

*Tim Claypool*  
University of Saskatchewan

*Christopher Marusiak*  
and

*Henry L. Janzen*  
University of Alberta

## Ability and Achievement Variables in Average, Low Average, and Borderline Students and the Roles of the School Psychologist

*This study contributes to ongoing research in the field of school psychology by examining some of the effects of using the Full Scale Intelligence Quotient (FSIQ) to classify students aged 6-16 years according to their results on an individual measure of intelligence, the Wechsler Intelligence Scale for Children, Third Edition (WISC-III, 1991). Levels of achievement in word-reading and mathematics calculation were compared for 196 students classified as Average (IQ ranging from 90-109), Low Average (IQ ranging from 80-89), or Borderline (IQ ranging from 70-79). In all cases the Low Average and Borderline groups' achievement levels differed significantly from that of the Average group. The fact that reading and mathematics abilities were not differentiated when Low Average and Borderline groups were compared calls into question the veracity of these labels.*

*Cette étude contribue à la recherche en cours dans le domaine de la psychologie scolaire en étudiant les effets d'employer le quotient intellectuel à l'échelle complète pour classer les élèves de 6 à 16 ans en fonction de leurs résultats sur une mesure individuelle d'intelligence, l'échelle d'intelligence de Wechsler pour enfants, 3<sup>e</sup> édition (WISCIII). Nous avons comparé les rendements en lecture et en mathématiques de 196 élèves classés comme moyens (QI entre 90 et 109), moyens faibles (QI entre 80 et 89), ou médiocres (QI entre 70 et 79). Dans tous les cas, le rendement des élèves dans les groupes «moyens faibles» et «médiocres» divergeait de façon significative de celui du groupe «moyens». Le fait que les habiletés en lecture et en mathématiques des élèves du groupe «moyens faibles» ne se distinguaient pas de celles du groupe «médiocres» remet en question la fiabilité de ces étiquettes.*

Students are differentiated semantically and psychometrically when their levels of intelligence are determined through measures of cognitive abilities. As one prominent test, the Wechsler Intelligence Scales for Children—Third

---

Tim Claypool is an assistant professor in the Educational Psychology and Special Education Department, College of Education. His research interests are in the area of school psychology, with emphasis on issues related to psychoeducational assessment, character education, and empathy.

Christopher Marusiak is a doctoral student in counseling psychology and a registered psychologist currently in private practice in Edmonton.

Henry Janzen is a professor emeritus in the Department of Educational Psychology, Faculty of Education. He is a registered psychologist known for his clinical practice and community service as well as his teaching and research.

Edition (WISC-III, Wechsler, 1991) employs the Full Scale Intelligence Quotient (FSIQ) to summarize an individual's overall performance according to the following classification system: *Very Superior* (130 and above), *Superior* (120-129), *Above Average* (110-119), *Average* (90-109), *Low Average* (80-89), *Borderline* (70-79), and *Intellectual Deficient* (69 and below). School psychologists invariably refer to these WISC-III classifications in their psychoeducational reports, and these labels often form the basis for communication with other professionals and non-professionals alike. As well, assumptions about students' academic capabilities stem in large part from these IQ-based labels and classifications.

If a common goal of intelligence test interpretation is the "classification of individuals according to their cognitive abilities" (Flanagan, McGrew, & Ortiz, 2000, p. 14), then it makes sense that the resulting classifications communicate meaningful information. Labels such as *Low Average* may assist communication between professionals and their clients, yet it begs the question regarding what exactly is being communicated. These "80s kids" may or may not be cognitively and academically similar, and earlier research has not fully elucidated the nature of this psychometric grouping. To this end several questions need to be asked. For example, will Low Average students' WISC-III profiles serve to distinguish them significantly from their Average and Borderline peers? Will patterns of academic achievement emerge that will be distinctive from students' IQ classifications?

Individuals classified as Low Average on the Wechsler Scales are often overlooked and understudied by educators and researchers. The Low Average classification used by the WISC-III presumes a normal distribution with approximately 16% of the population falling within the 80-89 standard score boundaries. This percentage indeed constitutes a significant portion of any given student population. Similar to their Average or Borderline peers, Low Average students may achieve either within, below, or above the low average range in measures of academic achievement. These individuals may be at risk for school failure even though Special Education services are available in school systems.

School psychologists and other school personnel have a common goal to identify students' needs and match them with available programs and services. However, problems arise when students do not meet the provincial and local eligibility criteria for these specialized programs and services based on their psychoeducational assessment results. Low Average students may have been excluded from being classified as learning disabled because their IQs did not fall in a normal range (Rourke, 1998). Typically, Low Average students cannot be diagnosed as learning-disabled because a severe discrepancy between ability and achievement is one of the main criteria for categorizing a learning disability (Evans, 1990). This would translate into standard scores on measures of academic achievement in the 50-60 range. By comparing actual achievement scores in reading and mathematics with IQ classification levels in the current study, discrepancies between measures of ability and actual achievement became more obvious. These comparisons added to the descriptive picture that was formed and could lead to conclusions that might question some of the exclusionary practices in special education funding of learning-disabled students based on IQ cut-off scores in the Average range.

