DOCUMENT RESUME

ED 291 518 PS 017 201

AUTHOR Bauer, Joseph J.; Smith, Douglas K.
TITLE Stability of the K-ABC and S-B:4 with Preschool Children.
PUB DATE Apr 88
NOTE 16p.; Paper presented at the Annual Meeting of the National Association of School Psychologists (Chicago, IL, April 5-10, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Early Childhood Education; *Educational Testing; *Intelligence Tests; Middle Class; *Preschool Children; Preschool Education; *Preschool Tests; *Test Reliability
IDENTIFIERS *Kaufman Assessment Battery for Children; *Stanford Binet Intelligence Scale

ABSTRACT

Stability of performance on the Kaufman Assessment Battery for Children (K-ABC) and the Stanford-Binet Intelligence Scale: Fourth Edition (S-B:4) over a 1-year interval was examined with a sample of 28 nonhandicapped preschoolers. Each child was administered both tests in counterbalanced order and retested in 1 year with either the K-ABC or the S-B:4. Results indicated high stability for the K-ABC global scales and individual subtests. The verbal reasoning cluster of the S-B:4 seemed to provide a stable measure at the preschool level, but adequate stability was not demonstrated by the other area scores or the test composite. (PCB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED291518

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Stability of the K-ABC and S-B: 4 with Preschool Children

Joseph J. Bauer and Douglas K. Smith

Department of Psychology

University of Wisconsin-River Falls

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Douglas K. Smith

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the Annual Meeting of the National Association of School Psychologists, Chicago, April 1988

Running head: STABILITY

PS 018 201

BEST COPY AVAILABLE

Abstract

Stability of K-ABC and S-B: 4 performance over a one year interval was examined with a sample of 28 nonhandicapped preschoolers. Each child was administered the K-ABC and S-B: 4 in counterbalanced order and retested in one year with either the K-ABC or S-B: 4. Global scale stability coefficients (corrected for restriction in range as needed) for the K-ABC were Mental Processing Composite= .93; Sequential Processing = .80; Simultaneous Processing = .80 and Achievement = .84 (all significant at $p < .001$). Stability coefficients for mental processing subtests were all significant ($p < .01$) and ranged from .55 (Hand Movements) to .96 (Matrix Analogies). On the S-B: 4 stability coefficients (corrected for restriction in range as needed) included: Test Composite = .20; Verbal Reasoning = .85 ($p < .001$); Abstract/Visual Reasoning = .31; Quantitative Reasoning = -.36; and Short-Term Memory = -.25. Stability coefficients for four of eight subtests were significant ($p < .01$): Vocabulary (.73), Comprehension (.54), Absurdities (.56) and Pattern Analysis (.54).

As the emphasis on preschool assessment has increased, new instruments that can be used to assess preschool children's abilities and skills have been developed. These include the Kaufman Assessment Battery for Children (K-ABC; A. Kaufman & N. Kaufman, 1983) and the Stanford-Binet Intelligence Scale: Fourth Edition (S-B: 4; Thorndike, Hagen & Sattler, 1986). In order to establish validity for these instruments, studies comparing the performance of preschoolers on the scales have been conducted and are reported in the respective test manuals and the professional literature.

The stability of performance among preschool children on these scales, however, has received little attention. Limited test-retest data are provided in the test manuals. For the S-B: 4, a preschool sample of 57 children (mean age of 5 years, 2 months) was retested after an interval of two to eight months and area score reliabilities ranged from .71 (Quantitative Reasoning) to .91 (Test Composite). A sample of 84 preschool children (ages 2 1/2 through 4) was retested with the K-ABC after an interval of two to four weeks and global scale reliabilities ranged from .77 (Sequential and Simultaneous Processing) to .95 (Achievement Scale). Of several recent studies that have examined stability of performance, only three (Lyon & Smith, 1987; Telzrow, Proefrock & Hartlage, 1985; Valencia, 1985) have used preschool samples. Telzrow et al.'s study included 26 preschool children who had been identified as high-risk infants and were tested at ages 3, 4, 5 and

6 with either the Stanford-Binet (Form L-M) or WPPSI. Correlations with age 6 Binet IQ ranged from .48 for age 5 Binet IQ to .56 for age 4 WPPSI IQ. Valencia's study involved 42 Mexican-American children enrolled in a Head Start program. Each student was tested with the K-ABC in Spring 1983 and retested in Fall 1983. Over this time interval, global scale stability coefficients ranged from .76 to .90. Lyon and Smith utilized a sample of 53 at-risk students enrolled in early intervention programs and tested with the K-ABC in Fall 1984 and retested in May 1985. Global scale stability coefficients ranged from .73 to .83. Stability studies using the K-ABC or the S-B: 4 with nonhandicapped students could not be located in the research literature.

