DOCUMENT RESUME

ED 291 519 PS 017 202

AUTHOR Smith, Douglas K.; And Others
TITLE K-ABC Stability in a Preschool Sample: A Longitudinal Study.
PUB DATE Apr 88
NOTE 21p.; 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; Longitudinal Studies; Middle Class; *Preschool Children; Preschool Education; *Preschool Tests; *Test Reliability; *Young Children
IDENTIFIERS *Kaufman Assessment Battery for Children

ABSTRACT

Over a 2-year period, stability of performance on the Kaufman Assessment Battery for Children of 25 nonhandicapped preschool children was examined. Overall, results indicated a high level of stability for both global scale scores and subtest scores. Less stability was indicated for performance patterns and subtest strengths and weaknesses. (PCB)

K-ABC Stability in a Preschool Sample: A Longitudinal Study

Douglas K. Smith

Jean A. Bolin

Donald L. Stovall

Department of Psychology

University of Wisconsin-River Falls

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

Running head: K-ABC STABILITY

Abstract

Stability of K-ABC performance over a two year interval with 25 nonhandicapped preschool children was examined. Global stability coefficients (corrected for restriction in range) for Time 1 testing (age 4) and Time 3 testing (age 6) were: MPC = .77; SEQ = .73; SIM = .76 and ACH = .80. Subtest stability coefficients ranged from .53 to .75 for the Time 1/Time 3 comparisons. Repeated measures analyses of variance revealed significant increases in mean SIM score from age 4 to age 6, significant decreases in mean SEQ scores from ages 4 and 5 to age 6 and a significant decrease in mean ACH score from age 4 to ages 5 and 6. Overall, the results indicated a high level of stability for both global scale scores and subtest scores. Less stability was indicated for performance patterns (SIM > SEQ, SIM < SEQ, SIM = SEQ) and subtest strengths and weaknesses with this sample of nonhandicapped preschool children tested at ages 4, 5 and 6.

In response to the emphasis on preschool assessment, many new instruments to assess preschool children's abilities and skills have been developed. Examples include the Stanford-Binet Intelligence Scale: Fourth Edition (Thorndike, Sattler & Hagen, 1986), the Kaufman Assessment Battery for Children (K-ABC; A. Kaufman & N. Kaufman, 1983) and the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R; Wechsler, in press). Although the performances of various groups of preschool children on these tests have been compared, the stability of performance among handicapped and nonhandicapped preschool children on these scales has received little attention. Limited stability data are included in test manuals and those data usually involve test-retest intervals of less than three months. Of several recent studies that have examined stability of performance, only three (Lyon & Smith, 1987; Telzrow, Proefrock & Hartlage, 1985; Valencia, 1985) have included preschool samples. Lyon and Smith (1987) utilized a sample of 53 at-risk preschool children, Telzrow et al. (1985) employed a sample of 26 preschool children who had been identified as high-risk infants and Valencia's sample (Valencia, 1985) involved 42 Mexican-American children enrolled in a Head Start program. No stability studies with nonhandicapped preschool children, however, were indicated in the current research literature.

Purpose of the Study

As new scales for preschool assessment are developed, it is important that the stability of performance on the instruments be examined. This issue assumes even greater importance for preschool students as academic performance is often predicted from children's performance on these instruments and "young children's behavior (limited attention, expressive problems, etc.) contributes to a generally less reliable and valid measure" (Prasse, 1983, p. 46). At the same time it is argued that the "preschool years of 3-5 represent the transition between poor and relatively good predictions" (Lidz, 1983, p. 18). Thus, the stability of performance or relationship between a child's score at age 4 and the score on the same instrument at ages 5 or 6 is a crucial issue for both handicapped and nonhandicapped children. Since children's performance across scales and subtests is often used to develop intervention plans, the stability of patterns of test performance or how children obtained their overall score is a critical issue, also. Therefore, the present study was designed to investigate K-ABC stability over a two year period by using a sample of nonhandicapped children from a middle-class, suburban community in the midwest.

