
Rex B. Kline
Concordia University
Montréal, Québec, Canada

Joseph Snyder
Concordia University and
La Commission Scolaire Jérôme-Le Royer
Montréal, Québec, Canada

Sylvie Guilmette
Concordia University
Montréal, Québec, Canada

Maria Castellanos
La Commission Scolaire Jérôme-Le Royer
Montréal, Québec, Canada

Scatter analysis of IQ profiles has a long and controversial history. We conducted this study to determine whether the validity of scatter information is any greater for 2 new IQ batteries, the Kaufman Assessment Battery for Children (K-ABC) and the Fourth Edition Stanford-Binet (SB4), than for the Wechsler Intelligence Scale for Children-Revised (WISC-R). Within a sample of referred children, we computed numerical indexes of profile elevation, shape, and variability for all IQ tests. Using hierarchical multiple regression with achievement scores as dependent variables, we found shape information had marginal incremental validity over elevation as a predictor for the WISC-R, even less for the SB4, and virtually none for the K-ABC. We discuss implications of these results.

One of the most robust findings in assessment research is that IQ scores covary with academic achievement. A much more controversial issue, however, concerns the value of scatter analysis of IQ test profiles, the formulation of hypotheses about abilities on the basis of patterns of subtest scores. Interest in scatter analysis dates to the 1930s and 1940s, and many reports about scatter analysis still appear in the clinical and school psychology literatures (e.g., McLean, Reynolds, & Kaufman, 1990; Reynolds & Clark, 1985, 1986).

Results of much research do not, however, support the value of scatter analysis (Kramer, Henning-Stout, Ullman, & Schnellenberg, 1987). For example, ability profiles obtained from some remedial groups (e.g., learning disabled children) often do not show distinct shapes or variabilities compared with those from normal control subjects (e.g., Clarizio & Veres, 1983; Kaufman, 1976, 1981). Other results suggest that scatter information has little incremental validity over summary IQ scores in the prediction of academic achievement (e.g., Hale & Saxe, 1983; Sexton & Street, 1983).

Although scatter analysis has been a problematic enterprise, it is still possible that new or recently revised IQ batteries may yield more useful scatter information than earlier scales. For example, the Stanford-Binet was recently revised (Fourth Edition; SB4; Delaney & Hopkins, 1987; Thorndike, Hagen, & Sattler, 1986a, 1986b) and is based on a four-factor ability model. Because it potentially provides more information than Wechsler scales or earlier Stanford-Binet editions, scatter information from the SB4 may have greater external validity. The Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983a, 1983b) is also based on a specific theoretical model, one that its authors hoped would be useful in generating recommendations for remediation.

In this investigation, we studied the relative usefulness of profile scatter information from the WISC-R, K-ABC, and SB4 in the prediction of academic achievement. We administered all three IQ batteries to children referred to school psychologists and calculated for each child and for each IQ test numerical indexes of profile elevation and scatter (shape, variability). Using hierarchical multiple regression, we evaluated whether scatter information had incremental validity over elevation in the prediction of scholastic test results. We predicted that profile elevation would be a significant predictor of achievement for all three IQ tests but that SB4 and K-ABC scatter information would have greater incremental validity than for the WISC-R.

Method

Subjects

The subjects were 146 children in English-language public schools in eastern and western Canada who were referred by their teachers to school psychologists. Their average age was 9.3 years (SD = 1.9; range = 6–12.5 years), most were boys (71%), and almost all were White (93%). A total of 75% were enrolled in regular classrooms and were referred to determine the need for special education services. The remaining 25%
were receiving learning disability services and had been referred as part of their reevaluations. (Three percent were in self-contained classrooms, 22% were mainstreamed in regular classes but were receiving "resource room" help.) Although some readers might object to inclusion of both regular and special education students in the same sample, we hope to generalize our results to referred children of normal overall cognitive ability.

On forms completed by teachers to request psychological testing, 54% of the children were reported to have poor reading skills, 32% were described as having poor spelling skills, and 23% were attributed math difficulties. These ratings covaried significantly with scores on the Wide Range Achievement Test—Revised (WRAT-R; Jastak & Wilkinson, 1984). Children described as poor readers had a mean Reading score of 80 versus 91 for all others, children reported to have math problems attained a mean Arithmetic score of 78 versus 88 for all others, and children reported to be poor spellers had somewhat lower Spelling scores (81 versus 87 for all other cases).