Purpose of the Study

Stability of scores on individual intelligence tests for preschool students is important as academic performance is often predicted on the basis of such scores. And yet this issue has received little attention in the research community. Therefore, the present study was designed to investigate the stability of the K-ABC and S-B: 4 using a sample of nonhandicapped children from a suburban, middle-class community in the midwest. The K-ABC and S-B: 4 were selected because they are the most recently developed individually administered intelligence tests that are used with preschool children.

Method

Subjects

The sample consisted of 28 children (15 males and 13 females) attending a daycare center located in a suburban area of a large midwestern city. The parents of 40 children were randomly selected and asked to participate in the study with the parents of 30 children agreeing to participate (participation rate of 75%). Two children subsequently withdrew from the project. Parent educational level ranged from high school to post college with the majority of parents having a college degree. At the time of initial testing (Time 1) the children ranged in age from 3 years, 11 months to 6 years, 2 months with a mean age of 4 years, 11 months.

Procedure

At Time 1 testing (Fall 1985), each child was administered the K-ABC and S-B: 4 in counterbalanced order. The average length of time between tests was 11 days with a range of 4 to 21 days. Time 2 testing occurred in Fall 1986. One-half of the sample was randomly selected to be retested with the K-ABC and the remainder with the S-B:4. All testing was completed by school psychologists trained in the administration and interpretation of the two tests. All protocols were checked for scoring accuracy prior to being included in the data analysis.

Results

Mean scores on the K-ABC global scales, Mental Processing Composite (MPC), Simultaneous Processing (SIM), Sequential Processing (SEQ) and Achievement (ACH), and the S-B: 4 Test Composite and area scores, Verbal Reasoning (VR), Abstract/Visual Reasoning (AVR), Quantitative Reasoning (QR) and Short-Term Memory (STM), were all in the average range at Time 1 testing. Similar results were obtained at Time 2 testing.

Pearson product moment correlations were computed for the global scales of the K-ABC and S-B:4 and corrected for restriction in range (as needed) using the formula developed by Guilford (1954). This produced stability coefficients ranging from .80 (SEQ and SIM) to .93 (MPC) on the K-ABC and from -.36 (QR) to .85 (VR) on the S-B:4. Mean scores, standard deviations and correlations for the K-ABC global scales and S-B:4 areas are presented in Table 1.

Insert Table 1 about here

Mean scores for the subtests of the K-ABC and S-B:4 were all in the average range at both Time 1 and Time 2 testing. Pearson product moment correlations were also calculated for the subtests. Stability coefficients (corrected for restriction in range) ranged from .55 (Hand Movements) to .96 (Matrix Analogies) on the K-ABC and from -.36 (Quantitative) to .73 (Vocabulary) on the S-B: 4.

Mean scores, standard deviations and correlations for the subtests of the K-ABC and S-B: 4 are presented in Table 2.

Insert Table 2 about here

Discussion

Stability coefficients (corrected for restriction in range) for the global scales of the K-ABC are all significant ($p < .001$) and range from .80 to .93. These results compare favorably with the test-retest results of 84 preschool children retested with the K-ABC at an average interval of 18 days. Those results, as reported in the K-ABC Interpretive Manual, ranged from .77 to .95 for the global scales. In the present study, stability coefficients (corrected for restriction in range) for individual subtests ranged from .55 to .96 as compared to a range of .62 to .87 in the test-retest study. Using $r = .70$ as a standard, all global scales exhibited high stability over the one year test-retest interval with four of seven mental processing subtests (57%) and three of four achievement subtests (75%) exhibiting high stability. A moderate degree of stability was displayed by the remaining subtests ($r = .55$ or above).

