Learning Disabilities Practice

Learning Disabilities Research & Practice, *19*(1), 49–56

Diagnostic Utility of the Bannatyne WISC-III Pattern

Courtney B. Smith and Marley W. Watkins

The Pennsylvania State University

Abstract. *Regrouping Wechsler Intelligence Scale for Children-Third Edition (WISC-III) subtests into Bannatyne's spatial, conceptual, and sequential patterns has been thought by many to identify children with learning disabilities (LD). This study investigated the prevalence and diagnostic utility of WISC-III Bannatyne patterns by comparing 1,302 children with LD to 2,158 children in the WISC-III normative sample. Further analysis was conducted on a subsample of students with specific reading disabilities. Results indicated that the presence of the Bannatyne WISC-III pattern would not lead to decisions that are useful in differentiating children with LD from children without LD. For example, receiver operating characteristic (ROC) analysis, measured by the area under the curve (AUC), indicated that the Bannatyne WISC-III pattern exhibited low diagnostic utility (AUC = 0.54–0.55). Due to its inaccuracy, use of the Bannatyne WISC-III pattern is not recommended.*

More than 250 million standardized tests are administered to public school children each year in the United States (Salvia & Ysseldyke, 1998). Although much school-based testing is accomplished in groups, a substantial number of standardized tests are also employed in individual evaluations. For example, millions of children served in special education programs have participated in individual psychoeducational evaluations (U.S. Department of Education, 2001).

Over 50 percent of the students enrolled in special education programs are diagnosed as learning disabled (LD). Given that LD diagnoses have commonly rested on an ability-achievement discrepancy (Mercer, Jordan, Allsopp, & Mercer, 1996), individual evaluations to examine special education eligibility often include a standardized measure of intellectual functioning. Of the available individual intelligence tests, the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991) is the most frequently used (Kaufman & Lichtenberger, 2000) and has become an integral part of psychological assessment in the schools (Politano & Finch, 1996).

Given the popularity of the WISC-III, much attention has been focused on its usefulness in discriminating average and exceptional children and in detecting specific areas of cognitive strength and weakness (Saklofske, Schmidt, & Yackulic, 1984). Typically, interpretation of the WISC-III is based on a hierarchical, top-down model that first considers global IQ scores. Next, to extract more information from the WISC-III, distinct patterns or profiles of WISC-III subtest scores that are presumed to be associated with intellectual or educational disabilities are analyzed. This practice of interpreting the pattern of subtest scores attained by children on individual measures of intelligence is known as profile analysis (Sattler, 1992).

More than 75 different Wechsler subtest patterns have been identified. One of the most popular was developed by Bannatyne (1968), who recategorized the subtest scores of the Wechsler Intelligence Scale for Children (WISC; Wechsler, 1949) to identify children with learning disabilities. Bannatyne believed that it did not serve a constructive purpose to divide the WISC performance of children with reading disabilities into verbal and performance IQs. Instead, he attempted to reanalyze the scaled scores by grouping them into three logical categories: spatial, conceptual, and sequential. According to Bannatyne, subtests in the spatial category (Block Design, Object Assembly, and Picture Completion) require the ability to manipulate objects in multidimensional space without sequencing, subtests in the conceptual category (Similarities, Vocabulary, and

Requests for reprints should be sent to Marley Watkins, Pennsylvania State University, Department of Educational and School Psychology and Special Education, 125 CEDAR Building, University Park, PA 16802. Electronic inquiries may be sent to mww10@psu.edu.

Comprehension) involve the use of concepts and abstract reasoning, and subtests in the sequential category (Digit Span, Picture Arrangement, and Coding) engage the ability to remember sequences of visual or auditory stimuli. Bannatyne (1971) reported that disabled readers had their highest scores in the spatial category, intermediate scores in the conceptual category, and lowest scores in the sequential category (spatial > conceptual > sequential).

Rugel (1974) reviewed 25 studies that reported WISC subtest scores of disabled readers and identified the Bannatyne pattern across 22 of the samples for which complete recategorization of the subtest scores was possible. Although Rugel found that children with reading disabilities demonstrated a clear deficit in the sequential category, they did not perform significantly lower than the general population on the Picture Arrangement subtest. In addition, children with reading disabilities scored lower on the Arithmetic subtest, which was not a part of Bannatyne's (1968) original model.

Subsequently, Bannatyne (1974) acknowledged that the Picture Arrangement subtest was erroneously included in the sequential category and substituted the Arithmetic subtest for the Picture Arrangement subtest to modify the sequential category. Thus, the revised Bannatyne pattern included the spatial category (Block Design, Object Assembly, and Picture Completion), conceptual category (Similarities, Vocabulary, and Comprehension), and sequential category (Digit Span, Arithmetic, and Coding).