The average WISC-R Full Scale IQ (FSIQ) score of the sample was 95 (SD = 15). A total of 4% attained IQ scores in the mentally impaired range (< 70); 9% had borderline (70–79); 29% had low average (80–89); 42% had average (90–109); 11% had high average (110–119); and 6% had superior or higher (> 119) range scores. These children also had similar mean scores and frequencies of scores within the same IQ ranges on the K-ABC and SB4. Correlations among summary IQ scores of the different batteries were high: r(WISC-R/K-ABC) = .79, r(WISC-R/SB4) = .74, and r(K-ABC/SB4) = .72.

Measures

Of all three IQ batteries, the WISC-R is most widely known and used (Chattin & Bracken, 1989), and detailed descriptions of its characteristics and validity have been extensively reported (e.g., critiques: Bortner, 1985; Detterman, 1985; Mehrens, 1984; Witt & Gresham, 1985; see also test references in Mitchell, 1985, pp. 1710–1713). We administered all 12 WISC-R subtests.

The K-ABC (Kauffman & Kaufman, 1983a, 1983b) is normed for ages 2.5–12.5 years, and its subtests are organized into two broad scales, Mental Processing and Achievement. Mental Processing subtests, intended as measures of fluid intelligence, are further partitioned into Sequential Processing and Simultaneous Processing scales. From the Sequential Processing scale we administered Hand Movements, Number Recall, and Word Order; from the Simultaneous Processing scale we administered Gestalt Closure, Triangles, Matrix Analogies, Spatial Memory, and Photo Series. K-ABC Achievement Scale subtests were designed as measures of crystallized intelligence, and to children in this study we administered Faces and Places, Arithmetic, Riddles, Reading—Decoding (i.e., letter and word recognition), and Reading—Understanding (i.e., sentence comprehension).

We did not use results from the K-ABC Riddles and Places and Faces subtests in the data analyses. The former is conceptually more like a verbal reasoning task (e.g., the Similarities test from the WISC-R) than a traditional scholastic achievement test, and the latter is problematic for Canadian children due to U.S.-specific content (e.g., pictures of John Wayne, Robert E. Lee). The mean Faces and Places score in our sample was only 76, about a full standard deviation below some of the children's other summary scores from all three IQ tests.

The SB4 (Delaney & Hopkins, 1987; Thordike, Hagen, & Sattler, 1986a, 1986b) is normed for persons ages 2–23 years, and its 15 subtests are partitioned into four scales. Not all SB4 tasks are given to all age groups, but to children in our sample we administered Vocabulary, Comprehension, and Absurdities (Verbal Reasoning scale); Quantitative (Quantitative scale); Pattern Analysis, Copying, and Matrices (Abstract–Visual Reasoning scale); and Bead Memory, Memory for Symbols (Short–Term Memory scale).

The literature for the validity of the K-ABC is now quite substantial (e.g., Kamphaus & Reynolds, 1987); the base of validity findings for the newer SB4 is smaller but is rapidly growing. The psychometric characteristics of both batteries are generally sound (K-ABC: Kaufman & Kaufman, 1983b; Mehrens, 1984; SB4: Sattler, 1988; Thordike et al., 1986b). Both tests have been criticized, however, regarding whether their factor structures support their respective theoretical models (K-ABC: Keith, 1985, Keith & Dunbar, 1984; SB4: Keith, et al., 1988; Kline, 1989; Ownby & Carmin, 1988), and some have proposed alternative systems to derive summary scores (K-ABC: Kamphaus & Reynolds, 1987; SB4: Sattler, 1988). Because analyses for this study are based on subtest scores, the problem of how best to calculate summary scores is not here so crucial.

The WRAT-R is a well-known and widely used academic screening measure (e.g., critique: Reinehr, 1987). The WRAT-R has three subtests, Reading (letter and word recognition), Arithmetic (mathematical arithmetic, paper-and-pencil calculations), and Spelling (paper-and-pencil), which are normed for ages 5–75 years.

Procedure and Interrater Reliability

All children were tested in their schools over 2 half-day (2–3 hr) sessions, usually conducted within the same week. The IQ and achievement measures were given in counterbalanced order either by Rex Kline, Joseph Snyder, and Maria Castellanos or by MA-level clinical psychologists trained and supervised by Rex Kline. Proportions of test administrations conducted by various examiners are as follows. WISC-Rs: Rex Kline (28%), Joseph Snyder (21%), Maria Castellanos (18%), Rex Kline-supervised (32%); K-ABCs: Rex Kline (73%), Rex Kline-supervised (32%); SB4s: Rex Kline (74%), Rex Kline-supervised (26%); WRAT-Rs: Rex Kline (65%), Joseph Snyder (1%), Rex Kline-supervised (34%). As a check for scoring errors, Rex Kline reviewed the protocols for accuracy before scores were entered into the database. To evaluate possible examiner effects, we compared mean IQ and composite scores from the WISC-R, K-ABC, SB4, and WRAT-R across all examiners. All these comparisons were nonsignificant, indicating that summary scores did not differ systematically by examiner.

