2004

Empirically derived ability-achievement subtypes in a heterogeneous clinic-referred sample.

Rosemary S. Waxman
University of Windsor

Recommended Citation
Waxman, Rosemary S., "Empirically derived ability-achievement subtypes in a heterogeneous clinic-referred sample." (2004). *Electronic Theses and Dissertations*. 1922.
https://scholar.uwindsor.ca/etd/1922

EMPIRICALLY DERIVED ABILITY-ACHIEVEMENT SUBTYPES IN A HETEROGENEOUS CLINIC-REFERRED SAMPLE

by

Rosemary S. Waxman

A Dissertation
Submitted to the Faculty of Graduate Studies and Research
through the Department of Psychology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy at the
University of Windsor

Windsor, Ontario, Canada

2004

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisisitons et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-92540-4
Our file Notre référence
ISBN: 0-612-92540-4

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

The purpose of the present study was to identify clinically meaningful and reliable patterns of ability and achievement using the WISC-III and WIAT. As an extension of the work of Saunders, Casey, and Jones (2001), it was anticipated that several of the derived subtypes would share a similar profile to many of the subtypes described in their research, and that many of the derived subtypes would demonstrate a predictable pattern of neuropsychological test results. Cluster analysis was used to group the 182 WISC-III and WIAT profiles (10 WISC-III subtests and 4 WIAT subtests) of children between the ages of 9 and 14 years. Theoretical and empirical considerations were used to identify a cluster solution, which involved comparison of several five-, six- and eight cluster solutions. Ultimately, a five-cluster solution was selected as being representative of the data, which was well-replicated across three hierarchical clustering methods (i.e., complete linkage, average linkage-within groups, and average linkage-between groups (UPGMA)). The clusters were labeled based on their most salient characteristics, which included a group of predominantly Low Ability and achievement, a group demonstrating a pattern of verbal processing deficits, a group demonstrating a pattern of visual spatial/processing speed deficits, and a group with deficits consistent with an ACID pattern. Three of the subtypes were found to be highly similar to subtypes of Saunders et al., and all five subtypes had been identified in the learning disabilities literature. The external validity of the five subtypes was assessed through evaluation of the relationship between cluster membership and neuropsychological profile. Most predictions regarding neuropsychological performance were supported by the data, providing further evidence of the validity of the five-cluster solution. Clinical implications of the ability-achievement typology and suggestions for future research are discussed.

DEDICATION

This work is dedicated to my inspiration, pillar of strength, and soul mate, Aaron Waxman.

ACKNOWLEDGEMENTS

There are many people who deserve mention for assisting and inspiring me through this very long and involved process. First, I would like to express my deepest gratitude for the patience, generosity, and insightful comments of my committee members, Dr. Joseph Casey, Dr. Cory Saunders, Dr. Darren Fuerst, Dr. Douglas Shore, Dr. Krista Chandler, and my external reviewer, Dr. Douglas Ris. I am especially grateful to my advisor, Dr. Joseph Casey, an excellent teacher and mentor, who offered many kind words of encouragement and advice. I would also like to extend a special thanks to Dr. Douglas Ris for his thought provoking comments, and for driving in from Ohio to attend my defense.

I owe much thanks to Dr. John Strang and staff at the Ozad Institute of the Windsor Regional Children's Centre who made available to me the many neuropsychological test files utilized for the purposes of this work. Dr. Cory Saunders was also instrumental in helping me navigate the many complicated steps involved in running my cluster analyses.

I am forever grateful for the continued encouragement and support of my family and friends in Toronto, psychology staff at the Bloorview MacMillan Children's Centre, and students and faculty of the University of Windsor. In particular, I would like to recognize individually, the members of my incredibly supportive family; Mom and Victor, Dad and Joanne, Ted and Perlan, Melanie, Tara, Victoria, Elizabeth, Mark, Marnie, Mandy, Justin, Randy and Jody – you guys are the best!! In addition, Ms. Barb Zakoor was especially helpful in facilitating my long distance communication with my committee and throughout my years in the program, for which I am grateful. Finally, this project could not have been completed without the ongoing confidence, encouragement and support of my husband, Aaron Waxman.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Learning disabilities represent the largest handicapping condition among children in the United States (Stanford & Oakland, 2000). In Canada, learning disability is similarly the most common long-term condition among children from birth to 14 years of age (Statistics Canada's Health and Activity Limitation Survey, 1991). Approximately 5-10% of the school-aged population has been found to demonstrate difficulty in the acquisition of reading skills despite average intelligence, as well as adequate social and educational opportunities (Habib, 2000).

Although public policy delineates the conditions that must be satisfied for a child to be considered disabled, and ultimately, eligible for educational accommodations or services, the absence of a well-accepted, satisfactory definition of learning disability continues to encumber progress in the field. Indeed, the search for a universal definition of learning disability has been controversial since its inception as a diagnostic entity (Hooper & Willis, 1989). However, there is little disagreement that most definitions incorporate some notion of discrepancy, which is typically identified as unexpected underachievement relative to performance on intelligence testing (Sattler, 1992). Thus, adherence to such a definition implies that administration of standardized intelligence and achievement tests are necessary in the identification of learning disability (Fletcher et al., 2002; Slate, 1994). Within this conceptual framework, the importance of psychometrically sound, reliable, and valid instrumentation is of paramount importance.

In this vein, the Wechsler series of intelligence tests, among the most widely used instruments in the evaluation of children's intellectual functioning (Anastasi & Urbina, 1997; Goh, Teslow, & Fuller, 1981; Stinnett, Havey, & Oehler-Stinnett, 1994), have been demonstrated to be of considerable predictive utility with respect to academic and occupational performance (Anastasi & Urbina). The Wechsler Individual Achievement Test (WIAT) has the unique advantage over other measures of academic achievement in that it was co-normed on a subset of the Wechsler Intelligence Scale for Children – III (WISC-III) standardization sample, facilitating the computation of ability-achievement discrepancies.