Following this revision, the Bannatyne pattern was applied to the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974). Initial investigations frequently found the spatial > conceptual > sequential pattern among children with learning and reading disabilities (Smith, Coleman, Dokecki, & Davis, 1977; Vance & Singer, 1979). Although later studies generally agreed that the pattern existed among some children with learning disabilities, it was not clear whether the pattern was useful in discriminating between children with different disabilities or those of different ethnic backgrounds (Clarizio & Bernard, 1981; D'Angiulli & Siegel, 2003; Dundon, Sewell, Manni, & Goldstein, 1986; Gutkin, 1979; Henry & Wittman, 1981; Kavale & Forness, 1984; Moore & Wilson, 1987; Zarske & Moore, 1982).

Research on the Bannatyne pattern extended to the WISC-III soon after its publication. Prifitera and Dersh (1993) compared the Bannatyne WISC-III pattern of children with LD, children with ADHD, and children without disabilities. Base rates of the Bannatyne WISC-III pattern in each sample were used to estimate the probability of LD given the presence of the pattern. The base rate for children with a LD was 33 percent, while the base rate for children with ADHD was 47 percent and for children without a disability was 14 percent. Given these relative proportions, Prifitera and Dersh suggested that the Bannatyne WISC-III pattern is useful for diagnostic purposes and recommended that "the presence of a pattern or patterns would suggest strongly that the disorder is present" (1993, p. 53).

Although Prifitera and Dersh (1993) recognized the possibility of misclassification if the Bannatyne pattern was used, the magnitude of this problem was not explicated. For example, the Bannatyne pattern correctly recognized only 33 of their sample of 99 children with LD. In contrast, it incorrectly identified 293 of the 2,158 children in the WISC-III normative sample as LD. Thus, only 33 of the 326 children marked by the WISC-III Bannatyne pattern were actually enrolled in LD programs. Conversely, two-thirds of the children with LD were missed by the Bannatyne pattern.

Most studies of the Bannatyne pattern utilized analyses that are useful in identifying group differences, but are not as informative for differential diagnosis of individuals within those groups. Typically, statistically significant differences between regular and special education groups have been interpreted as evidence of diagnostic accuracy for individuals. This illustrates reliance on classical validity methods instead of the more appropriate clinical utility approach (Wiggins, 1988). As noted by Elwood, "significance alone does *not* reflect the size of the group differences nor does it imply the test can discriminate subjects with sufficient accuracy for clinical use" (1993, p. 409). Little attention has been paid to the overlap in score distributions between regular and exceptional groups, although its importance has been known for decades (Meehl, 1973). In sum, group separation is necessary, but not sufficient, for accurate decisions about individuals.

To date, Prifitera and Dersh (1993) are the only researchers who have examined the Bannatyne WISC-III pattern in children with LD. Nevertheless, the Bannatyne pattern is still common in psychological training and practice. School psychologists report a reliance on subtest interpretations when analyzing intelligence tests (Pfeiffer, Reddy, Kletzel, Schmelzer, & Boyer, 2000; Watkins, 2000). Some assessment texts illustrate calculation of Bannatyne recategorized WISC-III scores (e.g., Cooper, 1995) or imply that the Bannatyne pattern has diagnostic significance (e.g., Aiken, 1996). Additionally, a common automated scoring program for the WISC-III computes Bannatyne scores (SAWS, 1995). Beyond its widespread use in the United States, the Bannatyne pattern has also been applied with non-English speaking populations (Alm & Kaufman, 2002; Chen, Yang, & Tang, 2002; Morad & Mahmoud, 2001).

Given that decisions based on the WISC-III Bannatyne profile may have a major impact on children, additional research regarding its validity is necessary. Consequently, the purpose of the present study was to investigate the prevalence of the Bannatyne WISC-III pattern in a large sample of children with learning disabilities. Additionally, appropriate diagnostic utility statistics were applied to determine whether the presence of the Bannatyne WISC-III pattern was useful in differentiating children with disabilities from children without disabilities.

METHOD

Instruments

The WISC-III is an individually administered measure of intellectual functioning designed to assess children from ages 6 years, 0 months to 16 years, 11 months. It has 13 individual subtests ($M = 10$, $SD = 3$), 10 standard and three supplementary, that combine to yield three composite scores: Verbal (VIQ), Performance (PIQ), and Full-Scale (FSIQ) IQs ($M = 100$, $SD = 15$). In addition, the WISC-III provides four factor-based index scores: Verbal Comprehension (VC), Perceptual Organization (PO), Freedom from Distractibility (FD), and Processing Speed (PS) ($M = 100$, $SD = 15$). Full details of the WISC-III and its standardization are presented in Wechsler (1991).