These results suggest that more caution should be used in the interpretation of individual subtests as their level of stability, although moderate to high, is somewhat less than the level of stability for the global scales. Since the global scales are

composed of clusters of individual subtests, these results are to be expected. Nevertheless, the practicing school psychologist is advised to exercise some degree of caution in the interpretation of such individual subtests as Hand Movements and Arithmetic as their stability coefficients were less than .60 (.55 and .58, respectively).

Stability results for the K-ABC in this sample of nonhandicapped, preschool students were similar to results in other studies with handicapped samples (e.g. Lyon & Smith, 1987; Telzrow, Proefrock & Hartlage, 1985; Valencia, 1985). For example, global scale stability coefficients ranged from .73 to .83 in one study (Lyon & Smith, 1987) and from .76 to .90 in another (Valencia, 1985). Although the research literature to date is limited, it is consistent in indicating a high level of stability for the K-ABC with handicapped and nonhandicapped preschool students. Since the number of studies is limited and sample sizes are small, stability of scores should continue to be investigated.

A more variable pattern of results is presented by the S-B: 4. Stability coefficients (corrected for restriction in range) ranged from -.36 to .85 for the Test Composite and area scores. Only the VR stability coefficient of .85 ($p < .001$) was significant. In the present study, much variability was exhibited in subtest stability coefficients also with a range of -.36 to .73. Using $r = .70$ as a standard, only the VR area score showed high stability with the Test Composite and other area scores showing little stability as

the stability coefficients were all less than .33. Similarly, high stability was only shown by one subtest (Vocabulary, r = .73) of eight (13%) and moderate stability was shown by three subtests (39%). Minimal stability was shown by four of the eight subtests (50%). This pattern is quite different from that reported in the S-B: 4 Technical Manual in which a preschool sample of 57 was retested after two to eight months and test-retest reliabilities were .71 to .91 for the Test Composite and area scores. Clearly, this issue merits further investigation as sample sizes were small and test-retest intervals were not the same.

In order to further explore the stability of scores on the S-B: 4, the results were reanalyzed using the factor scores for preschoolers proposed by Sattler (1988). Stability of the factor scores is important for Sattler (1988, p. 261) argues that they should be used for interpretation adding that since the "area scores are not supported by factor analysis, they should not be used for most interpretive purposes." At the preschool level there are two factors: Verbal Comprehension (VC), composed of Vocabulary, Comprehension, Absurdities and Memory for Sentences; and Nonverbal Reasoning/Visualization (NRV), composed of Pattern Analysis, Copying, Quantitative nd Bead Memory. Using the procedures outlined by Sattler (1988), these factor scores were calculated for both the test and retest scores. Stability coefficients were not significant and included .26 for VC and .11 for NRV. In the case of the VC factor, the addition of the Memory

for Sentences subtest appears to have greatly reduced the stability of the factor score as the original VR area score produced a stability coefficient of .85.

An examination of Table 2 indicates that the subtests Copying, Quantitative, Memory for Sentences and Bead Memory were the least stable for this group of preschool students. Thus, they should be interpreted cautiously. Likewise, only the VR area score demonstrated a sufficient level of stability. The Test Composite and other area scores including the factor scores proposed by Sattler (1988) should be interpreted cautiously. Whether these results are an artifact of the present sample or reveal a basic problem with the stability of the S-B: 4 at preschool ages is unknown. Additional studies investigating this issue are needed.

Conclusions

The results of this study of 28 nonhandicapped, preschool students indicate high stability for the K-ABC global scales and individual subtests. These results are consistent with results of previous studies using handicapped samples. Thus, the studies to date suggest the K-ABC provides a stable measure of the constructs it purports to measure at the preschool level for both handicapped and nonhandicapped students.

For the S-B: 4, high stability is shown by the VR area score and Vocabulary subtest with moderate stability demonstrated by three other subtests (Comprehension, Absurdities and Pattern Analysis). The VR cluster of the S-B: 4 seems to provide a stable

measure at the preschool level. Adequate stability was not demonstrated by the other area scores or Test Composite. Since the sample size in this study was small and other studies examining stability of the S-B: 4 at the preschool level are lacking, these results of the study should be interpreted cautiously and the issue investigated in future research. Until then, the practicing school psychologist may be wise to use the S-B: 4 in combination with other age-appropriate measures.