Given the emphasis on standardized ability and achievement testing within the educational system, there has been significant research on the utility of the WISC-III. One avenue of research has endeavored to identify a typology of learning disabilities based on performance on intellectual testing (e.g., Donders, 1996, Konold, Glutting, McDermott, Kush, & Watkins, 1999). To date, this profile and cluster analytic literature clearly indicates that there is no single pattern of test results that characterizes all learning disabled children (Joschko & Rourke, 1985). Instead, there are a number of test score profiles that have been identified in different subgroups of children with learning disabilities. Despite the recognized need for replication and validation of learning disability taxonomies (Joschko & Rourke), subtyping studies utilizing the WISC-III alone or in combination with scores from achievement testing are limited almost exclusively to research on the standardization sample.

In one of the few cluster analytic studies to evaluate WISC-III performance among a referred sample, Saunders, Casey and Jones (2001) generated a six-cluster solution which was demonstrated to have both internal and external validity. They also identified subgroups, such as a group with low scores on tasks involving sequencing and language, which have received little attention in the WISC-III literature. Accordingly, the purpose of the present study was to extend the research of Saunders et al. in a similarly heterogeneous clinical sample. Efforts to generate a valid ability-achievement typology were accomplished through evaluation of children's performance on the subtests from the WISC-III and WIAT. Following the derivation of reliable subtypes, the external validity of the cluster solution was evaluated on auxiliary measures of neuropsychological functioning.

A Brief History of Learning Disabilities: Definitional and Diagnostic Criteria

The concept of learning disabilities, first identified as “minimal brain dysfunction,” evolved from the field of medicine in the early 20th century, a time when distinctions between organic and so-called functional impairments were in the forefront of diagnostic and treatment issues (Francis, Fletcher, Shaywitz, Shaywitz, & Rourke, 1996; Kessler, 1980). Based on the widespread practice of describing the behaviours of brain injured children, children with learning disabilities became grouped into this same general category (Kessler). It was not until the 1960s that conceptualizations of learning disability within an educationally based context appeared (Francis et al.). Within this context, the first formal definition of learning disabilities was advanced by the National Advisory Committee on Handicapped Children in 1968, which later became incorporated in Public Law 94-142 (Torgesen, 1991). Specifically, this definition stated:

> 'Specific learning disability' means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations (Torgesen, 1991, p. 20).

Additional exclusionary clauses such as learning difficulties due to a primary visual or auditory impairment, Mental Retardation or emotional disturbance were also included, serving to prevent such children from receiving services that were designated for the learning disabled population (Meyer, 2000).

In the United States, the current definition of learning disability proposed by the most recent revision of the Individuals with Disabilities Education Act is relatively unchanged from the 1968 legislation (Meyer, 2000); but has nevertheless incorporated the notion of a discrepancy between ability and achievement (Francis et al., 1996). Since the passage of Public Law 94-142 and its reenactment as the Individuals with Disabilities Education Act of 1990, children are required to undergo psychological reevaluation at least every three years to be considered eligible to receive special education or accommodations (Education for All Handicapped Children Act; U.S. Department of Health, Education, and Welfare [USDEHW], 1977). Such an evaluation must identify a significant discrepancy between ability and achievement for a child to be diagnosed as learning disabled. Similarly, the Learning Disabilities Association of Canada (LDAC, 2001) has recently revised their national definition of learning disability. Like its American counterpart, this definition also states that for a learning disability to be identified there must be evidence of academic underachievement, resulting from deficiencies in psychological processes, with otherwise average intellectual abilities (LDAC).

Historically, diagnosis of a learning disability has involved the evaluation of ability-achievement discrepancies, in which academic performance in one or more areas is markedly below expectations based on psychometric intelligence (Flanagan & Alfonso, 1993; Fletcher et al., 1994; Glutting, McDermott, Prifitera, & McGrath, 1994; Hooper & Willis, 1989; Sattler, 1992). Although current definitions adopted by school boards and North American legislation share common themes, they nonetheless differ in their definitions of learning disability and the required conditions under which such a diagnosis is warranted (Frankenberger & Fronzaglio, 1991; Mercer, King-Sears, & Mercer, 1990). Despite the variability across definitions of learning disability, most seem to indicate that learning disabilities are primarily characterized by underlying processing or "psychological" deficits (Daley & Nagle, 1995). Indeed, a *statistically significant* discrepancy between performance on standardized ability and achievement tests are among the criteria proposed by the Diagnostic and Statistical Manual of Mental Disorders (APA, 1994), most states, provinces, and IDEA in order for a child to be identified as learning disabled (Frankenberger & Frozaglio; Gridley & Roid, 1998).

A longstanding debate regarding the relevancy of IQ scores in the identification of learning disabilities has persisted into the 21st century. In a recent study (Speece & Shekitka, 2002) designed to examine experts' opinions on the components that should be included in an operational definition of reading disability, an empirically-based survey was mailed to 218 editorial board members of four journals devoted to learning disabilities. Of the 113 respondents, 30.2% endorsed the inclusion of an IQ and reading achievement discrepancy and 42% agreed that inclusion of an IQ score was important. Although there was general agreement on some definitional features, there was no predominant component that was rated as most important. This study reflects the general climate of learning disability identification and the lack of cohesion both between and within various disciplines.