When looking beyond Wechsler's classification system the term *slow learner* was used to refer to those individuals who "can progress in school but cannot grasp higher abstractions and symbols and who advance one-half to three-quarters of a grade per year" (Shelton, 1971, p. 17). More recently, Swanson, Mink, and Bocian (1999) referred to slow learners as those individuals whose low reading scores align with a corresponding low IQ. Subsequently, these researchers report an average WISC-III Full Scale IQ of 73 for the 25 slow learners included in their study. In the past it was generally believed that the main cause of the "slow learner's" lack of progress in school was mainly genetic. It has, however, since been shown that the interactive effects of genetics and environment combine to determine an individual's IQ or global intellectual level (Bidell & Fischer, 1997; Sparrow & Davis, 2000).

Frequently the reporting of an individual's performance on standardized tests fails to describe adequately the individual completing these tests. When an ecological approach to learning is considered, the question always remains as to whether the specific learning difficulties of Low Average students are inherent in the individual or secondary by-products of insufficient or ineffective accommodation to their learning needs. It needs to be recognized that some of the Low Average students' academic problems are created and sustained by inflexible academic and evaluation standards (Kronick & Hargis, 1998). Along a related vein, once Low Average individuals are no longer in school they may not be easily detected in the general population (Shelton, 1971). If this is indeed the case, then the validity of classification systems based strictly on standardized test scores is questionable if not discriminatory. The efficacy of using IQ-based cut-offs for the purpose of limiting access to specialized programming and government funding could be challenged on ethical and psychometric, if not legal grounds.

Critics of the psychometric approach to the measurement of intelligence are quick to point out the damaging effects of labeling and mislabeling students (Sternberg & Grigorenko, 2001). In addition to the challenges discussed above, alternative yet equally valid ways of measuring IQ may lead to changes in a student's IQ-based classification. As one example, Kaufman (1994) has for several years advocated that Symbol Search should be substituted for Coding to calculate the Performance Scale and Full Scale IQs. Symbol Search's higher correlations with these scales, combined with the fact that it is mainly a measure of mental processing as compared with Coding's measure of fine motor skills, results in a strong case for this routine substitution. If IQ varies according to how it is measured, then the resulting IQ-based labels may convey little meaningful information.

What are the academic implications when a Low Average student achieves within predicted levels on a standardized measure of achievement? Using the Low Average student's WISC-III FSIQ, his or her predicted performance on standardized measures of academic achievement would be expected to fall within a corresponding low average range. Typically, this student is experiencing academic problems in school. Clearly the Low Average student has demonstrated an ability to learn, but at a different rate or pace than the Average student. It is generally accepted that many Above Average students are capable of progressing quickly through their assignments and courses

when given the opportunity. A similar flexibility needs to be applied to the Low Average student who may require more time to progress through the same curricula. When students fall into the Low Average classification, they may require some form of assistance to ensure continued success, yet not qualify for a learning disability designation and associated accommodations. The school psychologist is left with the task of recommending appropriate programming for the Low Average student even when special-needs funding might not be available to support its implementation.

#### *The Role of the School Psychologist*

The common ground for all stakeholders in the psychoeducational assessment process is an underlying concern for the welfare of the individuals being assessed. When students present with a constellation of complex educational needs, it is incumbent on school psychologists to ensure that all levels of their investigation respond appropriately to this complexity. School psychologists are responsible for translating assessment results into meaningful reports that communicate both psychometric and educationally relevant information and using these results to build a solid set of recommendations that can be enacted by parents and professionals. WISC-III results will assist in coding students as developmentally delayed, learning-disabled, or gifted, but the benefit that may come with the knowledge that a student is Low Average is less obvious. Restrictive government funding for financing special education programming seems to demand the classification of students. School personnel scramble to meet these ever-changing regulations and coding guidelines in the hope of better serving those students who have been judged to have more intense learning and behavioral needs.

The call for transparency and accountability in a school psychologist's practice has been traced to the increasingly litigious and political climate surrounding education and special needs issues (Woolfson, Whaling, Stewart, & Monsen, 2003). As a result, there has been a move toward evidence-based practice (Fox, 2003) as well as the need to look more closely at the political milieu in which school psychologists typically operate. Various levels of government may have a hand in drafting informed legislation surrounding the allocations of limited financial resources for targeted special education populations. The school psychologist as technician is non-political and merely applies his or her psychometric tools and report writing skills to the student eligibility task at hand. Yet at the same time the school psychologist could be viewed as a gatekeeper to valuable resources (Dennis, 2004) when the results of their assessment could translate into government funding for students with special needs.

The call for school psychology services to become more comprehensive in nature may result in a shift in role emphasis from a special education assessor to a problem-solver and change agent (Peterson, 2001). It is difficult to imagine the role of the school psychologist changing significantly unless there is a shift in mindset from one of technician to that of a social broker or interventionist. To orchestrate change, school psychologists cannot operate in a vacuum created by theory and philosophical allegiances. Instead they must be prepared to work directly with all stakeholders from a variety of disciplines, perspectives, and agendas. If the school psychologist's role is reshaped, individual

assessment and consultation may be deemphasized in favor of the development and evaluation of programs designed to meet the general learning and mental health needs of all students (Braden, DiMarino-Linnen, & Good, 2001).