Method

Subjects

Subjects for this study included 25 nonhandicapped children (13 boys and 12 girls), ranging in age from 44 months to 58 months (mean = 49, SD = 3.8) at the May 1985 (Time 1) testing, from 57 months to 68 months (mean = 62, SD = 3.6) at the June 1986 (Time 2) testing and from 71 months to 83 months (mean = 77, SD = 3.5) at the August 1987 (Time 3) testing. Parental educational levels ranged from high school to post college with the majority of parents having a college degree. Eight children who participated in the original testing (Time 1 and Time 2) moved from the area and did not participate in the Time 3 testing. (Their scores on the global scales of the K-ABC at Time 1 and Time 2 testing did not differ significantly from the scores of the 25 children in the present study).

Procedure

The original sample was selected randomly from children successfully completing the school district's screening program for four year old children. Participation rate in this phase of the study was approximately 95%. On a yearly basis, parents were contacted and permission for evaluation was obtained. Each child was evaluated with the K-ABC by school psychologists trained in the administration and interpretation of the K-ABC. All protocols were checked for scoring accuracy before being included in data analysis. Testing occurred in May 1985, June 1986 and August 1987.

Results

Mean scores on the K-ABC global scales, Mental Processing Composite (MPC), Simultaneous Processing (SIM), Sequential Processing (SEQ) and Achievement (ACH), were concentrated in the high average range at all three test times. The difference between highest and lowest mean global scale scores was 3.04 at Time 1, 4.72 at Time 2 and 5.84 at Time 3. Mean scores, standard deviations and range are presented in Table 1.

Insert Table 1 about here

Pearson product moment correlations were computed for the K-ABC global scales and corrected for restriction in range using the formula developed by Guilford (1954). This produced stability coefficients ranging from .56 ($p < .01$) for SEQ at Time 1/Time 2 to .93 ($p < .001$) for ACH at Time 2/Time 3. The complete table of correlations for K-ABC global scales is presented in Table 2.

Insert Table 2 about here

Pearson product moment correlations were also calculated for K-ABC mental processing and achievement subtests. Stability coefficients (corrected for restriction in range) for all subtests were significant ($p < .01$). The complete table of correlations is presented in Table 3.

Insert Table 3 about here

Differences in mean global scale scores over time were analyzed using a repeated measures analysis of variance. Significant results were obtained for SIM with $F(2,48) = 3.99$, $p < .05$, SEQ with $F(2,48) = 3.58$, $p < .05$ and ACH with $F(2,48) = 9.58$, $p < .001$. Post hoc comparisons using the protected t-test indicated that the mean SIM score at age 4 was significantly lower than the mean SIM scores at ages 5 and 6; mean SEQ scores at ages 4 and 5 were significantly higher than mean SEQ score at age 6; and mean ACH score at age 4 was significantly higher than mean ACH scores at ages 5 and 6. Thus, the MPC remained quite constant while the other global scales showed some variation.

In order to determine if the relationship of the global scales to each other differed over time, a repeated measures analysis of variance was conducted for global scale scores at Time 1, Time 2 and Time 3. Significant results were obtained at Time 3 only with $F(3, 72) = 6.42$, $p < .001$. Post hoc comparisons utilizing the protected t-test indicated that the mean SEQ and ACH scores were significantly lower ($p < .05$) than the mean SIM and MPC scores. At Time 1 and Time 2 the mean global scale scores did not differ significantly from each other, while at Time 3 differences were noted.

Chi square analyses of performance patterns (SIM > SEQ, SIM < SEQ or SIM = SEQ) across the three test times were not significant. At each testing approximately 63% of the children did not show a preference for SIM or SEQ processing as shown by a significantly higher SIM or SEQ score. Similar analyses of MPC/ACH patterns across test times were not significant with approximately 71% of the children showing nonsignificant differences between ACH and MPC scales.

Significant subtest strengths or weaknesses (as determined by the procedure outlined in the K-ABC manual) were exhibited by 15 students (60%) at Time 1 testing (age 4) and by 16 students (64%) at Time 3 testing (age 6). Although significant subtest strengths or weaknesses were exhibited by 10 students at both Time 1 and Time 3 testing, they were not the same subtest strengths and weaknesses. Five subtest strengths remained strengths across test times while five became nonsignificant (neither strengths nor weaknesses). Likewise, two subtest weaknesses remained weaknesses and six weaknesses became nonsignificant (neither strengths or weaknesses) at Time 3 testing. Finally, six new subtest strengths and six new subtest weaknesses emerged.