We also conducted a reliability study of our scoring of the IQ tests. We randomly selected 15 records from the entire 146 protocols (10% of the sample) that included tests administrations by all examiners. We removed the original scores, and all protocols were independently re-scored by Rex Kline, Joseph Snyder, and Maria Castellanos. We then derived intraclass correlation coefficients for all summary scores across the three scorers. Values of these coefficients are as follows: WISC-R: Verbal, .99; Performance, .99; Full Scale, .99; SB4: Verbal, .99, Abstract–Visual, .91; Quantitative, 1.00; Short–Term Memory, .99; Composite, 1.00; K-ABC: Sequential, .99; Simultaneous, .99; Mental Processing Composite, .99; Achievement Composite, .99. These uniformly high coefficients indicate good interrater agreement about test scoring.

Results

Representation of Profile Information

Ability profiles convey three types of information: elevation, variability, and shape. Elevation is the mean score across all subtests, variability is the variance of subtest scores around the mean, and shape is the particular pattern of high and low subtest scores in the profile. For each child, we calculated numerical indexes of elevation, variability, and shape for each IQ scale. For example, elevation information from the WISC-R was rep-
resented by each child's mean score across all 12 subtests. For the K-ABC, we used the mean of the eight Mental Processing subtests; for the SB4, the mean of its 11 subtest scores. We used the profile variability index (PVI; McLean, Reynolds, & Kaufman, 1990; Plake, Reynolds, & Gutkin, 1981) to represent variability. The PVI is the variance of subtest scores around a child's mean subtest score, and for the WISC-R it is derived as follows:

$$ WISC-R \text{ PVI} = \frac{\sum_{i=1}^{k} (X_i - M)^2}{K - 1} $$

where $M_i$ is the mean score of child $i$ across all WISC-R subtests, $X_i$ represents the subtest scores of child $i$, and $K$ is the number of subtests on the WISC-R (12). We calculated similar indexes for the K-ABC and the SB4.

We used the technique of cluster analysis to represent profile shape information. Briefly, cluster analysis identifies "prototypic" profile groups in a sample and classifies individual profiles into groups they most closely resemble. We conducted for each IQ scale a hierarchical cluster analysis using cosines as the similarity index (with the cluster program of SPSS Release 4.0). Hierarchical algorithms yield a series of solutions, $c_0, c_1, c_2, \ldots$, where $c_0$ represents the initial solution in which all $N$ profiles are distinct, separate groups, and $c_i$ is the final solution where all profiles are classified together into one large group. Selection of an intermediate solution depends on distinctiveness of group mean profiles and cluster size at each step (Blashfield, 1981).

Before we present results of the cluster analyses for each IQ test, readers should understand two important limitations. First, because cluster analysis can construct groups using random data, it is important to consider whether a particular solution would replicate across another sample (Blashfield, 1981). Our overall sample size is not large enough to split and conduct separate cluster analyses. Thus, we cannot convincingly demonstrate here that the particular WISC-R, K-ABC, and SB4 subtypes we identify are generalizable outside of our sample. Fortunately, other researchers have used similar techniques with the WISC-R, and at least we can compare the WISC-R profile types we find in our sample to those described by others. Second, our sample size will limit the number of distinct profile types we can reasonably find. That is, additional profile shapes may be discovered in a larger sample.

**Cluster Analyses**

Results of cluster analyses for all three IQ batteries indicated three-group solutions. Reported in Figure 1 are the mean profiles of clusters for all IQ tests. Although the metric of SB4 subtests scores is $\mu = 50$ and $SD = 8$, in Figure 1 mean SB4 scores are reported in the same metric as for WISC-R and K-ABC subtests, $\mu = 10$, $SD = 3$. Also, cluster sizes do not always sum to $N = 146$ because of missing data.