In the absence of provincial or national directives, the local political climate or context in which change is expected to occur would be jurisdiction-dependent. As an example, Edmonton Public Schools have determined that one of the eligibility criteria for their learning disability programs is that students have a minimum IQ of 100. Historically, researchers and practitioners have questioned the assumption that students with learning disabilities must have an IQ within the average range and recognized the political and budgetary considerations that might exclude Low Average students from this designation. Bryan and Bryan (1986) challenged both educators and researchers to provide empirical proof of the benefits of such practices by clearly demonstrating that, "knowing that a child received a low IQ score leads to school programs that are better suited to the skills of the child or simply discourages child and teacher alike" (p. 274). The utility of a learning disability label diminishes when evaluation is based on edometrics (Bryan & Bryan) rather than standardized scores on tests of intelligence. The ability of individuals to read and understand curriculum materials and meet teacher-determined standards of performance is not only more relevant to the student in question, but probably a more accurate predictor of his or her success or failure in school than IQ scores or vaguely worded DSM-IV definitions of specific learning disabilities.

#### *Purpose of the Study*

It is widely accepted that inter-individual differences exist in WISC-III classifications (Ward, Ward, Glutting, & Hatt, 1999), but the nature and extent of these differences requires further clarification. The main goal of this study is to determine the nature of students classified as Low Average by a standardized measure of cognitive ability and academic aptitude, the WISC-III. This study explored the psychometric implications of using IQ-based classifications to differentiate students and clarify the patterns of cognitive and academic performance both within and between three separate IQ classifications. When we combined descriptive statistics and academic achievement results in reading and mathematics with relevant personal and demographic data, a detailed picture of Borderline, Low Average, and Average students emerged. The efficacy and utility of using WISC-III standard scores as a basis of classifying Low Average students may be questioned if it cannot be shown that excessive differences in cognitive and achievement profiles exist in classifications. IQ classifications derived from WISC-III results may not accurately reflect Low Average students' academic abilities or intellectual potential. As well, if alternative but accepted methods of determining IQ result in changes of IQ-based classifications, then the validity of these classifications may be called further into question.

It is intended that this study will provide further insight into the extrapolation of meaning from IQ-based classifications and their utility in grouping children for instructional purposes. By investigating individual cognitive and achievement profiles, we highlighted the heterogeneity both within and between the Average, Low Average, and Borderline classifications. Results from this study could lead to increased accountability in the reporting of test results

through a deemphasis on classifications or categorizations of individuals based solely on their psychoeducational assessment results. This in turn could influence the types of recommendations that school psychologists may generate from their psychometric-based tests of intelligence. In addition, this article questions the government-sanctioned use of IQ cut-off scores to allocate special needs funding and provide further support for the *Rights without Labels* movement. As well, IQ classifications carry a host of stated and unstated assumptions. The present study strives to improve the transparency of some of the more obvious assumptions that apply to IQ-based classifications, in the hope that school psychologists might use results from intelligence tests more effectively and ethically.

#### *Method*

Archival data were collected and analyzed from the files of students assessed at a counselling and assessment center in a major Canadian city. The tests were administered and scored according to standardized procedures by graduate students enrolled in an educational psychology program under the supervision of a registered psychologist. Canadian norms were used to calculate all WISC-III scores. Average reliability coefficients for the Canadian norms as reported in the WISC-III Canadian Supplement Manual are: FSIQ=.95; Verbal Scale=.93; Performance Scale=.89; Verbal Comprehension Index=.93; Perceptual Organization Index=.88; Freedom from Distractibility Index=.85; and Processing Speed Index=.86.

#### *Participants*

The main criteria for selecting cases from these archives were: (a) a completed WISC-III with the FSIQ falling into one of three targeted classifications (Average, 90-109; Low Average, 80-89; Borderline, 70-79); and (b) a standardized measure of academic achievement in the areas of word recognition and mathematics calculation. In addition, cases with statistically significant differences between Verbal Intelligence Quotient (VIQ) and Performance Intelligence Quotient (PIQ) were excluded from the sample, as the wide variability in cognitive strengths and weaknesses evident in these types of WISC-III profiles make the resulting FSIQ difficult to interpret (Sattler, 2001). To ensure that the number of cases included in the Average, Low Average, and Borderline groupings remained relatively equal, files from each classification were randomly selected until the total for that group represented approximately one third of the entire sample. In addition, cases were selected in each IQ classification to provide a balance between male and female students.

Following a thorough review of archived clinic files from 1994 to 2003, a total of 196 cases were found to meet the specified criteria. Although the actual numbers in each classification varied (Average=69; Low Average=68; Borderline=59), these differences might be considered less significant when their percentage relative to the entire sample population is considered. Of the 196 participants, 88 (45%) were female and 108 (55%) were male. The distribution of boys and girls in each of the IQ classifications was relatively equal. The average age of the sample was 10 years 9 months ( $SD=2.67$ ), with ages distributed relatively equally across Average ( $M=10:1$ ,  $SD=2.77$ ), Low Average

( $M=11:1$ ,  $SD=2.69$ ), and Borderline ( $M=11:3$ ,  $SD=2.4$ ) groups. The average grade of participants for all classifications was within the grade 5 level.

Most of the students included in the study had some form of learning, behavioral, or emotional concerns (see Table 1). Probably in many cases the reason for referral to the assessment clinic was to investigate the possible presence of these problems, and therefore this finding may not be representative of Borderline, Low Average, and Average students in the general population. As well, many participants were enrolled in special education programs or repeating a grade. An expected trend emerged where individuals with lower FSIQ scores tended to repeat a grade more frequently than those with higher FSIQ scores (see Table 1). Also as expected, special programming was more predominant in the Borderline and Low Average groups, and more of these students had some formalized individual education plan in place compared with the Average group. It is acknowledged that there was little if any consistency across the sample regarding the terms used to describe the special education programs. For the purpose of the current study, modified and life skills programs, as well as Individual Program Plans and the use of a resource room were all included under the label *modified program*. Special needs assistants, bilingual, and ESL programs were included as separate labels.