Procedure

Requests to contribute to an investigation of the WISC-III were mailed to 9,227 school psychologists throughout the United States. Practitioners invited to participate all worked in a school setting and were members of the National Association of School Psychologists. These school psychologists were asked to report anonymous data from their five most recent evaluations that resulted in special education placement under the following categories: LD, emotional disability, or mental retardation. Responses were received from 492 school psychologists from 47 states. Respondents worked in rural ($n = 153$), urban ($n = 107$), suburban ($n = 179$), and mixed ($n = 29$) school districts and ranged in years of experience from 1 year to 37 years ($M = 12.26$, $SD = 8.35$). Forty-three percent of the responding school psychologists held a master's degree, 17 percent held a doctoral degree, and 39 percent held a specialist degree.

Participants

Anonymous scores were reported on 2,356 students; however, 28 of those students could not be included in this study due to insufficient or invalid data. The 2,328 remaining students included 1,557 identified as LD, 265 identified as emotionally disturbed, 258 identified as mentally retarded, and 248 with other or multiple diagnoses. Identification of students in these special education categories was accomplished according to diagnostic criteria used in the local setting.

Students with Learning Disabilities

Scores on eight standard subtests (Block Design, Object Assembly, Picture Completion, Similarities, Vocabulary, Comprehension, Arithmetic, and Coding) and one supplementary subtest (Digit Span) were necessary for computation of the Bannatyne WISC-III pattern. Because of this prerequisite, WISC-III data from 1,302 students with a local diagnosis of LD (882 male, 420 female) who had all eight subtest scores on record were included in the study. No other selection criteria (i.e., VIQ-PIQ differences, low IQ scores, etc.) were used.

Based on local diagnostic criteria, students were classified as having a LD in reading alone ($n = 200$); math alone ($n = 142$); written expression alone ($n = 222$); reading and math ($n = 66$); reading and written expression ($n = 383$); math and written expression ($n = 78$); reading, math, and written expression ($n = 192$); and unspecified ($n = 19$). Students ranged in age from 6 to 16 years ($M = 9.83$, $SD = 2.54$) and in grade from kindergarten to Grade 11 ($Mdn = 4$). Ethnic background of the students was 74.9 percent white, 7.8 percent Hispanic, 12.7 percent African American, 1.9 percent Native American, 0.8 percent Asian/Pacific, 0.8 percent other, and 1.2 percent unspecified. Students represented 46 states and were enrolled in rural (30.5 percent), urban (20.3 percent), suburban (37.1 percent), mixed (7.6 percent), and unspecified (4.5 percent) school districts. Level of parental education, as reported by responding school psychologists, included primary education (4.8 percent); high school education (42.9 percent); some college education (16.6 percent); college education (10.6 percent); graduate education (3.7 percent); and unspecified (21.4 percent).

Specific Learning Disability in Reading

Given the imprecision of diagnosis and classification of LD in practice (Kavale, Fuchs, & Scruggs, 1994), a subsample of students with specific reading disabilities was selected from the larger group of 1,302 students with a LD. These students were identified by the following criteria: (1) diagnosed as having a LD in reading by the local multidisciplinary team; (2) exhibited a significant discrepancy between expected and obtained reading achievement based on regression analysis of FSIQ and reading test scores with a 95 percent confidence interval (Reynolds, 1984–1985); and (3) displayed nonsignificant discrepancy between expected and obtained math achievement based on regression analysis of FSIQ and math test scores. Academic achievement was measured by a total of 62 tests or combinations of tests. However, the Woodcock-Johnson Tests of Achievement and the Wechsler Individual Achievement Test were used for around 85 percent of the cases. Achievement in reading and math was first determined by averaging the reading achievement scores (i.e., basic reading skills and reading comprehension subtests) and math achievement scores (i.e., math computation and math reasoning subtests). If only one achievement score was provided in reading or math, that score was used to determine the discrepancy in that academic area. Based on these criteria, 192 students were identified as having a specific reading disability.

Students Without Disabilities

The WISC-III normative sample served as the contrast group for this investigation. It was comprised of 2,200 children, which included 100 males and 100 females in each of 11 age groups ranging from 6 years, 0 months to 16 years, 11 months. The standardization sample was stratified by age, gender, race/ethnicity, geographic location, and parent education according to the 1988 U.S. Census. Children excluded from this normative sample were those who obtained a FSIQ less than 70. Thus, the final sample of children without disabilities consisted of 2,158 children, as extracted from the report of Prifitera and Dersh (1993).

Analyses

Bannatyne Pattern

Following Bannatyne's (1974) revised model, scaled scores of the WISC-III subtests were combined to produce spatial, conceptual, and sequential scores. The presence of the Bannatyne pattern was determined when the summed scores on the spatial subtests (Block Design, Object Assembly, and Picture Completion) were greater ($\geq$1 point) than the summed scores on the conceptual subtests (Similarities, Vocabulary, and Comprehension), which, in turn, were greater ($\geq$1 point) than the summed scores on the sequential subtests (Digit Span, Arithmetic, and Coding).