Type 1 WISC-R profiles were relatively flat; Type 2 profiles had lower scores on Verbal subtests than on Performance subtests, suggestive of poor language skills; and Type 3 profiles had low Block Design and Object Assembly scores, which may indicate poor visual–spatial skills. Other researchers who have used
statistical classification techniques with WISC-R profiles have described groups with mean profiles similar to Types 2 and 3 in Figure 1. Within a sample of children with verbal deficits, Group 3 of Richman and Lindgren (1980) is similar to our Type 2 (low Verbal scale scores). Among referred elementary school children, Profile 4 of Hale and Saxe (1983) is similar to our Type 3 (low Block Design, Object Assembly).

For the K-ABC, Type 1 profiles had relatively high Number Recall scores but low scores on Hand Movements; Type 2 profiles were relatively flat, except for marginally lower Hand Movements scores; and Type 3 profiles showed much variability, with somewhat lower scores on sequential tasks than simultaneous tasks.

As with the WISC-R, one of the SB4 clusters (Type 1) had relatively low scores on verbal subtests as well as Memory for Sentences, which loads on a verbal factor for children older than 7 years in SB4 factor analyses (Sattler, 1988). Type 2 profiles were generally flat except for low scores on Copying, a paper-and-pencil drawing test; Type 3 was generally a flat profile group except for marginally lower scores on Bead Memory.

Demographic differences among the WISC-R, K-ABC, and SB4 profile groups were limited. The three WISC-R clusters did not differ significantly by child age or sex but did vary by the proportion of children who were receiving learning disability services (of any kind, resource room or self-contained classroom): 38% of children with Type 2 profiles (poor verbal skills) were receiving remedial services, as opposed to 12% of children with Type 1 (flat profiles) and 18% of children with Type 3 (poor visual–spatial skills) profiles, χ²(2, N = 137) = 10.15, p < .01. The three SB4 clusters did not differ by child sex or the proportion receiving regular versus remedial services but did differ significantly by age, F(2, 134) = 14.46, p < .01: Children with Type 2 (poor hand–eye coordination) profiles (M = 10.7 years) were marginally older than those with Type 1 (poor verbal skills; M = 8.8 years) or Type 3 (flat profiles; M = 9.9 years) profiles. The K-ABC profile types did not differ significantly by age, sex, and classroom placement.

We also tested whether children with different WISC-R, K-ABC, and SB4 profile types differed in academic ability. We conducted three separate multivariate analyses of variance (MANOVAs), one for each IQ scale, with the six scholastic achievement tests (3 WRA-R, 3 K-ABC Achievement scale) as dependent variables. The WISC-R clusters differed significantly across the achievement measures (Wilks’ Lambda = .76, F(12, 214) = 2.58, p < .01), as did the SB4 groups (Wilks’ Lambda = .78, F(12, 214) = 2.42, p < .01). In contrast, the three K-ABC profile types did not differ significantly in their achievement level (Wilks’ Lambda = .93, F(12, 226) = .70, p > .05).

Mean WRA-R and K-ABC Achievement Scale test scores and results of ANOVAs and post hoc comparisons for WISC-R and SB4 clusters are reported in Table 1. The WISC-R groups differed significantly across all six achievement measures. As expected, children with Type 2 (poor verbal skills) profiles had significantly lower test scores than children with Type 1 (flat profiles) and Type 3 (poor visual–spatial skills) profiles on the “language arts” achievement tests (WRAT-R Reading, Spelling; K-ABC Reading–Decoding, Reading–Understanding). More surprising, the Type 3 group (poor visual–spatial skills) had somewhat higher scores than did the group with flat WISC-R profiles (Type 1) on WRAT-R Reading and Spelling tests. The performances of these two groups on K-ABC Achievement scale tasks were more similar, however.

The SB4 groups differed significantly on all scholastic measures except the WRAT-R Arithmetic test. Similar to results with the WISC-R, children who had “poor verbal skills” profiles on the SB4 (Type 4) also had the lowest mean achievement test scores. The mean achievement scores of children with Type 1 SB4 profiles were always significantly lower than the mean scores for children with Type 2 (poor hand–eye coordination) profiles, but Types 1 and 3 (flat profiles) differed significantly on only the WRAT-R Reading test. Children with Type 2 (poor hand–eye coordination) profiles always had the highest achievement scores.

These findings suggest that profile shape is related to academic achievement, at least for the WISC-R and SB4. Whether information about profile shape has any unique predictive capacity (i.e., incremental validity) beyond elevation or variability information is evaluated in the next section. Different profile configurations on the K-ABC, however, were unrelated to achievement. Although information about profile variability may still be an important predictor (evaluated below), profile shape information from the K-ABC will probably have little incremental validity.