Table 1  
Comparison of WISC-III FSIQ Classification and Comorbid Attentional, Behavioral, and Emotional Conditions, and Special Educational Programming

	Borderline (n=59)		Low Average (n=68)		Average (n=69)	
	f	%	f	%	f	%
<i>Comorbid Diagnoses</i>						
None	14	24.14	12	17.65	17	24.64
MDD	6	10.34	0	0.00	0	0.00
LD	2	3.45	1	1.47	2	2.90
ADHD	7	12.07	5	7.35	9	13.04
ODD	1	1.72	0	0.00	0	0.00
Depression/Anxiety	5	8.62	5	7.35	3	4.35
Attention/Behavior problems	18	31.03	31	45.59	30	43.48
Multiple Diagnoses	2	3.45	6	8.82	5	7.25
Other	1	1.72	3	4.41	2	2.90
<i>Special Programs</i>						
No Special Program	17	29.31	22	32.35	50	72.46
Modified Program	32	55.17	30	44.12	11	15.94
Bilingual Program	2	3.45	4	5.88	6	8.70
Special Needs Assistant	1	1.72	4	5.88	0	0.00
ESL Program	2	3.45	0	0.00	0	0.00
Grade Retention	18	31.03	14	20.59	7	10.14

Note. MDD=Mild Developmental Delay; LD=Learning Disability; ADHD=Attention Deficit/Hyperactivity Disorder; ODD=Oppositional Defiant Disorder.

Table 2  
Summary of WISC-III Results for IQ-Based Classifications

	Borderline (n=59)		Low Average (n=68)		Average (n=69)	
	M	SD	M	SD	M	SD
<i>WISC-III Factor Scores</i>						
FSIQ	75.19	3.16	84.59	2.58	98.17	5.55
VIQ	77.69	4.05	85.13	3.26	98.48	5.81
PIQ	77.81	4.95	87.62	4.87	99.09	6.46
VCI	78.19	4.88	86.35	4.87	98.61	6.16
POI	79.54	5.91	89.63	6.30	100.04	7.60
FDI	81.19	9.59	84.04	11.03	96.71	11.86
PSI	84.76	12.82	91.06	14.76	101.00	11.21
<i>Subtest Scores</i>						
Information	6.15	2.24	7.43	1.77	9.84	2.17
Similarities	6.80	2.05	8.32	1.78	10.45	1.97
Arithmetic	5.98	1.97	6.78	2.19	9.68	2.53
Vocabulary	5.85	2.07	7.75	1.72	9.93	1.85
Comprehension	5.76	2.42	7.25	2.13	9.01	2.13
Digit Span	7.64	2.18	7.85	2.46	9.01	2.69
Picture Completion	7.53	2.10	9.01	2.17	10.23	2.21
Coding	6.31	2.56	7.54	3.24	9.36	2.77
Picture Arrangement	6.78	2.74	8.84	2.69	10.28	2.56
Block Design	5.64	2.48	7.60	2.35	9.65	2.69
Object Assembly	6.47	2.30	7.97	2.11	9.77	2.33
Symbol Search	7.83	3.11	8.94	3.16	10.90	2.65

Note. FSIQ = Full Scale Intelligence Quotient, VIQ=Verbal Intelligence Quotient, PIQ=Performance Intelligence Quotient, VCI=Verbal Comprehension Index; POI=Perceptual Organization Index; FDI=Freedom from Distractibility; PSI=Processing Speed Index.

## Results

### Descriptive Results

*WISC-III.* Table 2 gives an overview of WISC-III results and demonstrates that expected trends were observed in relation to mean FSIQ, VIQ, and PIQ for each of the three IQ-based classifications. Similarly, the Verbal Comprehension Index (VCI) and Perceptual Organization Index (POI) results did not show significant variability in results with means scores falling in expected ranges for each classification. One slight variation did occur with the Freedom from Distractibility Index (FDI) for the Borderline group where mean performance fell in the Low Average range. Possibly a more significant result occurred with the Processing Speed Index (PSI) for the Low Average classification where the mean score fell within the Average range. Here the standard deviation exceeded that of the entire sample and indicated significantly more variability in individual performance in the Low Average group than either the Average or Borderline groups for the PSI or other WISC-III factors.

Sattler (2001) recommends that scaled scores ( $M=10$ ,  $SD=3$ ) for individual WISC-III subtests that range between 1 and 7 be classified as a weaknesses or below average, whereas values from 8 to 12 are described as average. The mean scaled score for each of the 12 subtests fell in the average range for the Average

Table 3  
Summary of Achievement Test Results for IQ-Based Classifications

	<i>n</i>	<i>Word Recognition</i>		<i>Math Calculation</i>	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Full Sample	196	93.68	13.15	84.77	12.80
Average	69	100.42	11.93	91.52	13.39
Low Average	68	91.29	11.97	81.75	11.33
Borderline	59	88.54	12.70	80.36	10.35

group and below average for the Borderline group as illustrated in Table 2. Results from the Low Average group were mixed, with eight subtests falling in the below average range and four subtests falling in the average range. In other words, the mean performance on WISC-III subtests remained within expected levels for the Average and Borderline groups, whereas considerably more variability in subtest results was observed in the Low Average group.