References

Guilford, J. (1954). Psychometric methods (2nd ed.). New York: McGraw-Hill.

Kaufman, A., & Kaufman, N. (1983). Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

Lyon, M., & Smith, D. (1987). Stability of the Kaufman Assessment Battery for Children for a sample of at-risk preschool children. Psychology in the Schools, 24(2), 111-115.

Sattler, J. (1988). Assessment of children (3rd ed.). San Diego, CA: Jerome Sattler, Publisher.

Telzrow, C., Proefrock, V., & Hartlage, L. (1985). Stability of preschool scales as predictors of six year IQ. Los Angeles. (Paper presented at Annual Meeting of the American Psychological Association)

Thorndike, R., Hagen, E., & Sattler, J. (1986). Stanford-Binet Intelligence Scale: Fourth Edition. Chicago, IL: Riverside Publishing Company.

Valencia, R. (1985). Stability of the Kaufman Assessment Battery for Children for a sample of Mexican-American children. Journal of School Psychology, 23(2), 189-193.

Table 1

Means, Standard Deviations and Correlations for the K-ABC and S-B:4 at Time 1 and Time 2 Testing

	Time 1		Time 2		
	Mean	SD	Mean	SD	r
K-ABC					
Mental Processing Composite	109.50	13.17	111.84	11.56	.93*
Simultaneous Processing	110.07	12.63	110.93	11.91	.80*
Sequential Processing	106.57	16.97	110.50	13.51	.80*
Achievement	106.50	10.88	105.43	10.98	.84*
S-B: 4					
Test Composite	107.29	9.52	105.14	8.12	.20
Verbal Reasoning	112.50	8.71	109.29	10.18	.85*
Abstract Visual Reasoning	105.50	13.48	104.36	10.02	.31
Quantitative Reasoning	100.86	11.99	99.00	15.23	-.36
Short-Term Memory	105.50	10.82	104.64	14.79	-.25

Note. Number of children for each testing was 24. Time 1 = Fall 1985; Time 2 = Fall 1986. Correlations have been corrected for restriction in range as needed (based on standard deviation at Time 1 testing) using Guilford's (1954) procedure.

*p < .001 **p < .01 ***p < .05

Table 2

Means, Standard Deviations and Correlations for the K-ABC and S-B:4 Subtests at Time 1 and Time 2 Testing

	Time 1		Time 2		
	Mean	SD	Mean	SD	
K-ABC					
Hand Movements	10.50	3.21	10.79	2.16	.55**
Gestalt Closure	10.71	2.27	11.93	2.87	.82*
Number Recall	11.07	3.67	12.50	2.28	.75*
Triangles	11.71	3.07	12.36	2.41	.67*
Word Order	11.50	2.35	11.64	2.82	.67*
Matrix Analogies (n = 8)	11.25	1.28	11.13	1.46	.96*
Spatial Memory (n = 8)	12.38	2.67	11.63	2.07	.81*
Faces & Places	102.14	10.88	102.29	12.79	.91*
Arithmetic	104.29	13.57	101.79	13.05	.58**
Riddles (n = 13)	108.23	8.94	108.23	8.63	.91*
Reading/Decoding (n = 8)	107.75	15.22	104.63	19.68	.71***
S-B: 4					
Vocabulary	53.86	5.76	55.36	5.79	.73*
Comprehension	55.64	5.57	53.36	5.29	.54**
Absurdities	56.71	3.12	53.64	5.89	.56**
Pattern Analysis	52.43	6.96	50.86	6.88	.54**
Copying	52.43	7.18	53.00	7.64	-.08
Quantitative	50.43	6.00	49.50	7.61	-.36
Memory for Sentences	51.36	4.57	53.71	7.10	-.02

					Stability
					15
Bead Memory	53.29	6.96	51.00	9.46	.22

Note. Number of children for each testing was 24 unless otherwise noted. Time 1 = Fall 1985; Time 2 = Fall 1986. Correlations have been corrected for restriction in range as needed (based on standard deviation at Time 1 testing) using Guilford's (1954) procedure.

*p < .001 **p < .01 ***p < .05