DOCUMENT RESUME

ED 306 271 TM 013 104

AUTHOR Kluever, Raymond C.; Green, Kathy E.
TITLE Identification of Gifted Children: A Comparison of the Stanford Binet 4th Edition and Form LM.
PUB DATE Mar 89
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Abstract Reasoning; *Academically Gifted; *Children; Comparative Analysis; Elementary Education; Elementary School Students; *Intelligence Tests; *Scores; *Visual Measures
IDENTIFIERS Stanford Binet Intelligence Scale; *Stanford Binet Intelligence Scale Fourth Edition

ABSTRACT

This study compared the performance of 51 gifted Caucasian children on the Stanford Binet LM and the Stanford Binet Fourth Edition (SB4) to determine whether significantly different scores would be obtained. The 33 male and 18 female subjects were from middle and upper-middle class families in a western urban area. Their developmental histories, derived from parents' reports, indicated precocious early development of cognitive linguistic abilities. When composite scores were compared, the mean difference between the two tests was 10 points. The LM intelligence quotient was significantly higher than that for the SB4, and the Kaufman Achievement Test total score showed insignificant differences. Correlations between the LM total and the SB4 area scores were significant. The results suggest that the SB4 can be used as part of a system for identifying gifted children, but the composite score cut-off value may have to be adjusted. Further study of the visual abstract reasoning tests is recommended since this sample of gifted children seemed to function quite differently on them than on the verbal and quantitative reasoning tests. Four data tables are included. (TJH)

Identification of Gifted Children
A Comparison of the Stanford Binet 4th Edition and Form LM

Raymond C. Kluever, Ph.D.
Associate Professor

Kathy E. Green, Ph.D.
Associate Professor

University of Denver

Paper presented at the American Educational Research Association annual meeting, San Francisco, March 1989.

ABSTRACT

This study compared the performance of 51 gifted children on the Stanford Binet LM and the Stanford Binet 4th Edition. When composite scores were compared, the mean difference between the two tests was 10 points ($t=5.46$, $p<.001$). The LM IQ was significantly higher than the SB4 and the Kaufman Achievement total score. But, the SB4 and the Kaufman Achievement total score showed insignificant differences. Correlations between the LM total and 4th Edition area scores were significant. These results suggest that the SB4 can be used as a part of a system for identifying gifted children but the composite score cut-off value may have to be adjusted. Further study of the visual abstract reasoning tests is recommended since this sample of gifted children seemed to function quite differently on them than on the verbal and quantitative reasoning tests.

Identification of Gifted Children

In many educational programs, the IQ score serves as a basis for identifying gifted children, generally in conjunction with other scores or assessments. Several variables are used to identify gifted children but among all of them, the IQ is probably the major factor used to establish program eligibility. A cut-off score on a standardized IQ test may be used to initially screen students for further consideration. The Stanford Binet Intelligence Scale has been used as one of the dominant screening measures for young children. With the advent of the Stanford Binet 4th Edition (SB4: Thorndike, Hagen, & Sattler, 1986), it is likely that the Stanford Binet LM (LM: Terman & Merrill, 1973) will be phased out. Little research is available regarding the equivalence of the two tests. In fact the theoretical bases of tests differ and existing data suggest total scores on the two tests to differ. Substantial differences in scores across the two tests would necessitate a re-analysis of criteria for identifying gifted children. The relationship between the LM and SB4 area scores and relationships with achievement measures were also investigated.

Tests designed for use with a general population may not provide measures that are as accurate at the extremes of the distribution--either for high or low scoring individuals. This problem is one possible explanation for score differences between tests: the score obtained from each test is less accurate for an extreme group than for an average group and the difference scores when measures are compared are even less reliable. Disparity in scores across different IQ tests have been found by several researchers working with mentally retarded persons (Cochran & Pedrini, 1969; Kroske, Fretwell, & Cupp, 1965; Spitz, 1983). Spitz (1986) found that scores between the WAIS-R and the WISC-R differed by .76 of the test's standard deviation, with the WAIS giving higher scores than the WISC and the WISC higher scores than the WISC-R. However, Lukens (1988) found no differences between LM and SB4 scores for a group of 31 retarded adolescents.