*Achievement tests.* The Wechsler Individual Achievement Test (WIAT) was by far the most predominant measure of academic achievement used, although the Canada Quiet, Wide Range Achievement Test (WRAT), and the Woodcock-Johnson-II (WJ-II) were also employed. Other standardized tests will provide information on academic achievement and therefore introduce another source of measurement error. This obvious limitation could not be avoided given the individual nature of each assessment battery included in the study. However, we attempted to balance the number of individual cases selected from each of the three IQ-based classifications on four of the five achievement tests. The Wechsler Individual Achievement Test—Second Edition (WIAT-II) was the exception, with just one Low Average student completing this achievement measure. Of the remaining four achievement tests, the WRAT was the only test where a less even distribution of cases among IQ classifications was evident. With data gathered from the WRAT constituting only 10.7% of the total sample, it was assumed that the following distribution of scores would not prove to be a significant confound in the subsequent data analysis: Borderline=2, Low Average=10, and Average=9. When the total number of cases from the three IQ-based classifications were grouped according to each of the remaining three achievement tests (WJ-II; WIAT, Canada Quiet), the difference between these classifications and the corresponding achievement measures used never exceeded six cases. For example, with most cases selected completing the WIAT ( $n=103$ ) the breakdown was as follows: Borderline=31, Low Average=37, and Average=35.

In Ackerman's (1998) review of the WIAT, criterion-related evidence was reported. The WIAT's moderate to high correlations with the WJ-R ( $r_s$  .68 to .88) and WRAT-R ( $r_s$  .69 to .87) suggests some commonality in the underlying constructs being measured by all three word-reading tasks. In addition, Ackerman points out that multitrait-multimethod comparisons provide sufficient evidence of convergent validity. Sattler (2001) indicates that content validity is supported when there is evidence that mean raw scores increase with age. This was the case with all achievement measures included in this study. Furthermore, the fact that parallel forms of statistical analysis using results from either

all five achievement measures or the WIAT only subsample yielded essentially the same results adds some support to the argument that the introduction of more than one measure of achievement was not a significant confound.

The use of only the word recognition and mathematics calculation tasks might serve to limit some validity concerns given the similar nature of these measures across the various tests included in the study. This perceived limitation might in fact provide a more accurate reflection of school psychologists' work in the field. Typically, when they are called to consult with school personnel, student test information is provided by many instruments. Adherence to standardized testing procedures helps to ensure the reliability of these test results. In turn the standard scores obtained ( $M=100$ ,  $SD=15$ ) can be used to make informed comparisons.

As Table 3 indicates, the mean reading scores fell in the Average range for the sample, and mean mathematics scores fell in the Low Average range. A similar trend was also observed when achievement tests results were compared according to the three classifications. Table 3 illustrates that reading achievement always exceeds mathematics achievement for Average, Low Average, and Borderline individuals.

Interestingly, a significant negative correlation between reading achievement and IQ was obtained for Low Average individuals. Increases in IQ in the 80-89 range were significantly associated with corresponding reductions in levels of word-reading ability ( $r=-.32$ ,  $p<.05$ ). Although this might constitute a spurious finding, it is interesting to speculate on possible explanations. The obvious implication is that as the IQ for Low Average individuals increased in the 80-89 range, their corresponding scores on word-reading tasks declined. It is reasonable to predict that increases in IQ would be accompanied by improved reading abilities. This unexpected trend in underachievement is reminiscent of more recent definitions of learning disabilities that stress similar patterns of performance (Learning Disabilities Association of Canada Definition, 2002). In the past, Low Average individuals may have been excluded from being designated as learning disabled due to insufficient discrepancies between ability and achievement. Assessment and diagnostic protocols will need to be reconfigured to align more closely with current conceptualization and definitions of the learning-disabled. In contrast, the relationship between IQ and mathematics achievement for Low Average individuals was non-significant. The opposite occurred with the Average group where a significant positive correlation occurred between IQ and reading achievement ( $r=.24$ ,  $p<.05$ ) as well as mathematics achievement ( $r=.36$ ,  $p<.01$ ). Results for the Borderline group demonstrated non-significant correlations between IQ and achievement scores in both reading and mathematics.

*Symbol search substitution.* Substituting symbol search for coding in the measurement of IQ as advocated by Kaufman (1994) resulted in substantial changes in students' classifications (see Table 4). The symbol search substitution changed 18% of the IQ classifications. Of note is that most of this variability between classifications was demonstrated in the Borderline and Low Average groups, with 29% of the Borderline group and 21% of the Low Average group being reclassified whereas only 7% of the Average group was reclassified. It can only be assumed that if new tasks are added and others are

**Table 4**  
**Comparison of IQ Classifications between Standard FSIQ and FSIQss Methods**

	<i>Standard FSIQ Classification</i>					
	<i>Borderline</i>		<i>Low Average</i>		<i>Average</i>	
	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>	<i>f</i>	<i>%</i>
<i>FSIQss Classification</i>						
Extremely Low	9	15.25	0	0.00	0	0.00
Borderline	42	71.19	3	4.41	0	0.00
Low Average	8	13.56	54	79.41	1	1.45
Average	0	0.00	11	16.18	64	92.75
High Average	0	0.00	0	0.00	4	5.80

Note. FSIQ = Full Scale Intelligence Quotient, FSIQss=Full Scale Intelligence Quotient Symbol Search Substitution.

eliminated, further changes in individual IQ-based classifications will result when the WISC-IV and SB5 are employed.