Literature documenting several fundamental problems related to use of ability-achievement discrepancies in the diagnosis of learning difficulties has been prolific, spanning over several decades. Inherent in the use of IQ scores in the definition and diagnosis of learning disability is the assumption that intelligence and academic achievement can be measured independently of one another (Siegel, 1989). As many researchers and clinicians have similarly indicated, measures of aptitude and achievement are highly intercorrelated and as such, likely measure similar cognitive processes (Francis et al., 1996; Sternberg & Grigorenko, 2002). Accordingly, underlying cognitive deficits would be expected to have a similar effect on a child's academic and intellectual performance, resulting in a diminished ability-achievement discrepancy.

Francis, Epsy, Rourke, and Fletcher (1991) reviewed several important issues related to the validity of IQ scores in the definition of learning disability. Among the key problems identified were the concept of intelligence as a reflection of learning potential, psychometric shortcomings, and limitations associated with the assumption that individuals with reading disability commensurate with intellectual ability were distinct from those with discrepant ability-achievement profiles. Other problems inherent in the use of ability-achievement discrepancies are that children with lower IQ scores are automatically deemed ineligible for services designed to assist learning disabled children because they are less likely to demonstrate a significant discrepancy between ability and achievement. Children of minority, non-English, or culturally diverse backgrounds are thus placed at a disadvantage in comparison to their English-speaking peers due to qualities of the standard testing instrumentation rather than deficiencies that are unique to the child. That is, children from minority or different cultural backgrounds who exhibit difficulties reading and low IQ scores are penalized in not being labeled reading disabled due to their low IQ scores and hence, do not qualify for remedial services (Siegel, 1989).

Although there is strong empirical support that IQ and academic performance are highly correlated in normal populations, the literature to date has demonstrated that ability predicts no more than approximately 50% of the variability in achievement (Gridley & Roid, 1998), further obfuscating issues pertaining to learning disability diagnosis and eligibility determinations. Indeed, in considering the significance of IQ scores in the diagnosis of learning disability, it is important to consider the expected lack of predictive validity of IQ scores in learning disability assessment. The underlying assumption is that learning disabilities are due to some processing deficiency that prevents the child from achieving his or her maximum potential (Meyer, 2000).

Accordingly, potential is measured by one's performance on the WISC and therefore, one must assume that the WISC is invalid in the prediction of academic achievement in learning disabled populations.

In an attempt to reconcile the differences in conceptualization and definitions of learning disability, several diagnostic approaches have been advocated (Meyer, 2000). Indeed, discrepancy formulae are as varied as the definitions and approaches in the identification of learning disabilities. Some of the reasons for these differences include lacking consensus about the relevance of discrepancy in the diagnosis of learning disability, variation in operational definitions of discrepancy, and a number of psychometric problems inherent in their usage (Rispens, van Yperen, & van Duijn, 1991). Many researchers have therefore examined the influence of different discrepancy models and computations on ability-achievement outcomes and learning disability diagnoses.

While an in depth discussion of ability-achievement discrepancy models is beyond the scope of the current work, several comprehensive reviews and empirical investigations have been published (e.g.,Fletcher et al., 1994; Francis et al., 1991; Francis et al., 1996; Kamphaus, Frick, & Lahey, 1991; Sternberg & Grigorenko, 2002). Briefly, these critical analyses have consistently found that comparison between various discrepancy models yields highly discrepant results. Thus, different procedures used in the evaluation of learning disabilities have a profound effect on the accuracy of learning disability classifications and the determination of children's eligibility for educational accommodations or services. This unavoidable quandary underscores the need for consistent definitional and diagnostic criteria, agreement around the importance and use of IQ tests in the assessment of learning disabilities, and which IQ test should be employed.

The Wechsler Intelligence Scales

Wechsler's test of intelligence was originally designed to assess global intelligence in adults (Wechsler Bellevue Intelligence Scale), from which complementary measures designed for school-aged children (Wechsler Intelligence Scale for Children; WISC) and preschoolers (Wechsler Preschool and Primary Scale of Intelligence; WPPSI) were later modelled. Wechsler's series of intelligence tests are among the most widely used instruments in the evaluation of children's intellectual functioning (Goh et al., 1981; Stinnett et al., 1994). Their popularity is firmly based on a comprehensive literature indicating that global psychometric intelligence is one of the most enduring predictors associated with academic and occupational success (Anastasi & Urbina, 1997; Kaufman, 1994; Weiss & Prifitera, 1995).

Instead of the scales being derived from a specific theory of intellectual functioning, Wechsler's measures are pragmatically based and were designed to measure global intellectual functioning through a broad array of verbal and nonverbal tasks (Anastasi & Urbina, 1997; Kamphaus, 1993). It was not until the 1950s that several theories of intelligence were advanced, and ultimately applied to interpret intelligence tests already existing, such as the Wechsler scales (Kamphaus).

The Wechsler series evolved from the Binet-Simon scales (Binet & Simon, 1905, cited in Reynolds & Kaufman, 1990), the first comprehensive individual intelligence test to adopt a point-scale and the inclusion of several brief tasks, which in combination yielded a composite score (Boake, 2002; Reynolds & Kaufman; Sattler, 1992). However, Wechsler's scales differed in that they provided a means by which a number of summary measures could be computed in addition to the traditional single summary score provided by the original Binet-Simon (1905) and its revised formats (as cited in Reynolds & Kaufman). It is important to note, however, that instead of being originally created by Wechsler, many of the Wechsler subtests were derived from existing testing instruments in addition to those developed by Binet, including the Army Alpha and Beta Examinations (Sattler).