Diagnostic Utility

Diagnostic utility statistics were used to examine the relationship between the Bannatyne WISC-III pattern and a diagnosis of LD. First, sensitivity and specificity indices were calculated. Sensitivity was the percentage of children with LD who were classified by the Bannatyne pattern as having a disability (Fisher & Van Belle, 1993). That is, if a child has a LD, how likely is the Bannatyne pattern? It was computed as follows: the number of true-positive results $\div$ (the number of true-positive results plus the number of false-negative results).

Directly related to sensitivity is specificity. Specificity was the percentage of children without a LD who were classified by the Bannatyne pattern as not having a disability (Fisher & Van Belle, 1993). That is, if a child does not have a LD, how unlikely is the Bannatyne pattern? It was computed as follows: the number of true-negative results $\div$ (the number of true-negative results plus the number of false-positive results). Generally, sensitivity values of at least 0.70 and specificity values of at least 0.80 are considered necessary for an accurate and economical diagnostic test (Matthey & Petrovski, 2002).

Sensitivity and specificity convey important information, but it is also useful to know what proportion of positive tests show the presence of a LD and what proportion of negative tests show the absence of a LD. Therefore, positive and negative predictive values of the Bannatyne WISC-III pattern were also calculated. The positive predictive value was the percentage of children with a positive Bannatyne WISC-III pattern who truly had a LD. That is, if a child has the Bannatyne pattern, how likely is that child to have a LD? It was computed as follows: the number of true-positive results $\div$ (the number of true-positive results plus the number of false-positive results). Similarly, the negative predictive value was the percentage of children without the Bannatyne WISC-III pattern who truly did not have a LD. That is, if a child does not have the Bannatyne pattern, how likely is that child to not have a LD? It was computed as follows: the number of true-negative results $\div$ (the number of true-negative results plus the number of false-negative results).

Although these diagnostic utility indices provide valuable information about a diagnostic test, they have several limitations. Most importantly, each of these indices is influenced by the prevalence of the disability or the cut-off value used (McFall & Treat, 1999). Thus, if prevalence rates or cut-off values vary, then sensitivity, specificity, and predictive values might also change.

In the present study, this limitation was important to consider for two reasons. First, the two groups compared in this study had relatively similar prevalence rates. That is, there were a large number of students with a LD. In actual practice, the number of children with a LD would be much smaller than those without a LD (U.S. Department of Education, 2001) and, thus, relative prevalence rates would be quite different than those suggested by the present analysis.

Second, Bannatyne (1968) did not specify cut-off values for his pattern. A child displayed the Bannatyne pattern if the spatial score was greater than the conceptual score, which, in turn, was greater than the sequential score. With only one cut-off value, it could not be determined how other cut-off values would influence sensitivity, specificity, and predictive values.

To ameliorate these limitations, receiver operating characteristic (ROC) methods were applied. The ROC is independent of prevalence rates and cut-off values (McFall & Treat, 1999). Essentially, a ROC is a graph of the percentage of true-positive decisions against the percentage of false-positive decisions across all possible cut-off values. ROC analysis is conducted through three steps: (1) calculate true-positive and false-positive rates across an entire range of cut-off scores, (2) plot the resulting pairs of true-positive and false-positive rates to form a curve, and (3) calculate the area under the curve (AUC), which provides an accuracy index of the test (Henderson, 1993).

AUC values can range from 0.5 to 1.0. An AUC value of 0.5 signifies that the true-positive rates and false-positive rates are equal across all possible cut-off scores and no discrimination exists (McFall & Treat, 1999; Swets, 1988). In this case, the ROC curve lies on the main diagonal of the graph and the diagnostic system is functioning at the level of chance. In contrast, an AUC value of 1.0 denotes perfect discrimination. AUC values of 0.5 to 0.7 indicate low test accuracy, 0.7 to 0.9 indicate moderate test accuracy, and 0.9 to 1.0 indicate high test accuracy (Swets, 1988).

In the present study, ROC analysis was applied to the Bannatyne WISC-III pattern by assigning cut-off values based on the magnitude of spatial, conceptual, and sequential score differences. Thus, the first cut-off score was one or more points difference between each Bannatyne category. That is, a one-point or greater difference between the spatial and conceptual score and a one-point or greater difference between the conceptual and sequential score. Other cut-off values were successively set from two points or greater difference, three or more points difference, and so on. At each cut-off value, true-positive and false-positive rates were calculated and plotted. The AUC values were then calculated in order to measure the overall accuracy of the Bannatyne WISC-III pattern for discriminating between children with and without LD.