*Data Analyses*

Separate one-way ANOVAs were employed to compare the performance of Average, Low Average, and Borderline students on measures of math and reading achievement (word recognition and mathematics calculation). A two-tailed significance level of  $\alpha=.05$  was used in the analyses. IQ-based classifications served to differentiate the Average group from the Low Average and Borderline groups for both reading ( $F_{(2,182)}=14.11, p<.05$ ) and mathematics achievement ( $F_{(2,182)}=15.35, p<.05$ ). Subsequent Games-Howell post hoc analyses indicated that Average students were statistically different from both Low Average and Borderline students in measures of reading and mathematics achievement. However, Low Average students were not statistically different from Borderline students on either reading ( $p=.47$ ) or mathematics achievement ( $p=.60$ ). Performance on these achievement measures did not result in the emergence of significant differences between Low Average and Borderline individuals in the sample.

The effect of age on IQ and achievement levels was investigated with the entire sample ( $n=196$ ) divided into three age groupings: Younger (ages 6, 7, 8, 9 years)  $n=72$ ; Middle (ages 10, 11, 12)  $n=61$ ; and Older (ages 13, 14, 15, 16). With regard to age groupings and reading achievement, the null hypothesis was accepted indicating that observed differences in scores were not statistically significant. Conversely, the alternate hypothesis was accepted when only IQ classification was considered,  $F_{(2,196)}=16.25, p=.01$  and when only the results from the WIAT were considered,  $F_{(2,94)}=15.73, p=.01$ . The interaction between age and IQ grouping was not statistically significant. All forms of post hoc analysis (Tukey HSD, Scheffe, Bonferroni, Games-Howell) indicated that differences between the Average and Low Average or Average and Borderline classifications were found to be significant. However, reading achievement levels were not significantly different when Low Average and Borderline groups were compared.

Age groupings had a different effect than gender groupings when mathematics achievement levels were compared across IQ classifications. As the age of the participants increased, their corresponding mathematics abilities decreased. This was consistent across all three IQ classifications. Rejection of the null hypothesis indicated that a main effect for age groupings was significant,  $F(2,187)=9.27$ ,  $p=.01$  for the full sample as well as the WIAT results  $F(2,94)=7.61$ ,  $p=.01$ . With no significant interactions, it appeared that mathematics achievement levels did not depend on a combination of age and IQ classification. Instead, it was obvious that Average individuals' achievement levels in mathematics were consistently greater than those of their Low Average and Borderline peers with older individuals tending to score significantly lower than individuals from either the middle or younger groups.

WISC-III factor scores were also examined in an attempt to determine similarities and differences between the Average, Low Average, and Borderline groups. Separate one-way ANOVAs demonstrated significant differences between the performance of these groups on each of the four factors: VCI ( $F(2,183)=221.58$ ,  $p<.001$ ), POI ( $F(2,183)=146.49$ ,  $p<.001$ ), FDI ( $F(2,183)=35.79$ ,  $p<.001$ ), and PSI ( $F(2,183)=26.39$ ,  $p<.001$ ). Post hoc analyses indicated that the Average group significantly differed from both the Low Average and Borderline groups on each of these factors. Low Average and Borderline groups also differed on three of the four groups with the exception of FDI where Low Average ( $M=84.04$ ,  $SD=11.03$ ) and Borderline groups ( $M=81.19$ ,  $SD=9.59$ ) were not statistically different ( $p=.30$ ). Similarities between the Borderline and Low Average groups were more obvious when analysis of the mean levels of performance on the two subtests (Arithmetic and Digit Span) that comprise the FDI was completed. Mean Digit Span scores were similar for Borderline ( $M=7.64$ ) and Low Average ( $M=7.85$ ) whereas the Average group's results ( $M=9.01$ ) fell as expected into the average range. This was also the case for their arithmetic results ( $M=9.01$ ) whereas the significantly lower Borderline ( $M=5.98$ ) and Low Average ( $M=6.78$ ) groups were not well differentiated by this subtest.

Throughout these analyses, a similar trend emerged. IQ-based classifications served to differentiate the Average group from the Low Average and Borderline groups on all measures; however, post hoc analyses demonstrated no significant differences between Low Average and Borderline groups on several measures. Therefore, it might be assumed that the WISC-III FSIQ cannot be expected to separate individuals reliably into homogeneous groupings of academic aptitude unless a broader spectrum of standard scores are considered as evidenced in the Average IQ classification.

### *Discussion*

This study attempts to provide some needed clarification of what if any differences distinguish individuals classified as Average, Low Average, and Borderline according to their WISC-III FSIQ. This obtained score (Charter & Feldt, 2001) does not consider measurement error when establishing IQ classifications. As a result, a host of assumptions about the homogeneity of group membership could be made based solely on this classificatory label. Results from this study serve to highlight individual differences as well as similarities both within and between these three groups based on a sample of 196 children between the ages of 6 and 16 who have completed psychoeducational testing.