The WISC-III (Wechsler, 1991), like its predecessors the WISC (Wechsler, 1949) and WISC-R (Wechsler, 1974), was intended for use in children aged 6 through 16 years, and includes a number of verbal and nonverbal subtests, yielding three IQ scores (i.e., Verbal, Performance, and Full Scale) and four factor scores (Verbal Comprehension Index [VCI], Perceptual Organization Index [POI], Freedom from Distractibility Index [FDI], Processing Speed Index [PSI]). The WISC-III is comprised of 13 subtests in total; 10 of which are mandatory, and three which are supplementary (two performance and one verbal subtest) of which two contribute to the computation of factor scores and one (Mazes) is not included in any of the composite scores. Despite the retention of the 10 mandatory subtests and many of the specific test items, the WISC-III includes substantial changes to the WISC-R, and is viewed as the leading measure of intellectual functioning in school-aged children (Kaufman, 1994).

Since the advent of the most recent edition of the Wechsler scales for school-aged children, efforts to validate and replicate the construct dimensions of the instrument have been numerous (e.g., Kamphaus, Benson, Hutchinson, & Platt, 1994; Roid & Worrall, 1997). This interest in the latent structure of the Wechsler scales is predicated, in part, upon a long standing debate surrounding what the scales actually measure and their preferred level of interpretation. However, in the absence of a strong theoretical grounding, the search for a coherent basis of interpretation is difficult at best (Anastasi & Urbina, 1997). Although the most recent version of the WISC has evolved considerably from its predecessors in areas such as item content, subtest length and normative sampling, and statistical properties (Anastasi & Urbina; Kaufman, 1993; Roid, Prifitera, & Weiss, 1993), there remains considerable debate regarding the interpretability and structure of the index scores and their predictive utility in the assessment of learning disability.

Psychometric Properties of the WISC-III

The WISC-III manual (Wechsler, 1991) provides a rather comprehensive section on validity and a new and more ethnically and culturally diversified standardization sample in comparison to the WISC-R. Some have offered significant praise of the instrument, asserting that the improvements in test materials "are matched by the psychometric excellence that went into test development and by the meticulous care that characterized the preparation of the thorough, intelligently written manual" (Kaufman, 1993, p. 345). Despite the obvious benefits in revising testing instrumentation, which include the updating of normative data, contemporary stimulus materials, and extension of ceiling and floor levels, there are also costs associated with test revisions which go beyond the associated monetary expenditures. Researchers have commented on some of the weaknesses of the WISC-III, which include an increased emphasis on speed in the determination of children's IQ scores, the elimination of clinically relevant questions, and a reduction in the stability of some of the subtests (Kaufman).

In the context of these fundamental differences, investigations of the psychometric properties of the WISC-III have been both extensive and fundamental to its success as the leading measure of children's intellectual functioning. Research has consistently demonstrated that children initially assessed for learning difficulties with the WISC-R exhibit comparatively lower performance on the WISC-III of approximately five IQ-points (Bolen, Aichinger, Hall, & Webster, 1995; Carlton, & Sapp, 1997; Wechsler, 1991). This phenomenon has been attributed to a statistical artifact known as regression towards the mean (Sternberg & Grigorenko, 2002). As a consequence of outdated norms, mean decreases in performance on measures of intellectual ability (when comparing performance on earlier and updated versions of the same measure), reflecting an upward drift of approximately one third of one point per year have also been documented by Flynn (1984). Reported decreases in performance on the WISC-III compared to the WISC-R are consistent with this phenomenon, given the 17 years between the publication of the WISC-R and the WISC-III. These findings clearly warn of the potential disservice that can be done to children evaluated with the WISC-R and re-evaluated with the WISC-III, particularly in the context of determination of educational placement and special services.

Apart from the studies reported in the WISC-III manual, which account for a substantial proportion of the research documenting IQ score differences between the WISC-R and WISC-III, an additional study using an independent sample of learning disabled children (n = 53) found similar results (Slate, 1995a). In his study, Slate found that although WISC-R and WISC-III IQ scores correlated strongly, lowered IQ scores of 8, 5, and 7 points for the Full Scale, Verbal, and Performance scales, respectively, were evident. Bolen et al. (1995) and Carlton and Sapp (1997) have also reported similar findings.

The stability of WISC-III IQ scores was examined in a sample of children (n = 34) receiving special education (Slate, Jones, & Saarnio, 1997). Children were administered the WISC-III as part of their initial assessment and were re-evaluated three years later. Results of paired t-tests indicated that there was no significant difference between any of the IQ scores [i.e., Full Scale IQ (FSIQ), Verbal IQ (VIQ), Performance IQ (PIQ)] between initial diagnosis and re-evaluation with significant correlations between each of the IQ scores. The authors indicated that their results confirm previous research which has indicated that declines in IQ from the WISC-R to the WISC-III are an artifact of norming and not a reflection of diminished intelligence. Obvious implications of these research findings are that many children who may have been near discrepancy cut-off scores and diagnosed as learning disabled with the WISC-R may no longer meet eligibility requirements when reevaluated within three years of their initial assessment. It is therefore incumbent upon clinicians reevaluating children originally assessed with the WISC-R to take appropriate precautions to ensure that children are not unduly denied services.

Canivez and Watkins (2001) obtained findings in keeping with those of Slate and colleagues (1997) in their study examining the differential stability of WISC-III IQ scores of children from independent disability subgroups. Children diagnosed with Specific Learning Disability, serious emotional disability, and Mental Retardation demonstrated stable IQ scores across evaluations with FSIQ reflecting the most stability over time. Reliability of subtest scores, Freedom from Distractibility and Processing Speed Index Scores (FDI and PSI, respectively), and VIQ-PIQ discrepancy scores were inadequate and not recommended for use in the determination of students' eligibility for special educational services. Similar findings were reported in an earlier study examining the long-term stability of the WISC-III in a large heterogeneous sample of learning disabled children (Canivez & Watkins, 1998).