This ROC analysis was supplemented with an AUC calculation for a binary diagnostic test that "evaluates the discriminative ability of a test in its simplest dichotomous version" (Cantor & Kattan, 2000, p. 469). That is, a ROC based only on the dichotomous Bannatyne pattern (present vs. absent) described by Bannatyne (1974). Although only testing the effect of one cut score, positive results should obtain if the Bannatyne pattern is effectively discriminating between groups of students.

RESULTS

Descriptive statistics for intelligence and achievement test scores, including Bannatyne recategorizations, for each group of participants are provided in Table 1. As anticipated, the sample of students with reading disabilities was marked by average scores on intelligence and math tests and very low scores on reading tests. In contrast, the group with learning disabilities exhibited average scores on the WISC-III, but moderately low reading and math test scores. When group averages were considered, students with disabilities showed the Bannatyne pattern of spatial > conceptual > sequential. However, when considered individually,

TABLE 1
Descriptive Statistics for IQ and Achievement Tests for Students with Learning (LD) and Reading (RD) Disabilities and Students Without Disabilities (Norm)

	LD Group		*RD Group*		*Norm Group*	
	M	*SD*	*M*	*SD*	*M*	*SD*
FSIQ	94.7	12.4	96.4	11.8	100.8	14.1
Reading	84.5	12.0	77.4	8.3	—	—
Math	89.9	12.2	94.0	10.9	—	—
Spatial	28.5	7.4	29.4	6.7	30.2	7.9
Conceptual	27.5	7.6	27.9	7.4	30.2	8.0
Sequential	25.1	5.6	26.1	5.1	30.3	6.7

TABLE 2
Diagnostic Utility Statistics for the Bannatyne WISC-III Pattern

			Predictive Value	
Group	*Sensitivity*	*Specificity*	*Positive*	*Negative*
Learning disability	0.224	0.861	0.493	0.648
Reading disability	0.240	0.861	0.133	0.927

only 291 of the 1,302 children with LD (22.4 percent) and 46 of the 192 children with specific reading disabilities (24 percent) displayed the Bannatyne WISC-III pattern. Of the 2,158 children in the WISC-III normative sample, 299 (13.9 percent) displayed the Bannatyne WISC-III pattern.

Sensitivity, specificity, positive predictive, and negative predictive indices are reported in Table 2. Sensitivity of the Bannatyne WISC-III pattern was low for both groups of students with disabilities, while specificity was good. Generally, sensitivity values of at least 0.70 and specificity values of at least 0.80 are considered worthwhile (Matthey & Petrovski, 2002). In comparing children with LD to the children in the WISC-III normative sample, the positive and negative predictive values were generally poor (Cicchetti, 2001). As expected, the low prevalence of specific reading disabilities resulted in an increase in negative predictive power and decrease in positive predictive power (Meehl & Rosen, 1955).

ROC analysis comparing true-positive and false-positive rates of children with learning disabilities and children in the WISC-III standardization sample resulted in an AUC of 0.543 (see Figure 1). The AUC for children with specific reading disabilities compared to children in the WISC-III normative sample was 0.553

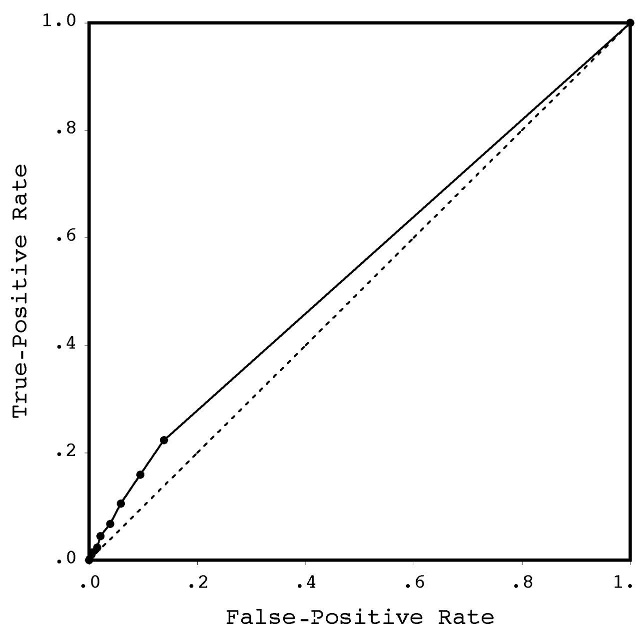

FIGURE 1 ROC curve comparing true-positive and false-positive rates of children with learning disabilities ($n = 1{,}302$) and children in the WISC-III standardization sample ($n = 2{,}158$).