Results from the current study also demonstrated that IQ-based classifications can change simply by applying accepted but alternative methods of determining the global indicator of cognitive ability. When a student is classified as Low Average by the WISC-III Full-Scale IQ this should not automatically exclude him or her from needed services or limit academic options as the method used to calculate an aggregate WISC-III score is a significant factor in determining who meets predetermined IQ cut-offs.

The main intention of this study was to focus on a specific subpopulation of students as defined by their results on a commonly used psychometric instrument, the WISC-III. There is no inherent belief that Low Average students possess any intrinsic qualities that would serve to differentiate them from students classified by the WISC-III as Average or Borderline. Rather, it may be more accurate to hypothesize that their similarities far outweigh their differences (Glutting & McDermott, 1994; Ward et al., 1999). Further comparative studies may choose to pursue an analysis of students from various IQ classifications on any number of academic and personality dimensions to determine whether more qualitative differences exist.

#### *Limitations*

Academic achievement data for the study were limited to only two subtests gathered from five standardized measures. It could be argued that including a reading comprehension task might have revealed more information than simple word recognition. However, there is considerably less consistency in how an individual's comprehending abilities are measured among various tests of reading as compared with the more straightforward identification of isolated words. Similarly, the mathematics calculation task included in the study would produce more consistent data than various forms of reasoning or problem-solving. Therefore, it is recognized that reading and mathematics achievement levels reported in the study may be considered basic, but we hope more consistent, if not reliable, when the data obtained from the various standardized instruments is combined and compared.

In addition, the research sample comprised students referred to a psychological counseling and assessment center for a variety of academic, behavioral, and emotional concerns. It could be argued that including students with these comorbid conditions may introduce a series of confounding variables, and as such the results of this study may not accurately represent the full range of abilities of students in the Average, Low Average, and Borderline ranges. On the other hand, this sample may be more representative of a real-world setting and may reflect the types of cases seen by school psychologists given the nature of the targeted student population. No attempt was made to form a causal link between these conditions and specific IQ-based classifications other than simply reporting trends that appeared to emerge from reviewing the 196 assessment files included in the study.

#### *Rationale Revisited*

This study attempts to survey the psychometric landscape. As a result of the current analysis of a decade of psychoeducational assessments conducted at a psychological assessment center, an archival map composed of 196 individual profiles has been formed. Can this historical information influence current

practices and decision-making in school psychology? This is not a completely rhetorical question. Its answer lies in the hearts and minds of those school psychologists who continue to use today's revised and renormed tests of intelligence on a regular basis. Albeit new and improved with recent changes in the WISC-IV and Stanford-Binet-5, these psychological instruments continue to reinforce basic assumptions about the measurement of individual differences. Any IQ-based classifications that might result from these tests will be reported in subsequent reports and continue to reinforce if not legitimize educational decisions about programming and expected outcomes for the students being assessed.

At times considered a second-order purpose (Field, 2000), educational research can advance knowledge by refuting or at least calling into serious doubt earlier assumptions and practices that continue to affect students, teachers, and educational systems. As a case in point, current assessment practices employed by educational psychologists frequently use standardized measures of intelligence as a means of classification, categorizing, and coding school-aged children and adolescents. Establishing estimated IQ levels might help determine if government funding for students with special needs can be accessed. Sattler's (2001) four pillars of assessment are often reduced to one or possibly two elements, with standardized test results or IQ often carrying the most weight or political punch. When cast into the same category as phrenology and physiognomy, the measurement of IQ appears to be the lone survivor in the game of human classification (Gifford & O'Connor, 1992).

It is assumed that some traits such as intelligence are stable enough that they can be measured reliably. Yet when classification systems based on cut-off scores are introduced into this rather tenuous measurement process, it could be argued that the presence of error is being ignored if not forgotten. This seemingly flagrant violation of basic psychometric theory is so commonly practiced that school psychologists involved in psychoeducational assessment rarely question it. Consumers of psychoeducational reports expect—even demand—classifications or diagnostic labels to ensure that applications for funding will be accepted by government auditors. But the *end justifies the means* argument has worn too thin to withstand public scrutiny. An individual's right to needed educational services without him or her being categorized with diagnostic labels is on the horizon of a growing human rights movement (National Association of School Psychologists, 2004).

Ignoring the tacit knowledge (Sternberg, 2000) of below average students is tantamount to academic discrimination. If they are perceived as *less than* their higher-IQ peers, then one must expect *less than* adequate performance and progress in school. The school psychologist is in a unique position to weave together the psychometric with the observational and qualitative to produce a balanced summary of each individual student assessed. Placing too much weight on any variable will skew the results and sway perceptions of the student in question. Such influence is not to be taken lightly, and yet the demands of the job are frequently such that referrals for psychoeducational assessments continue to rise while the time available to complete thorough assessments becomes an increasingly precious commodity.

The presumed precision of an individual's IQ may be a somewhat fallacious assumption. The fact that the Borderline and Low Average classifications exist produces an expectation that the 10-point ranges in IQ that they represent are indeed meaningful and theoretically sound. The largely misunderstood Low Average label conveys an element of lowered expectations combined with the need for likely school curriculum modifications or adaptations without financial support from special-needs funding formulas and guidelines. The unspoken truth remains: unless your IQ is judged to be average or above, then expected deficits in academic achievement are viewed more as symptoms of possessing a less-than-average IQ rather than starting points for intervention.