In addition to changes in IQ scores, the WISC-III manual (Wechsler, 1991) reports adequate reliability with standard errors of measurement of 3.54, 4.54, and 3.2 for VIQ, PIQ, and FSIQ, respectively. These are, however, somewhat lower than those reported for the WISC-R. The manual also includes reference to a number of validity studies of both the predictive and concurrent variety.

Weiss and Prifitera (1995) examined the predictive validity of WISC-III FSIQ scores across ethnic and gender groups in 1000 children from the WIAT standardization sample. The authors found that FSIQ was significantly correlated with each of the WIAT composite measures across ethnic and gender groups. Furthermore, differential prediction of WIAT scores was observed in only 4 of the 12 comparisons, which were deemed to reflect very small effects. The authors concluded that their results provide strong support for the hypothesis that WISC-III FSIQ adequately predicts academic performance on the WIAT across racial and ethnic groups, although they caution that ability-achievement discrepancies should not be used independent of corroborating information in the diagnosis of learning disability.

In an independent study, Slate (1995a) reported significant correlations between the WISC-III FSIQ, VIQ and PIQ with the Reading subtest of the Wide Range Achievement Test-Revised (WRAT-R) in a sample of children receiving special education services (n = 32). He similarly found that WISC-III FSIQ, VIQ and PIQ scores were significantly correlated with performance on the Peabody Picture Vocabulary Test-Revised (PPVT-R).

Although the WIAT manual (Wechsler, 1992) reports moderate to high correlations between WISC-III IQ and index scores and subtest measures from the WIAT, few studies have examined their relationship within a predominantly clinical or learning disabled sample. This is particularly problematic in that several researchers have appropriately indicated that IQ scores are the result or product of a range of cognitive processes and, as such, may be differentially affected depending on the underlying processing deficiency of the child (Francis et al., 1996; Slate, 1994).

Slate (1994) was the first researcher to investigate the relationship between obtained WISC-III and WIAT scores in a group of children with learning difficulties.

Comparison between the WISC-III IQ and index scores with WIAT subtest scores in a heterogeneous group of children (202 children with Specific Learning Disabilities, 115 children with Mental Retardation, and 159 children who did not meet eligibility criteria for special services) was performed. He found that the majority of correlations between WISC-III and WIAT scores were higher than those reported in the WIAT manual. Interestingly, the weakest correlations between FSIQ and WIAT subtests was among children who did not qualify for special education services, suggesting that other variables may be interfering with the learning potential of these children. The relationships between FSIQ and VIQ scores and WIAT subtests were similar across the three samples justifying their use in regression formulas. However, the correlations between the WISC-III Index (particularly the VCI and POI) and WIAT subtest scores were significantly lower than those computed with the WISC-III IQ scores, suggesting that the WISC-III indexes should not be used in discrepancy analyses. Slate also cautioned against the use of the WIAT to diagnose learning disabilities until research has resolved the differences between the findings reported in the WIAT manual and those of his own research.

Factor Structure of the WISC-III: What is the WISC-III Measuring?

Research on the WISC-III has followed apace in the footsteps of its predecessors. Consistent with results of factor analytic research on the WISC-R, the extant literature on the WISC-III clearly supports the validity of the global measure of intelligence (FSIQ), the verbal and nonverbal dichotomy (VIQ and PIQ, respectively), and the Verbal Comprehension and Perceptual Organization factors (Hynd, Cohen, Riccio, & Arceneaux, 1998). Despite the fact that the Symbol Search subtest was originally developed to clarify the contribution of the Processing Speed factor, its inclusion instead resulted in two separate factors (Processing Speed and Freedom from

Distractibility; Wechsler, 1991). Thus, rather than providing clarification, the addition of Symbol Search, and hence the addition of a fourth factor, confuses rather than clarifies the constructs measured by the WISC-III (Keith & Witta, 1997). While some researchers maintain that the four-factor model reported in the WISC-III manual (Wechsler, 1991) best defines the structure of the test (Konold, Kush, & Canivez, 1997; Roid et al., 1993), others have found this solution inadequate (Burton et al., 2001; Kush, 1996).

Roid and colleagues (1993) sought to extend the factor analytic research on the WISC-III and replicate the factor solution reported in the WISC-III manual. Study participants consisted of 1,118 children who were administered the WISC-III and WIAT in conjunction with the standardization of the WIAT. From this independent group of children, approximately 100 were selected at each full year of age. The authors performed both exploratory and confirmatory factor analyses on the 13 subtests of the WISC-III, in which they compared numerous extraction and rotation methods. Initial analysis of the intercorrelation matrix revealed the presence of a global factor, "g," consistent with results of previous research. Comparison of factor extraction methods was performed using various quantitative criteria. Their findings supported the well-documented four-factor solution presented in the WISC-III manual (Wechsler, 1991). Specifically, their resulting factor structure consisted of two major factors (Verbal Comprehension and Perceptual Organization), accounting for 54% of the variance, and two smaller supplementary factors (Freedom from Distractibility and Processing Speed) accounting for 6% and 8% of the variance, respectively.

Roid et al. (1993) tested seven models based on a variety of goodness-of-fit statistics. Results of the confirmatory factor analysis similarly corroborated the adequacy of the four-factor model. Criterion validity of the four factors was evaluated through analysis of the relationship between the WISC-III factor indexes and composite achievement scores from the WIAT. It was determined that all four factor indexes of the WISC-III have clinical utility in the assessment of individual differences related to school performance. However, it is important to note that these authors included the Mazes subtest, an optional subtest not included in the computation of any of the index or IQ scores, with omission of this subtest reflecting the norm in test administration (Konold, Glutting, McDermott, Kush, & Watkins, 1999). Thus, generalizability of these results may be limited to settings that routinely administer the Mazes subtest.