When gifted and non-exceptional children are considered, there are also some data suggesting score instability across measures. Thorndike, Hagen, and Sattler (1986) compared 82 gifted children's LM and SB4 scores and found the LM scores to be an average of 13.5 points higher ($t=8.66$, $p<.01$). The average difference for a sample of 19 gifted children between the SB4 and the WISC-R was, in contrast, only 1.4 points ($p>.05$). For four samples of non-exceptional children, differences between the SB4 and other measures (LM, WPPSI, WISC-R, and WAIS-R) ranged from 2.3 to 5.0 points, scores on the SB4 being lower in each case. Hartwig, Sapp, and Clayton (1987) found no significant difference between LM and SB4 Composite for a sample of 30 non-exceptional children. Rothlisberg (1987), however, found WISC-R full scale scores to be significantly higher (7 point difference) than SB4 composite scores for a nonexceptional sample of 32 elementary school children. Kitano and DeLeon (1988) compared the

composite scores for a nonexceptional sample of 32 elementary school children. Kitano and DeLeon (1988) compared the proportion of preschool children identified as gifted using the LM in 1985 and 1986 with the proportion identified as gifted using the SB4 in 1987 and found that a significantly lower proportion met the criterion of IQ>124 with the SB4. They also reported use of the SB4 partial composite derived from the SB4 verbal reasoning and short term memory subareas. Using the partial composite, no significant differences between observed and expected proportions of preschoolers identified as giftd were found. The same children were not tested with both instruments and so differences may be due to differential selection. However, the authors argue that there was no evidence of a decline in cognitive abililty in the community children referred for assessment. Livesay (1986) compared LM and SB4 scores for a group of 120 six year old gifted children and found a significant (8 point) difference, with LM scores higher than SB4 scores. These results suggest that SB4 scores may be lower than LM scores and, in particular, that the abstract/visual reasoning and quantitative reasoning area scores are decreasing the SB4 composite score. If SB4 scores are, indeed, lower, and are used without adjustment of the criterion for consideration as gifted, the result would be a lower proportion of children identified as gifted. Also, if the tests are assessing different constructs, the nature of the groups identified as gifted would change. For example, if children are identified as gifted using the LM, they are likely to be highly verbal and perhaps giftd in typical academic pursuits. If children are identified as gifted using the SB4, in contrast, they may be highly verbal but may also be adept at using spatial reasoning or quantitative reasoning and may form a group requiring different instructional methods.

The SB4 and the LM differ in content and emphasis. The LM emphasizes verbal reasoning skills. The 4th edition provides four subarea scores (SAS) based on a hierarchical structural model of ability as well as a composite and does not emphasize verbal reasoning as heavily. The subarea scores are based on the examiner's selection of tests to administer from the 15 subtests, each of which yields a scale score. The administration manual provides suggestions on subtests to use with gifted children. The four areas are verbal reasoning, abstract/visual reasoning, quantitative reasoning, and short term memory. Correlations reported in the technical manual between LM total score and SB4 area scores for the norming sample of 82 gifted children ranged from .09 to .40. The correlations between LM total and SB4 composite was only .27. The correlation between LM total and SB4 composite for a non-exceptional sample of 139 children was .81. Hartwig et.al. (1987) reported a correlation of .72 between LM total and SB4 composite for a sample of 30 non-exceptional children. The tests may be measuring different constructs, then, at least for qifted children.

The primary purpose of this study was to determine whether significantly different scores would be obtained from the two

Stanford Binet tests for a sample of gifted children. The relationships between LM and SB4 area scores and relationship with an achievement measure were also investigated. The following research questions were posed:

1. Are the LM IQ's, the SB4 composite scores, and the Kaufman Achievement Total Scores significantly different?

2. Are the SB4 area scores significantly different from each other?

3. What are the relationships among the LM IQ, the Binet 4th Ed. Composite score, and the Kaufman Total Achievement scores?

METHOD

The sample consisted of 51 young Caucasian children from middle and upper middle class families in a western urban area. Their developmental histories, derived from parent's reports, indicated precocious early development of cognitive linguistic abilities. Parents reported that their children spoke earlier than one would anticipate and that many of them began reading at around 4 years of age. There were no known medical or handicapping conditions that would significantly impact the child's early development. All of the children were from intact two parent families. On a scale of 0 to 9 with 9 representing a highly responsible position with a high level of education and a 0 representing the opposite condition, the father's occupation's were scaled with a mean score of 7.2 and the mother's occupations at 5.7 (Mercer & Lewis, 1977).

Table 1 is a description of this sample of children. The data were gathered from a University Child Assessment Center where a large number of bright young children are seen regularly to provide educational program recommendations for their parents and to determine eligibility for these children for enrollment in special programs for gifted children.

Parents referred their children to the clinic for evaluations because of their interest in their child's cognitive developmental status or at the recommendation of their child's pediatrician or preschool teacher. Advanced graduate students, supervised by University faculty members, completed the evaluations in 4 to 5 testing sessions. A summary, recommendations, and a written report of findings was given to parents upon completion of the evaluation. The data presented in this study are based on all children from the University Center who had taken both the LM and SB4.