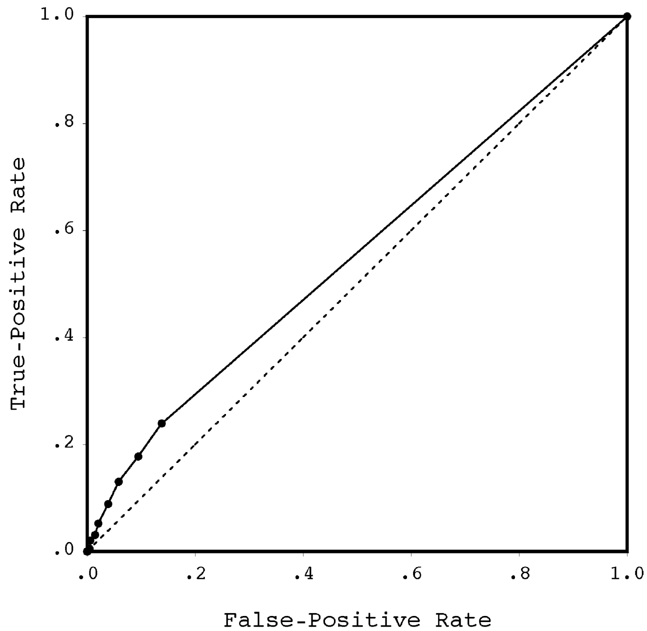

FIGURE 2 ROC curve comparing true-positive and false-positive rates of children with reading disabilities ($n = 192$) and children in the WISC-III normative sample ($n = 2{,}158$).

TABLE 3
Binary Area Under Curve (AUC) and Proportion of Students with Learning (LD) and Reading (RD) Disabilities and Students Without Disabilities (Norm) Showing Bannatyne Pattern Based on Magnitude of Spatial, Conceptual, and Sequential Differences

	LD Group		*RD Group*		*Norm Group*
Magnitude	%	*AUC*	%	*AUC*	%
1	22.4	0.543	24.0	0.551	13.6
2	15.9	0.532	17.7	0.541	9.3
3	10.5	0.523	13.0	0.536	5.7
4	6.8	0.514	8.9	0.525	3.9
5	4.5	0.512	5.2	0.516	2.0
6	2.4	0.505	3.1	0.508	1.5
7	1.5	0.504	2.1	0.507	0.6
8	0.6	0.502	0.5	0.502	0.2
9	0.3	0.501	0		0.2
10	0.2	0.501	0		0

(see Figure 2). The proportion of students displaying the Bannatyne pattern based on the increasing magnitude of spatial, conceptual, and sequential score differences are found in Table 3. Considering the Bannatyne pattern as a binary diagnostic test produced AUCs (see Table 3) consistent with those generated by varying cut-off values based on the severity of the differences between Bannatyne components.

DISCUSSION

This study investigated the prevalence and diagnostic utility of the Bannatyne WISC-III pattern in children with LD. Similar to previous research on the Bannatyne pattern (Clarizio & Bernard, 1981; Dundon et al., 1986; Gutkin, 1979; Henry & Wittman, 1981; Kavale & Forness, 1984; Moore & Wilson, 1987; Prifitera & Dersh, 1993; Smith et al., 1977; Vance & Singer, 1979), the Bannatyne WISC-III pattern was found in 22–24 percent of children with LD. However, it missed 76–78 percent of the children previously diagnosed with a LD and incorrectly identified around 14 percent of the WISC-III normative sample as LD.

Removing the confounding influence of cut scores and prevalence rates via a ROC analysis allowed a clearer demonstration of the discriminative power of the Bannatyne pattern. From this perspective, the probability that the Bannatyne WISC-III pattern accurately discriminated children with learning disabilities from children without learning disabilities was near chance. That is, if one child was selected at random from the sample of children with learning disabilities and another chosen randomly from the WISC-III normative sample, the probability of the Bannatyne WISC-III pattern correctly identifying the student with LD was 0.54 to 0.55. According to Swets (1988), AUC values of this magnitude indicate low test accuracy.

Prifitera and Dersh suggested that the Bannatyne WISC-III pattern is useful for diagnostic purposes and recommended that "the presence of a pattern or patterns would suggest strongly that the disorder is present" (1993, p. 53). The current results do not support their conclusion. Although the Bannatyne WISC-III pattern was present in around 22 percent of the children with LD, it was also found in 14 percent of the normative sample. Taken together, there were 590 children with the Bannatyne WISC-III pattern, but 291 were from the LD sample and 299 from the normative sample. Thus, presence of the Bannatyne WISC-III pattern was *not* a strong indicator of LD. Therefore, this study, like previous research on the Bannatyne WISC-R pattern, did not support the validity of the Bannatyne WISC-III pattern for determining the presence of a LD (Clarizio & Bernard, 1981; D'Angiulli & Siegel, 2003; Dundon et al., 1986; Gutkin, 1979; Henry & Wittman, 1981; Kavale & Forness, 1984; Moore & Wilson, 1987; Zarske & Moore, 1982). Professionals who use this pattern to make important decisions about the educational planning and placement of children would therefore be "acting in opposition to the scientific evidence" (Kamphaus, 1998, p. 41).