Kush (1996) similarly sought to extend the factor analytic research of the WISC-III to an independent sample of learning disabled children (n=327) in grades 1 through 11. Children were identified as learning disabled based on criteria established by the Arizona Department of Education which stipulates a severe discrepancy (i.e., 1.5 standard errors of estimate) between IQ and academic achievement based on regression-derived criteria. The 12 core WISC-III subtest scores were subjected to maximum-likelihood factor analysis followed by orthogonal (Varimax) and oblique (direct oblimin) rotations of all factors that exceeded eigenvalues of 1.0.

Results of Kush's (1996) factor analyses revealed an underlying factor, "g," which accounted for 34% of the total WISC-III variance. The remaining three factors consisted of two major and one smaller factor. The major factors, which appeared to correspond with the Verbal Comprehension (VC) and Perceptual Organizational (PO) dimensions of the WISC-III, collectively accounted for 39% of the variance, consistent with the findings reported for the standardization sample. The third factor, reflecting the Processing Speed Index, was defined by loadings from Coding and Symbol Search; however, this factor did not account for enough variance to be considered an independent factor. The Arithmetic subtest loaded with the first factor (VC), and Digit Span stood alone, thus, Kush's research failed to identify the FD factor. Based on these findings, Kush concluded that the WISC-III, like its predecessor, offers a robust measure of global intellectual functioning with well-defined verbal (VIQ) and nonverbal (PIQ) dimensions. Caution regarding the interpretation of profiles based on information other than the Verbal and Performance components was offered with suggestions for future research to examine the differential factor structure of the WISC-III across diverse populations. In keeping with the findings of Kush, Kamphaus et al. (1994), subjected the WISC-III standardization data to confirmatory factor analyses at each age group and did not consistently find support for the four-factor solution.

A more recent study (Konold et al., 1997) examined the WISC-III factor structure in three independent samples of children from special education classrooms. The first sample was comprised of 229 children between the ages of 6 and 15 of whom 81% had been diagnosed as learning disabled. Sample two consisted of 246 learning disabled children between the ages of 3 and 13 years. Two hundred and forty learning disabled children ranging in age from 8 to 13 years comprised the third sample. Performance on the 12 primary WISC-III subtests for each sample were evaluated through a series of five confirmatory factor models. Consistent with the extant literature, when a one-factor model was specified, all 12 subtests loaded on a single dimension. Comparison between the five different models across samples revealed that the four-factor solution presented in the WISC-III manual (i.e., VC, PO, FD, PS) consistently demonstrated the best overall fit. The authors concluded that their research provided strong support for the interpretation of WISC-III factor scores. They suggested that future research investigate the implications of deriving factor scores from only the 10 mandatory subtests, and that more homogenous samples of learning disabled children be utilized (Konold).

Support for the four-factor solution has also been demonstrated in the Canadian normative sample of 1100 children between the ages of 6 and 16 years (Roid & Worrall, 1997). In this study, all 13 WISC-III subtests were included in analyses which involved exploratory principal-axis factor analysis with oblique rotation. Keith and Witta (1997) similarly found support for the four-factor model. They subjected the WISC-III standardization data to hierarchical confirmatory factor analysis. They found that the FDI emerged as the best predictor of general intelligence, and indicated that this calls into question the accuracy of its title, which suggests cognitive characteristics of a less intellectually demanding nature. Indeed, their findings were later supported in the study by Reinecke, Beebe and Stein (1999) in which the FDI was found to be related to learning difficulties but not diagnostic of inattention or ADHD.

Riccio, Cohen, Hall, and Ross (1997) similarly examined the clinical utility of the FDI and PSI in differentiating between clinical groups (learning disabled, ADHD inattentive, ADHD combined type) and the relationship between these factors and other neuropsychological and behavioural measures. They found that scores on the FDI and PSI were significantly correlated with each other across clinical samples and that each of the indexes was also correlated with VCI and POI. Neither of the indexes correlated with measures of attention and concentration or behaviors specific to inattention or hyperactivity, nor were they helpful in the differential diagnosis of ADHD and learning disability. Their results further indicate that poor performance on the FDI or PSI should not be interpreted as indicative of the presence or absence of ADHD. Of note, the authors found a significant correlation between the FDI and measures of working memory and recommended that additional research specifically explore the clinical utility of this relationship (Riccio et al.).

In examining the consistency of the four-factor solution across the 11-year age span of the instrument, Keith and Witta (1997) found that the WISC-III did, indeed, measure the same characteristics across all age groups. However, the authors acknowledged that their goal was to evaluate the hierarchical model outlined in the WISC-III manual (Wechsler, 1991) and they therefore did not examine the viability of other potential models.