Table 1

Characteristics of the sample

Variable	N	Range	Mean	s.d.
Number of subjects	51			
Sex				
Male	33			
Female	18			
Age (months) on Binet LM		36-131	75	32
Age (months) on Binet 4th Ed.		44-144	81	32
Socio-Economic Status (0-9 scale)				
Fathers			7.2	
Mothers			5.7	

Both the LM and the SB4 data indicate that this is a sample of very bright young children (Table 2). Their mean LM IQ of 144 is more than two standard deviations above the mean. The distribution of IQ's was negatively skewed. Forty two of the 51 children had LM IQ's of 132 or higher. The mean SB4 composite score of this sample was 134 which is 2 standard deviations above the mean.

Table 2

Ranges, means, and standard deviations of the Binet LM, Binet 4th ed., and Kaufman Total Achievement standard scores

Variable	Range	Mean	s.d.
LM IQ	110 - 164	144	14
SB4 Verbal Reasoning SAS	112 - 164	135	13
SB4 Abstract/Visual Reasoning SAS	92 - 164	121	15
SB4 Quantitative Reasoning SAS	94 - 164	132	17
SB4 Memory SAS	110 - 164	127	19
SB4 Composite Score	110 - 164	134	15
Kaufman Total Achievement SAS	110 - 164	131	12

RESULTS

LM IQ scores and SB4 composite scores were significantly different (Table 3) but were also significantly positively correlated (r = .57).

There were significant differences in mean SAS scores among the subtests of the SB4 for this sample of children. The mean Abstract/Visual Reasoning score is 121 and was significantly lower than the mean Verbal Reasoning and Quantitative Reasoning scores of 135 (t=5.51, p<.000) and 132 (t=-5.50, p<.000).

A scatter plot of LM IQ's with the SB4 composite scores identified very few outliers. In all but one case those that were found represented children with substantially lower SB4 than LM scores. For instance, one child had an LM IQ of 160 and an SB4 composite score of 110. Another had an LM IQ of 144 and an SB4 score of 112 and a third child had an LM IQ of 160 and a SB4 score of 121. There were only 9 children with higher SB4 composite scores than LM IQ's. One child's pattern was exactly opposite of that described above. He had a SB4 score of 146 and LM score of 119. His verbal reasoning score was at the 93rd percentile but the other area scores were at or above the 99th percentile. His Kaufman Arithmetic standard score was 130 but the other achievement scores ranged from 101 to 114.

Table 3

Differences between the Binet LM, Binet 4th Edition, and Kaufman Total Achievement scores

Variables	N	Mean	s.d.	t	df	p
Binet LM IQ	51	144.09	14.42	5.46	50	p<.000
Binet 4th ed.	51	133.8	14.56			
Kaufman Total Ach.	27	131.3	12.4	-3.56	26	p<.001
Binet LM IQ	27	141.4	15.3			
Kaufman Total Ach.	27	131.3	12.4	.53	26	p<.597
Binet 4th ed.	27	129.6	13.7			

An examination of Kaufman Achievement subtest and total scores indicates that they are more similar to SB4 scores than to the Binet LM IQ scores. Twenty-seven of the 51 children had taken the Kaufman Achievement test. This result differs in part from that reported by Hayden, Furlong, and Linnemeyer (1988) who found the 32 gifted children they tested scored significantly higher on the SB4 than on the Kaufman ABC. The sample used in both the present study and the Hayden et.al. study had approximately equal SB4 composite scores. The Hayden et.al.

sample was on the average two years older than the sample used in the present study.

The relationship of the SB4 area scores and the composite score to the LM IQ was examined. The correlations ranged from .25 to .63. The finding of major interest was the significant relationship of the LM IQ with the SB4 composite score and the significant relationship of the Verbal Reasoning SAS and Quantitative Reasoning SAS with the LM IQ. Livesay (1986) found correlations of .64 (LM - SB4 composite), .56 (LM - Verbal Reasoning), and .30 (LM - Quantitative Reasoning). The LM scale has often been regarded as a measure with a high verbal component. These findings would seem to support this idea. The lower correlation of the Abstract/Visual Reasoning SAS with the Binet-LM scores seems to be a further indication of its distinctiveness when compared to the other three area scores.

Table 4

Relationships among scores on the two Binet scales

Binet 4th ed.	Binet LM IQ
Verbal Reasoning SAS	.63**
Abstract/Visual Reasoning SAS	.25*
Quantitative Reasoning SAS	.48**
Memory SAS	.39**
Composite Score	.57**

* $p<.05$, ** $p<.01$.