These results also graphically illustrate that reliance on group statistics to infer individual discrimination can be misleading (Elwood, 1993). In the current study, the group of students with LD displayed the Bannatyne pattern at a higher rate than did the children without disability, but extensive overlap of score distributions made it impossible to accurately identify individuals with and without LD. Similar group versus individual distinctions have been found in research with other IQ subtest profiles (Watkins, 1996; Watkins, Kush, & Glutting, 1997a, 1997b; Watkins, Kush, & Schaefer, 2002). Consequently, research on profiles and recategorizations should always include diagnostic utility methods in addition to classical validity methods.

ACKNOWLEDGMENTS

This article is based on the Master's thesis of the first author. This study was made possible by a research award from the Society for the Study of School Psychology.

REFERENCES

Aiken, L. R. (1996). *Assessment of intellectual functioning* (2nd ed.). New York: Plenum.

Alm, J., & Kaufman, A. S. (2002). The Swedish WAIS-R factor structure and cognitive profiles for adults with dyslexia. *Journal of Learning Disabilities, 35,* 321–333.

Bannatyne, A. (1968). Diagnosing learning disabilities and writing remedial prescriptions. *Journal of Learning Disabilities, 1,* 242–249.

Bannatyne, A. (1971). *Language, reading, and learning disabilities: Psychology, neuropsychology, diagnosis and remediation*. Springfield, IL: Charles C. Thomas Publisher.

Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC scale scores. *Journal of Learning Disabilities, 7,* 272–273.

Cantor, S. B., & Kattan, M. W. (2000). Determining the area under the ROC curve for a binary diagnostic test. *Medical Decision Making, 20,* 468–470.

Chen, H., Yang, Z., & Tang, X. (2002). Cognitive function of Chinese children with reading disorder: II: Subtypes of reading disorders in Chinese children. *Chinese Mental Health Journal, 16,* 52–54.

Cicchetti, D. V. (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology, 23,* 695–700.

Clarizio, H., & Bernard, R. (1981). Recategorized WISC-R scores of learning disabled children and differential diagnosis. *Psychology in the Schools, 18,* 5–12.

Cooper, S. (1995). *The clinical use and interpretation of the Wechsler Intelligence Scale for Children-Third Edition*. Springfield, IL: Charles C. Thomas.

D'Angiulli, A., & Siegel, L. S. (2003). Cognitive functioning as measured by the WISC-R: Do children with learning disabilities have distinctive patterns of performance? *Journal of Learning Disabilities, 36,* 48–58.

Dundon, W. D., Sewell, T. E., Manni, J. L., & Goldstein, D. (1986). The Bannatyne recategorization assessment procedure: Is it valid for individual diagnosis of LD children? *Learning Disability Quarterly, 9,* 208–213.

Elwood, R. W. (1993). Psychological tests and clinical discriminations: Beginning to address the base rate problem. *Clinical Psychology Review, 13,* 409–419.

Fisher, L. D., & Van Belle, G. (1993). *Biostatics: A methodology for the health sciences*. New York: John Wiley & Sons, Inc.

Gutkin, T. B. (1979). Bannatyne patterns of Caucasian and Mexican-American learning disabled children. *Psychology in the Schools, 16,* 178–183.

Henderson, A. R. (1993). Assessing test accuracy and its clinical consequences: A primer for receiver operating characteristic curve analysis. *Annals of Clinical Biochemistry, 30,* 521–539.

Henry, S. A., & Wittman, R. D. (1981). Diagnostic implications of Bannatyne's recategorized WISC-R scores for identifying learning disabled children. *Journal of Learning Disabilities, 14,* 517–520.

Kamphaus, R. W. (1998). Intelligence test interpretation: Acting in the absence of evidence. In A. Prifitera & D. H. Saklofske (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 40–57). New York: Academic Press.

Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Wechsler scale profiles and recategorizations: Patterns or parodies? *Learning Disability Quarterly, 7,* 137–156.

Kavale, K. A., Fuchs, D., & Scruggs, T. E. (1994). Setting the record straight on learning disability and low achievement: Implications for policymaking. *Learning Disabilities Research & Practice, 9,* 70–77.

Kaufman, A. S., & Lichtenberger, E. O. (2000). *Essentials of the WISC-III and WPPSI-R assessment*. New York: John Wiley & Sons, Inc.

Matthey, S., & Petrovski, P. (2002). The Children's Depression Inventory: Error in cutoff scores for screening purposes. *Psychological Assessment, 14,* 146–149.

McFall, R. M., & Treat, T. A. (1999). Quantifying the information value of clinical assessments with signal detection theory. *Annual Review of Psychology, 50,* 215–241.