However, evaluation of additional factorial models was conducted in a study by Burton et al. (2001) in which the authors performed a maximum-likelihood confirmatory factor analysis on all 13 WISC-III subtest scores of 318 children referred for psychoeducational assessment. Their results were cross-validated by reevaluating the same nine latent models in the WISC-III standardization sample. Unlike most prior research (e.g., Kamphaus et al., 1994; Roid & Worrall, 1997), the authors examined more complex models of intelligence than the commonly cited four-factor structure. Their results supported a five-factor solution in which they clearly established that Mazes did not account for a significant proportion of the variability in any of the factors. Instead, the best fitting model across both the clinical and standardization samples was that in which Mazes was excluded from the analysis. The following descriptions and subtest groupings comprised their five-factor solution: Verbal Comprehension (Vocabulary, Similarities, Information, Comprehension); Constructional Praxis (Picture Completion, Block Design, Object Assembly); Visual Reasoning (Picture Arrangement); Freedom from Distractibility (Arithmetic, Digit Span) and Processing Speed (Coding, Symbol Search). In contrast to research reported in the WISC-III manual (Wechsler, 1991) in which analyses were restricted to a maximum of four to five factors, Burton et al. evaluated nine models with up to seven factors. Indeed, the authors cited this as one of the possible reasons for their disparate findings. Additional support for their model was demonstrated by its correspondence to Kaufman's application of Horn's Gf-Gc theory (Fluid and Crystalized Intelligence) to the WISC-III. Similarities between the model and Horn's theory further suggests that the WISC-III contains very few measures of fluid intelligence, with the Visual Reasoning factor (i.e., Picture Arrangement) reflecting some measure of fluid intelligence. Furthermore, apart from this factor being defined by a single subtest, this subtest in particular has relatively low reliability (Kaufman, 1994) and caution in using it as a primary measure of fluid ability is therefore warranted.

Glutting, Youngstrom, Ward, Ward, and Hale (1997) examined the incremental efficacy of the four WISC-III factor scores. That is, they performed regression analyses to determine whether and to what extent each of the WISC-III IQ scores or factor indexes predict performance on the WIAT composite measures in both clinical (n = 636) and nonreferred (n = 283) samples. Their results indicated that FSIQ was the best predictor of achievement scores with no significant improvement in prediction using VIQ, PIQ, VCI, POI, FDI, or PSI when FSIQ was partialed out. Curiously, the FDI exhibited the largest unique relationship to reading, math, and writing performance represented by WIAT composites, but it accounted for only 1.4 - 5.2% of the variance in achievement. Glutting and colleagues indicated that in the absence of evidence to the contrary, clinicians are justified in abbreviating the WISC-III by omitting Mazes and Symbol Search, neither of which contribute to the FSIQ.

In summary, there is little disagreement that the WISC-III is a stable and adequate measure of global intellectual functioning, with an ability structure comprised of both verbal and nonverbal components. The first two factors, Verbal Comprehension and Perceptual Organization, have remained relatively unchanged across revisions of the instrument (Little, 1992). Some studies (e.g., Glutting et al., 1997) have also indicated that among the factor scores, the FDI may provide unique information about the cognitive functioning of children with learning difficulties, particularly with respect to academic achievement as measured by the WIAT. Nonetheless, the literature is relatively inconsistent despite there being a number of studies that support the four-factor-solution reported in the WISC-III manual (Wechsler, 1991). These studies seem to suggest that the WISC-III is sampling an adequate range of constituent skills that contribute to a child's global intellectual functioning.

Conversely, other researchers have indicated that the WISC-III four-factor structure is not as clear or reliable as described in the manual, and that the FDI "should be ignored" (Little, 1992, p. 153). Furthermore, few studies have examined models of greater than four factors and, thus, the possibility remains that the latent structure of the WISC-III may be better explained by a more complex factorial model. In addition, the issue of which subtests to include in analyses continues to hamper the generalizability of research findings, consequently limiting their clinical utility. Despite the fact that the WISC-III is composed of 13 subtests, in practice, most clinicians do not administer all 13. Instead, only those subtests factored into the calculation of IQ and index scores are typically given (Konold et al., 1999). This procedure has further empirical support in that the Mazes subtest: a) has been shown to contribute little to the factorial structure of the WISC-III (Burton et al., 2001); b) loads minimally on the global factor (Kaufman, 1994); and c) has low reliability (Kaufman). Taken together, the inconsistencies in the literature indicate that further investigation of the WISC-III factor structure across diverse populations is warranted.

The Profile Analytic Literature

Analysis of children's patterns of performance on WISC subtests, factor scores and/or IQ scores, often referred to as profile analysis, has its origins in the earliest versions of the Wechsler scales. The approach emerged from a widespread assumption that the average child, free of any significant learning difficulties or psychopathology, should exhibit a "flat" profile (Smith, Smith, Matthews, & Kennedy, 1993). This belief was dispelled following the early work of Kaufman (1976) who demonstrated that normal children do in fact exhibit subtest scatter of between five and nine points on the WISC-R. Following the publication of the first Wechsler scale, profile analysis became the predominant interpretive practice, characterized by an emphasis on deriving numerous subscores in an attempt to identify significant profiles and factors (Kamphaus, 1998).

The profile analytic movement has persisted into the 21st century and is fairly widespread in clinical practice, despite its reported lack of empirical support (Kamphaus, 1998; Macmann & Barnett, 1997b; Watkins, Kush, & Glutting, 1997). There is a large following of clinicians who are strongly opposed to profile analysis, citing the empirical and psychometric inadequacies as support for their position (e.g., Kamphaus, Macmann & Barnett). Conversely, strong proponents of profile analysis, some of whom have developed comprehensive interpretive systems and devoted textbooks to an in-depth treatment of the topic, acknowledge its clinical and diagnostic utility (e.g., Kaufman, 1994).

Bannatyne (1974, 1979) was among the first independent researchers to offer an auxiliary configuration of WISC subtest scores designed to facilitate the identification of children with genetic dyslexia. Several interpretive approaches and configurations have since been proposed, with an upward of 75 patterns of subtest variation having been devised for the Wechsler series (McDermott, Fantuzzo, & Glutting, 1990). These have varied in their degree of complexity, theoretical grounding, and most importantly, subjection to empirical investigation. Interestingly, despite the variability across interpretive approaches, many seem to have taken a particular interest in the significance of the conglomerate of subtests comprising the WISC-R FDI, namely, Arithmetic, Coding, and Digit Span. The paragraphs to follow will review a selection of predominant and representative profile analytic approaches devised for the Wechsler Scales for Children. It should be noted that the majority of these interpretive systems were originally developed for the WISC and/or WISC-R, although they have been extended for use with the WISC-III.