In the educational psychology field, it is widely accepted that variability in test performance is axiomatic, and each referral must be treated as a separate individual (Kaufman, 1994). Perhaps the time has come for the interpretation and reporting of these results to concentrate more on the individual assessed than on their normative peers. This would mean relinquishing a strict adherence to IQ-based classifications and cut-off scores that serve to diminish individual differences in favor of group comparisons. Placing the educational needs of individual students ahead of normative comparisons with their peers is a futuristic notion, but one with both form and substance.

#### References

- Ackerman, T. (1998). Review of the Wechsler individual achievement test. In J.C. Impara & B.S. Plake (Eds.), *The thirteenth mental measurement yearbook* (pp. 1125-1128). Lincoln, NE: University of Nebraska Press.
- Bidell, T.R., & Fischer, K.W. (1997). Between nature and nurture: The role of human agency in the epigenesis of intelligence. In R.J. Sternberg & E. Grigorenko (Eds.), *Intelligence, heredity and environment* (pp. 193-242). New York: Cambridge University Press.
- Braden, J.S., DiMarino-Linnen, E., & Good, T.L. (2001). Schools, society, and school psychologists' history and future. *Journal of School Psychology, 39*, 203-219.
- Bryan, T.H., & Bryan, J.H. (1986). *Understanding learning disabilities* (3rd ed.). Mountain View, CA: Mayfield.
- Charter, R.A., & Feldt, L.S. (2001). Confidence intervals for true scores: Is there a correct approach? *Journal of Psychoeducational Assessment, 19*, 350-364.
- Denis, R. (2004). So far so good? A qualitative case study exploring the implementation of consultation in schools. *Educational Psychology in Practice, 20*, 16-29.
- Evans, L.D. (1990). A conceptual overview of the regression discrepancy model for evaluating severe discrepancy between IQ and achievement scores. *Journal of Learning Disabilities, 23*, 406-411.
- Field, A. (2000). *Discovering statistics using SPSS for windows: Advanced techniques for the beginner*. London: Sage.
- Flanagan, D.P., McGrew, K.S., & Ortiz, S.O. (2000). *The Wechsler intelligence scales and gf-gc theory: A contemporary approach to interpretation*. Needham Heights, MA: Allyn & Bacon.
- Fox, M. (2003). Opening Pandora's box: Evidence-based practice for educational psychologists. *Educational Psychology in Practice, 19*, 91-102.
- Gifford, B.R., & O' Connor, M.C. (1992). *Changing assessments: Alternative views of aptitude, achievement and instruction*. Norwell, MA: Kluwer Academic.
- Glutting, J.J., & McDermott, P.A. (1994). Core profile types for the WISC-III and WIAT: Their development and application in identifying multivariate IQ-achievement discrepancies. *School Psychology Review, 23*, 629-643.
- Learning Disabilities Association of Canada. (2002). *Official definition of learning disabilities*. Ottawa: Author. Retrieved May 1, 2007, from: [http://www.ldactaac.ca/Defined/defined\\_new-e.asp](http://www.ldactaac.ca/Defined/defined_new-e.asp).
- Kaufman, A.S. (1994). *Intelligence testing with the WISC-III*. New York: Wiley Interscience.
- Kronick, R.F., & Hargis, C.H. (1998). *Dropouts: Who drops out and why and the recommended action*. Springfield, IL: Charles C. Thomas.

- National Association of School Psychologists. (2004). *Position statement: Rights without labels*. Retrieved February 19, 2004, from: [http://www.nasponline.org/information/pospaper\\_rwl.html](http://www.nasponline.org/information/pospaper_rwl.html).
- Peterson, S.A. (2001). School psychologists in the political maelstrom. *School Psychology International, 22*, 291-302.
- Rourke, B.P. (1998). Significance of verbal-performance discrepancies for subtypes of children with learning disabilities: Opportunities for the WISC-III. In A. Prifitera & D. Saklofske (Eds.), *WISC-III clinical use and interpretation* (pp. 139-156). London: Academic Press.
- Sattler, J.M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Jerome M. Sattler.
- Shelton, B.O. (1971). *Teaching and guiding the slow learner*. New York: Parker.
- Sparrow, S.S., & Davis, S.M. (2000). Recent advances in the assessment of intelligence and cognition. *Journal of Child Psychology and Psychiatry, 41*, 117-131.
- Sternberg, R.J. (2000). Intelligence and wisdom. In R.J. Sternberg (Ed.), *Handbook of intelligence* (pp. 631-649). New York: Cambridge University Press.
- Sternberg, R.J., & Grigorenko, E.L. (2001). *Environmental effects on cognitive abilities*. Mahwah, NJ: Erlbaum.
- Swanson, H.L., Mink, J., & Bocian, K.M. (1999). Cognitive processing deficits in poor readers with symptoms of reading disabilities and ADHD: More alike than different. *Journal of Educational Psychology, 91*, 321-333.
- Ward, T.J., Ward, S.B., Glutting, J.J., & Hatt, C.V. (1999). Exceptional LD profile types for the WISC-III and WIAT. *School Psychology Review, 28*, 629-643.
- Wechsler, D. (1991). *Wechsler intelligence scale for children – third edition*. San Antonio, TX: Psychological Corporation.
- Woolfson, L., Whaling, R., Stewart, A., & Monsen, J. (2003). An integrated framework to guide educational psychologist practice. *Educational Psychology in Practice, 19*, 283-302.