Meehl, P. E. (1973). Why I do not attend case conferences. In P. E. Meehl, *Psychodiagnosis: Selected papers* (pp. 225–302). Minneapolis, MN: University of Minnesota Press.

Meehl, P. E., & Rosen, A. (1955). Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychological Bulletin, 52,* 194–216.

Mercer, C. D., Jordan, L., Allsopp, D. H., & Mercer, A. R. (1996). Learning disabilities definitions and criteria used by state education departments. *Learning Disability Quarterly, 19,* 217–232.

Moore, D. W., & Wilson, B. J. (1987). On the search for characteristic WISC-R subtest profile of reading/learning disabled children. *Reading Research and Instruction, 26,* 133–140.

Morad, S., & Mahmoud, A. (2001). Cognitive levels and Bannatyne categorizations of IQ subscores among special needs children. *Journal of the Social Sciences, 29*(4), 63–86.

Pfeiffer, S. I., Reddy, L. A., Kletzel, J. E., Schmelzer, E. R., & Boyer, L. M. (2000). The practitioner's view of IQ testing and profile analysis. *School Psychology Quarterly, 15,* 376–385.

Politano, P. M., & Finch Jr., A. J. (1996). The Wechsler intelligence scale for children (3rd ed.) (WISC-III). In C. S. Newmark (Ed.), *Major psychological assessment instruments* (2nd ed., pp. 294–319). Boston, MA: Allyn and Bacon.

Prifitera, A., & Dersh, J. (1993). Base rates of WISC-III diagnostic subtest patterns among normal, learning disabled, and ADHD samples. *Journal of Psychoeducational Assessment, ISC-III Monograph*, 43–55.

Reynolds, C. R. (1984–1985). Critical measurement issues in learning disabilities. *Journal of Special Education, 18,* 451–476.

Rugel, R. P. (1974). WISC subtest scores of disabled readers: A review with respect to Bannatyne's recategorization. *Journal of Learning Disabilities, 7,* 48–55.

Saklofske, D. H., Schmidt, H. P J., & Yackulic, R. A. (1984). Variations in WISC-R patterns of learning disabled children. *Perceptual and Motor Skills, 59,* 415–421.

Salvia, J., & Ysseldyke, J. E. (1998). *Assessment* (7th ed.). Boston, MA: Houghton Mifflin Company.

Sattler, J. (1992). *Assessment of children: Revised and updated third edition*. San Diego, CA: Jerome M. Sattler, Publisher, Inc.

SAWS. (1995). Scoring assistant for the Wechsler scales [Computer software]. San Antonio, TX: Psychological Corporation.

Smith, M. D., Coleman, J., Dokecki, P. R., & Davis, E. E. (1977). Recategorized WISC-R scores of learning disabled children. *Journal of Learning Disabilities, 10,* 437–443.

Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science, 240,* 1285–1293.

U.S. Department of Education. (2001). *Twenty-first annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Jessup, MD: Author.

Vance, H. B., & Singer, M. G. (1979). Recategorization of the WISC-R subtest scaled scores for learning disabled children. *Journal of Learning Disabilities, 12,* 63–67.

Watkins, M. W. (1996). Diagnostic utility of the WISC-III developmental index as a predictor of learning disabilities. *Journal of Learning Disabilities, 29,* 305–312.

Watkins, M. W. (2000). Cognitive profile analysis: A shared professional myth. *School Psychology Quarterly, 15,* 465–479.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997a). Discriminant and predictive validity of the WISC-III ACID profile among children with learning disabilities. *Psychology in the Schools, 34,* 309–319.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997b). Prevalence and diagnostic utility of the WISC-III SCAD profile among children with disabilities. *School Psychology Quarterly, 12,* 235–248.

Watkins, M. W., Kush, J. C., & Schaefer, B. A. (2002). Diagnostic utility of the learning disability index. *Journal of Learning Disabilities, 35,* 98–103.

Wechsler, D. (1949). *Wechsler intelligence scale for children*. New York: Psychological Corporation.

Wechsler, D. (1974). *Wechsler intelligence scale for children* (rev. ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *Wechsler intelligence scale for children* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wiggins, J. S. (1988). *Personality and prediction: Principles of personality assessment*. Malabar, FL: Krieger Publishing Company.

Zarske, J. A., & Moore, C. L. (1982). Recategorized WISC-R scores of learning disabled Navajo Indian children. *Psychology in the Schools, 19,* 156–159.

About the Authors

Marley W. Watkins is Professor-in-Charge of Graduate Programs in School Psychology at The Pennsylvania State University. His research interests include: roles and functions of school psychologists; educational and psychological interventions; diagnostic utility and accuracy; and consultation.

Courtney B. Smith is a school psychologist for the Bucks County Intermediate Unit No. 22. She is currently completing her doctoral studies in school psychology at The Pennsylvania State University.