Discussion

A high level of stability between and among scores on the global scales of the K-ABC at ages 4, 5 and 6 is indicated by the results of this study. All correlations were significant ($p < .01$) and ranged from .73 to .80 with an average correlation of .76 for the age 4/age 6 comparison. Individual subtest scores also showed a high level of stability, although the overall magnitude of correlation was somewhat less than the global scale correlations (average correlation of .66 and range of .63 to .75 for the age 4/age 6 comparison).

At age 5 the composition of the Mental Processing and Achievement scales are changed with Magic Window and Face Recognition replaced by Matrix Analogies and Spatial Memory on the SIM scale and Expressive Vocabulary replaced by Reading/Decoding on the ACH scale. Despite these changes in global scale composition, the SIM and ACH stability coefficients for age 4/age 5 and age 4/age 6 comparisons were significant and ranged from .71 to .89. Thus, a high degree of stability is indicated with the K-ABC global scales. Although these results need to be supported by other studies with different samples, the preliminary indication is that the measurement of the constructs represented by the global scales has not been compromised by varying the subtest composition of the scales. In addition, the K-ABC appears to provide continuity of measurement at this preschool age range (approximately four to six years).

These results compare favorably with the test-retest data for preschool children provided in the K-ABC Interpretive Manual. Those correlations, obtained with 84 preschool children retested with the K-ABC at an average interval of 18 days, ranged from .77 to .95 for the global scales and from .62 to .87 for the individual subtests. Likewise, results of the present study are consistent with Valencia's (1985) study of Mexican American preschoolers and Lyon and Smith's (1987) study of at-risk preschoolers. In these studies global scale stability coefficients (corrected for restriction in range) ranged from .76 to .90 (Valencia, 1985), from .78 to .88 (Lyon & Smith, 1987) and from .73 to .80 in the present study. Despite differing preschool samples and retest intervals (Head Start children retested at four to six months, at-risk children retested at a nine month interval and nonhandicapped children retested with a two year interval), very similar global scale stability coefficients are indicated. In addition, the stability coefficients compare favorably with studies of other tests such as the Stanford-Binet (Telzrow et al., 1985).

Of the four global K-ABC scores, the MPC exhibited the highest level of stability across time. The ACH scale, on the other hand, exhibited the greatest variability with age 4 scores significantly higher than age 5 and age 6 scores (approximately .4 of a standard deviation), suggesting that scores on this scale at younger ages may not be as predictive of future achievement as scores at older ages. Since the ACH scale was designed to reflect children's learning, these changes may be a result of the varied learning experiences of the children in the study. Additional explanations for this finding include the relatively high educational level of the parents in this study, preschool experience or, perhaps, the nature of the ACH subtests at younger ages. The majority of the parents had a college degree and they may have provided a somewhat enriched home environment for the children. Likewise, 88% of the children had preschool experiences ranging from a few hours per week to daily experiences and this may have affected the age four ACH scores. Finally, the nature of the ACH scale itself may be responsible as Kamphaus and Reynolds (1987) suggest that the ACH scale may be composed of a verbal intelligence component (Expressive Vocabulary, Faces & Places, Riddles, Arithmetic) and a reading component (Reading/Decoding, Reading/Understanding). Thus, at Time 1 and Time 2 testing for this sample (ages 4 and 5), the ACH score may have been more a measure of verbal skills than achievement. It should be noted that for individual children the correlational results suggest considerable stability in ACH scores even though as a group the mean ACH score at Time 3 is significantly lower than the mean ACH score at Time 1 and Time 2. This issue (stability of the achievement scale) is complex and merits further investigation.

As previously indicated the MPC remained stable from ages 4 to 6. This stability was apparently accounted for by the increase in SIM scores and the accompanying decrease in SEQ scores. For both the SIM and SEQ scales the change in levels was approximately one-third of a standard deviation. Whether this change is age-related or an artifact of the present sample is unknown but should be explored.