DISCUSSION

The significant LM and SB4 composite score differences in bright young children may have an impact on decisions concerning educational programming and placement. Analyses from this study indicate that the two test scores are related but the mean scores are significantly different. A possible effect of this finding is that children who are eligible for placement in classes for gifted childen on the basis of the LM IQ (IQ 132 is a commonly used figure), may not be eligible on the basis of the SB4 composite score. Ten of the 51 children in this sample would have been affected by this difference in scores. Although program eligibility guidelines typically state that admission is based on a number of characteristics, the numerical score (IQ) seems to be heavily weighted in practice. The quantitative basis for admission may need to be redefined and different scores will constitute eligibility depending on the instrument used.

The two Binet tests are different in format and construction. Different items and the manner in which items and subsections are weighted may affect the total score. Intellect,

as a construct, is represented differently on the two instruments. On the LM, a substantial verbal component has been identified. The SB4 has 4 area scores which along with the total score are said to represent intellect. Children in this sample performed very well on verbal and quantitative tasks but tended to be less successful with abstract/visual reasoning tasks. Considering the present day concern for multi-measure bases for determining program eligibility and the long standing concept of individual differences, might we have gifted children with strong verbal talents but a different kind of gifted child with strong visual spatial abilities? This has curricular implications for programs for gifted children. Perhaps the SB4 Abstract/Visual Reasoning subtests will serve as a screening instrument to identify the visual spatial child and the more verbally oriented gifted child will show a superior performance on the verbal reasoning subtests.

The mean standard scores of the Kaufman Achievement test are more similar to the SB4 than the LM IQ. If the Binet scales are to be used as predictors of academic success in gifted programs, then the SB4 would seem to be a better choice than the LM for identifying gifted children.

The SB4, a new instrument, has much to recommend it as a basis for identifying and describing some of the intellectual characteristics of young gifted children. Redefinitions of program eligibility based on the different scales may be needed and a child with certain different intellectual characteristics may become more common in our gifted classes. These changes are possible and the outcomes can be very positive for children.

REFERENCES

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

Hartwig, S.S., Sapp, G.L., and Clayton, G.A. (1987). Comparison of the Stanford-Binet intelligence scale, Form L-M and the Stanford-Binet intelligence scale fourth edition. Psychological Reports, 60, 1215-1218.

Hayden, D.C., Furlong, M.J. & Linnemeyer, S. (1988). A comparison of the Kaufman Assessment Battery for Children and the Stanford-Binet IV for the assessment of gifted children. Psychology in the Schools, 25, 239-243.

Kaufman, M., & Kaufman, M. (1983). Kaufman Assessment Battery for Children. Circle Pines, Minnesota: American Guidance Services.

Kitano, M.K., & deLeon, J. (1988). Identification of gifted children. Roeper Review, 10, 156-159.

Kroske, W.H. Fretwell, L.N., & Cupp, M.E. (1965). Comparison of the Kahn Intelligence Tests: Experimental Form, the Stanford-Binet and the WAIS for familial retardates. Perceptual and Motor Skills, 21, 428.

Livesay, K. (1986). Comparisons of the Stanford Binet: Fourth Edition to the Stanford Binet: L-M and WISC-R with gifted referrals. Paper presented at the Annual Conference of the Florida Association of School Psychologists, Jacksonville.

Lukens, J (1988). Comparison of the Fourth Edition and the L-M Edition of the Stanford-Binet used with mentally retarded persons. Journal of School Psychology, 26, 87-89.

Mercer, J. & Lewis, J. (1977). System of multicultural pluralistic assessment: Parent interview manual. New York: The Psychological Corporation.

Rothlisberg, B.A. (1987). Comparing the Stanford-Binet, Fourth Edition to the WISC-R: A Concurrent validity study. Journal of School Psychology. 25, 193-196.

Spitz, H.H. (1983). Intratest and intertest reliability and stability of the WISC, WISC-R, and WAIS Full Scale IQs in a mentally retarded population. Journal of Special Education, 17, 69-80.

Spitz, H.H. (1986). Disparities in mentally retarded persons' IQs derived from different intelligence tests. American Journal of Mental Deficiency, 90, 588-591.

Terman, L.M. & Merrill, M.A. (1973). Stanford Binet Intelligence Scale: Manual for the third revision Form L-M. Chicago: The Riverside Publishing Company.

Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986). Technical manual Stanford-Binet Intelligence Scale: Fourth Edition. Chicago: The Riverside Publishing Company.