### Table 1: Mean Academic Test Scores of WISC-R and SB4 Clusters

<table>
<thead>
<tr>
<th>Tests</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>F*</th>
<th>Post hoc*</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRAT-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>87</td>
<td>78</td>
<td>94</td>
<td>12.31*</td>
<td>2 1 3</td>
</tr>
<tr>
<td>Spelling</td>
<td>86</td>
<td>79</td>
<td>95</td>
<td>11.11*</td>
<td>2 1 3</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>87</td>
<td>81</td>
<td>94</td>
<td>6.07*</td>
<td>2 1 3</td>
</tr>
<tr>
<td>K-ABC Achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading-Decoding</td>
<td>92</td>
<td>84</td>
<td>96</td>
<td>10.58*</td>
<td>2 1 3</td>
</tr>
<tr>
<td>Reading-Understanding</td>
<td>92</td>
<td>85</td>
<td>95</td>
<td>6.20*</td>
<td>2 1 3</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>96</td>
<td>86</td>
<td>92</td>
<td>6.89*</td>
<td>2 3 1</td>
</tr>
</tbody>
</table>

SB4 clusters

<table>
<thead>
<tr>
<th>Tests</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>F*</th>
<th>Post hoc*</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRAT-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>79</td>
<td>94</td>
<td>85</td>
<td>13.21*</td>
<td>1 3 2</td>
</tr>
<tr>
<td>Spelling</td>
<td>82</td>
<td>92</td>
<td>84</td>
<td>5.64*</td>
<td>1 3 2</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>83</td>
<td>90</td>
<td>88</td>
<td>2.17</td>
<td></td>
</tr>
<tr>
<td>K-ABC Achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading-Decoding</td>
<td>86</td>
<td>97</td>
<td>89</td>
<td>8.15*</td>
<td>1 3 2</td>
</tr>
<tr>
<td>Reading-Understanding</td>
<td>85</td>
<td>95</td>
<td>89</td>
<td>6.82*</td>
<td>1 3 2</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>87</td>
<td>96</td>
<td>92</td>
<td>5.04*</td>
<td>1 3 2</td>
</tr>
</tbody>
</table>


* Results of Neuman-Keuls at .05. Groups are listed from lowest mean score to highest. Discontinuous lines indicate significant pairwise comparison. *F* degrees of freedom are (2, 143) except for Reading–Understanding, which are (2, 113). Children younger than 7 years were not administered this test.

* p < .01.
Relative Usefulness of Elevation, Shape, and Variability in Predicting Achievement

Once we calculated for every child indexes of profile elevation (mean score), variability (PVI), and shape (cluster membership, represented with dummy variables), we evaluated with hierarchical multiple regression whether variability and shape have incremental validity beyond elevation in the prediction of academic achievement. For each IQ battery, we entered elevation at Step 1 as a predictor of academic achievement; at Step 2, we entered variability and shape data, and tested the increase in the multiple correlation from the prior step. A significant incremental $R^2$ would indicate that variability and shape contributed to the prediction of achievement beyond elevation alone. At Step 2 we also examined tests of significance of individual predictors (i.e., whether beta weights for elevation, variability, and shape differed significantly from 0).

Results (significant at the .01 level) of the regression analyses are presented in Table 2. For all three IQ scales, elevation was always a significant predictor of achievement at Step 1. Although this result was expected, the relations of mean subtest scores from the K-ABC to academic achievement were somewhat lower than for the WISC-R and SB4. Elevation from the K-ABC explained an average of 29% of achievement test variance (range: 21–45%); the same value for the WISC-R was 38% (28–55%); for the SB4, 43% (36–53%).

Profile variability was not a significant predictor of academic achievement: Beta weights for the PVI at Step 2 of the regressions were never significantly different from 0. This means that significant incremental $R^2$ values at Step 2 would be attributable to the incremental validity of profile shape and not variability. For the WISC-R, profile shape information had incremental validity beyond elevation in the prediction of four of six achievement measures. For these four measures, incremental $R^2$ values at Step 2 ranged from 7% (for WRAT-R Arithmetic and K-ABC Reading-Decoding) to 11% (WRAT-R Spelling). Although these increases in proportions of explained variance due to profile shape are moderate in size, these results nevertheless suggest that information about profile shape does have some unique predictive capacity for the WISC-R.