Another element of stability involves the relationship of the global scales to each other over time. At ages four and five there were no significant differences among the mean global scale scores. At age six, however, mean SEQ and ACH scores were significantly lower than the mean MPC and SIM scores. The difference in mean scores were modest and ranged from 4.92 (ACH/MPC) to 8.12 (SEQ/SIM) and may be related to a decrease in mean SEQ and ACH scores from ages four and five. The reasons for these decreases are not clear and may be the result of age-related changes, an artifact of the K-ABC or an artifact of the current sample. Clearly, this issue deserves further investigation, since programming decision may be based on comparisons of SIM with SEQ and MPC with ACH.

A final element of stability involves performance patterns such as SIM > SEQ, SIM < SEQ or SIM = SEQ. In the current study the percentage of children exhibiting these patterns was consistent at each age level with the majority of children displaying a SIM = SEQ pattern at each age level. Of greater importance, however, is the stability of a pattern for individual children. Seven children (28%) exhibited a SIM = SEQ pattern across the three test ages, while two children (8%) exhibited a SIM < SEQ or SIM > SEQ pattern and 16 children (64%) exhibited a variable or inconsistent pattern across the three test ages. Although the number of children demonstrating a consistent pattern from one test age to the next increased with age, the results were not statistically significant. The important finding, however, is that less than half the children exhibited a stable SIM/SEQ pattern from age four to age six. Thus, a child's processing preference at age four has less than a fifty percent chance of being the same at age six. Therefore, programming decisions based on processing style at early ages should be monitored on a regular basis. Stability of processing style is an important issue that demands further investigation using samples of different ages. For example, it is possible that more stability in processing style is present with older children. It is also crucial that this issue be investigated with samples of handicapped children

The presence of subtest strengths or weaknesses was not uncommon with 14 (56%) to 16 (64%) of the children having at least one subtest strength or weakness at each age level. However, the analysis of subtest strengths and weaknesses in individual children should be undertaken cautiously. Although in no cases did subtest strengths become weaknesses or subtest weaknesses become strengths, there was considerable variability across the age range of four to six in the patterns of strengths and weaknesses. Only 50% of the indicated subtest strengths at age four remained strengths at age six and only 25% of the subtest weaknesses at age four remained weaknesses at age six. The generalizability of these results to older children or to handicapped populations is unknown and should be investigated in future studies.

Summary and Conclusions

The present study involved a middle class sample of preschool children tested with the K-ABC at ages four, five and six. Considerable stability for the global scales of the K-ABC was indicated with stability coefficients (corrected for restriction in range) ranging from .73 to .80 for age four/age six comparison. Subtest stability was also high with stability coefficients (corrected for restriction in range) ranging from .53 to .78. Mean MPC scores were quite stable across the age range of four to six, while some variation in SIM, SEQ and ACH scores was noted (.3 to .4 standard deviations). Less stability was indicated for analysis of performance patterns (SIM < SEQ, SIM > SEQ, SIM = SEQ) and subtest strengths or weaknesses. Therefore, caution is recommended in analyzing performance patterns and subtest strengths and weaknesses. Programming decisions based on these analyses should be monitored regularly to assess any changes in performance patterns of individual students. Overall, the results of the study were consistent with previous K-ABC test-retest studies (over a shorter time interval) and indicated that the K-ABC global scales and subtests were stable measures for the preschool children in this sample. Additional studies are needed to examine the stability of the K-ABC with preschool students over a longer time period and to examine performance patterns over time.

References

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

Kamphaus, R., & Reynolds, C. (1987). Clinical and research applications of the K-ABC. Circle Pines, MN: American Guidance Service.

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

Lidz, C. (1983). Issues in assessing preschool children. In K. Paget, & B. Bracken (Eds.), Psychoeducational assessment c preschool children (pp. 17-27). New York: Grune & Stratton.

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.

