIM, D. (1985). Review of the PPST. In J. Mitchell (Ed.), *The ninth mental measurements yearbook* (pp. 1187-1188). Lincoln, NE: Buros Institute of Mental Measurements.
- QUELLMALZ, E. (1985). Review of the PPST. In J. Mitchell (Ed.), *The ninth mental measurements yearbook* (pp. 1188-1189). Lincoln, NE: Buros Institute of Mental Measurements.