### Table 2
Hierarchical Regressions of Achievement on Profile Elevation, Variability, and Shape

<table>
<thead>
<tr>
<th>Step</th>
<th>Entered</th>
<th>WRAT-R tests</th>
<th>K-ABC tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Reading</td>
<td>Spelling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$R^2$</td>
<td>change</td>
</tr>
<tr>
<td>WISC-R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Shape, Variability</td>
<td>44</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Significant predictors at Step 2 (*$p &lt; .01$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>Variability</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>K-ABC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Shape, Variability</td>
<td>26</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>Significant predictors at Step 2 (*$p &lt; .01$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>Variability</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>SB4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>Elevation</td>
<td>48</td>
<td>—</td>
</tr>
<tr>
<td>2.</td>
<td>Shape, Variability</td>
<td>55</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Significant predictors at Step 2 (*$p &lt; .01$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Shape</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td>Variability</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

Different patterns of results for profile shape were apparent for the K-ABC and the SB4. For the SB4, only two of six incremental $R^2$ values were significant at Step 2; these significant values were 5% (K-ABC Reading-Understanding) and 7% (WRAT-R Reading). In general, profile shape had negligible incremental validity over elevation for the SB4. For the K-ABC, the only significant predictor of scholastic achievement was elevation; information about profile shape had no unique predictive power.

Discussion

The results of this study are quite clear: For all three IQ batteries, elevation was clearly the most important predictor of achievement. Profile variability had essentially no incremental validity, and the only IQ scale for which profile shape had even moderate unique predictive power was the WISC-R. Although other researchers have reported similar results for the WISC-R, we were surprised that the importance of profile shape information from the K-ABC and the SB4 was even less. For theoretical and psychometric reasons we mentioned earlier, we expected the opposite result.

We can offer some speculations as to why profile shape information from the K-ABC and SB4 was much less important than such information from the WISC-R. The only way profile shape was clearly related to achievement for the WISC-R was when poor language skills were indicated. Of the three IQ scales, the WISC-R has the most verbal subtests (a total of 5, not counting Digit Span); the SB4 has only three for children ages 2–11 years and four for older persons; the K-ABC Mental Processing scale has, of course, none. If the WISC-R taps a wider variety of verbal abilities that are related to school success, then profile shapes that indicate weak verbal skills would have greater external validity than for the other IQ tests.

In defense of the K-ABC, its Mental Processing scale was designed to minimize language requirements, in part to enhance assessment of minority and language-delayed children (Kaufman & Kaufman, 1983b). This characteristic of the K-ABC, however, has been controversial. For example, some authors have suggested that the K-ABC Mental Processing scale measures cognitive processes less complex than those measured by more traditional IQ batteries (Jensen, 1984). In any event, our results suggest that examiners who administer the K-ABC as their sole IQ measure for school assessments should supplement their batteries with tests of verbal skills (e.g., the Verbal scale of the WISC-R).

Elevation information from the SB4 was an excellent predictor of achievement, but the usefulness of profile shape was only marginally greater than for the K-ABC. Like the WISC-R, the only shape information that seemed relevant beyond elevation was the "poor verbal skills" pattern. Our results suggest that the WISC-R and SB4 could be used interchangeably as overall predictors of scholastic achievement, but examiners should be wary that profile shape information from the SB4 may be even less useful than that from the WISC-R.

The findings of this study add one more voice to the following caution cited by many researchers: For predicting academic achievement, the most useful information from IQ-type tests is the overall elevation of the child's profile. Profile shape information adds relatively little unique information, and therefore examiners should not overinterpret particular patterns of scores. The only clear exception we found was for the WISC-R: Children with "language deficit" profiles have lower achievement scores than their peers.

We have already mentioned some limitations of this study, including lack of a replication sample for cluster analyses. Another limitation is that one can use procedures other than cluster analysis to identify profile subtypes, including, for example, Kaufman's (1979) tables of interpretive hypotheses for the WISC-R. Other classification procedures may yield results such that shape and variability information are more useful than indicated by our results. The academic achievement measures we used in this study pose another limitation. Although the psychometric characteristics of WRAT-R and K-ABC Achievement scale subtests are quite good, they are, after all, only screening measures. There are other measures (e.g., KeyMath Diagnostic Arithmetic Test) that provide more in-depth assessments of scholastic abilities. Finally, researchers should determine whether profile shape and variability information from the new WISC-III has greater validity than for the WISC-R.

Considering the above limitations, our analyses should be replicated with other samples and measures. In view of the many findings in the literature about the limited usefulness of profile shape and variability information, we would be surprised, however, if others will obtain dramatically different results.

References


**Received May 2, 1991**

Revision received January 4, 1992

Accepted January 9, 1992