Prasse, D. (1983). Legal issues underlying preschool assessment. In K. Paget, & B. Bracken (Eds.), Psychoeducational assessment of preschool children (pp. 29-50). New York: Grune & Stratton.

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 American Psychological Association)

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

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.

Wechsler, D. (in press). Wechsler Preschool and Primary Scale of Intelligence-Revised. San Antonio, TX: Psychological Corporation.

Table 1

Means, Standard Deviations and Minimum/Maximum Values for the K-ABC at Times 1, 2, and 3 for 25 Nonhandicapped Students

	MPC	SEQ	SIM	ACH
Time 1 (Age 4)				
Mean	113.52	112.88	110.88	115.56
Standard Deviation	12.28	11.16	12.48	12.93
Range	94-146	98-141	93-137	92-137
Time 2 (Age 5)				
Mean	116.64	112.64	115.36	111.56
Standard Deviation	12.36	8.48	14.67	12.36
Range	99-148	95-131	96-150	86-134
Time 3 (Age 6)				
Mean	114.64	107.92	116.04	109.72
Standard Deviation	11.88	12.83	11.84	11.56
Range	97-142	91-135	100-142	84-127

Note. MPC = Mental Processing Composite; SEQ = Sequential Processing; SIM = Simultaneous Processing; ACH = Achievement; Time 1 = May 1985; Time 2 = June 1986; Time 3 = August 1987.

Table 2

Correlations of K-ABC Global Scales at Ages 4, 5 and 6

	Age 5(Time 2)	Age 6(Time 3)
MPC at Age 4 (Time 1)	.67(.75)*	.70(.77)*
MPC at Age 5 (Time 2)		.81(.85)*
SEQ at Age 4 (Time 1)	.45(.56)**	.62(.73)*
SEQ at Age 5 (Time 2)		.61(.67)*
SIM at Age 4 (Time 1)	.64(.71)*	.70(.76)*
SIM at Age 5 (Time 2)		.82(.83)*
ACH at Age 4 (Time 1)	.86(.89)*	.76(.80)*
ACH at Age 5 (Time 2)		.92(.93)*

Note. Correlations in parentheses are corrected for restriction in range. This table presents correlations of test scores of K-ABC Global Scales at ages 4 and 5 with test results obtained at ages 5 and 6 for the same 25 children. For example, the MPC correlation of age 4 testing with age 5 testing was .75 and the MPC correlation of age 4 testing with age 6 testing was .77

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

Table 3

Correlations of K-ABC Subtests at Ages 4, 5 and 6

	Age 5	Age 6
Mental Processing Subtests		
Hand Movements at Age 4	.34(.44)**	.62(.73)*
Hand Movements at Age 5		.26(.41)**
Gestalt Closure at Age 4	.58(.61)*	.68(.71)*
Gestalt Closure at Age 5		.77(.78)*
Number Recall at Age 4 (n = 24)	.58(.67)*	.48(.58)*
Number Recall at Age 5		.56(.76)*
Triangles at Age 4 (n = 17)	.74(.78)*	.70(.75)*
Triangles at Age 5		.64(.64)*
Word Order at Age 4 (n = 17)	.62(.71)*	.53(.63)*
Word Order at Age 5		.58(.76)*
Matrix Analogies at Age 4	--------	--------
Matrix Analogies at Age 5 (n = 18)	--------	.56(.69)*

Spatial Memory at Age 4	--------	--------
Spatial Memory at Age 5 (n = 18)		.48(.55)**
Achievement Subtests		
Faces & Places at Age 4	.83(.84)*	.59(.61)*
Faces & Places at Age 5		.61(.68)*
Arithmetic at Age 4	.80(.80)*	.53(.53)**
Arithmetic at Age 5		.65(.70)*
Riddles at Age 4	.68(.75)*	.64(.71)*
Riddles at Age 5		.75(.85)*
Reading/Decoding at Age 4	--------	--------
Reading/Decoding at Age 5 (n = 18)		.79(.77)*

Note. Correlation coefficients in parentheses are corrected for restriction in range. Correlations involving $n = 10$ or less have been omitted and $n = 25$, unless otherwise noted

*$p < .001$

**$p < .01$

***$p < .05$