While a 12-point difference between VIQ and PIQ is considered statistically significant, base rate data reported in the WISC-III manual (Wechsler, 1991) suggest that more than one-third of the children in the standardization sample obtained VIQ-PIQ discrepancies of 12-points or greater. Therefore, it is common practice that only discrepancies which occur very infrequently (i.e., in less than 10% of the standardization sample) be considered to reflect a clinically meaningful difference (Prifitera, Weiss, & Saklofske, 1998). Some have indicated that this cut-off is somewhat arbitrary and instead, recommend a VIQ-PIQ discrepancy of 20-points or greater to be considered of clinical significance (Prifitera et al.).

Despite the general acceptance that significant discrepancies between Verbal and Performance IQ scores render the Full Scale IQ meaningless as a composite measure of general intellectual capacity, there are several potential approaches in analyzing and deriving meaning from VIQ-PIQ differences (Reynolds & Kaufman, 1990). Research at the University of Windsor has investigated the heterogeneity of learning disabilities, through which reliable and valid subgroups have been identified. This research program originated with the evaluation of VIQ-PIQ discrepancies in children referred for neuropsychological assessment. It is important to note that this procedure was initiated as a research model from which hypotheses could be formulated and tested and not as a diagnostic tool in and of itself. Accordingly, in the first of a series of studies (Rourke, Young, & Flewelling, 1971), three groups of children were formed based on the relationship between their VIQ and PIQ scores. The three groups were characterized as having high VIQ and low PIQ (10-point difference), high PIQ and low VIQ (10-point difference), and VIQ and PIQ within 4-points of each other (V = P). The authors found significant differences in neuropsychological and academic test performance of the three groups such that the high VIQ and low PIQ group performed comparatively better on measures of verbal and auditory perceptual skills, the high PIQ and low VIQ group performed comparatively better on measures of visual-perceptual skills, and the V = P group performed intermediate to the other two groups.

Subsequent investigations (Rourke & Telegdy, 1971; Rourke, Dietrich, & Young, 1973) utilized the same VIQ-PIQ criteria to classify children, and independently examined differences in performance on motor and psychomotor measures in younger (aged 5 to 8 years) and older (aged 9 to 14 years) children. As predicted, older children exhibiting the high PIQ low VIQ pattern demonstrated superior performance on most measures of complex motor and psychomotor abilities in comparison to the V = P, and high VIQ low PIQ groups. These results were perceived as providing support for the hypothesis that WISC VIQ-PIQ discrepancies reflected the differential integrity of the cerebral hemispheres in older learning-disabled children. Although there were few significant differences among the younger groups on measures of motor and psychomotor abilities, this was thought to reflect the substantial developmental variability in children aged 9 to 14 years and the emergence of progressive differentiation of abilities within these subgroups of learning-disabled children. The results of these studies stimulated a comprehensive research program that ultimately led to the identification and classification of the syndrome of nonverbal learning disabilities. The primary author of this research program indicated in a recent review of the literature that in order to permit greater generalizaiblity of the research findings to date, these studies should be replicated using the WISC-III (Rourke, 1998).

Among the first published studies to evaluate IQ and index score discrepancies on the WISC-III within a clinical sample, Slate (1995) found statistically significant IQ and index score discrepancies across each of the clinical samples employed (Specific Learning Disability, Mental Retardation, and children who did not meet criteria for either diagnosis). Consistent with previous research with the WISC-R (Kavale & Forness, 1984; Smith et al., 1993), mean PIQ was higher than mean VIQ for each of the three groups. The magnitude of the derived discrepancies, which were curiously smaller than those reported for the standardization sample, nonetheless indicate that significant VIQ-PIQ discrepancies are found in non-disabled and disabled populations alike and should therefore not be viewed as diagnostic of learning disability. Similarly, Humphries and Bone (1993) compared WISC-R subtest scatter and academic performance on the WRAT among a group of learning-disabled children exhibiting a VIQ-PIQ discrepancy (favoring the latter) and a group of slow learners characterized by low VIQ and PIQ. The authors indicated that the groups did not demonstrate significant cognitive differences beyond the higher PIQ score among the learning-disabled children, which was one of the primary differentiating variables on which the groups were initially formed. These results would further attest to the limited utility of VIQ-PIQ discrepancies in the differential diagnosis of learning disability.

Kaufman (1979, 1994), one of the most vocal advocates of profile analysis, proposed that WISC subtest recategorization provided information beyond the global measures that was useful in the diagnosis of learning disability. His system involved the recategorization of WISC-R subtests and subsequent comparison of composite groupings as follows: Reasoning (Similarities, Arithmetic, Comprehension) versus Recall (Information, Vocabulary, Digit Span); Long Stimuli (Information, Arithmetic, Comprehension) versus Brief Stimuli (Similarities, Vocabulary, Digit Span); Much Expression (Similarities, Vocabulary, Comprehension) versus Little Expression

(Information, Arithmetic, Digit Span); and Meaningful Stimuli (Picture Completion, Picture Arrangement, Object Assembly) versus Abstract Stimuli (Block Design, Coding).

Smith et al. (1993) found that 22 of the 29 students in their group of learning-disabled children exhibited the Kaufman WISC-R regrouping scores in the expected sequence (Reasoning > Recall, Much Expression > Little Expression, Long Stimuli > Brief Stimuli, and Meaningful Stimuli > Abstract Stimuli) in comparison to 10 of the 30 non-disabled children. The authors concluded that the Kaufman regrouping may be of diagnostic significance.

Grossman and Galvin (1987) found that performance on Kaufman's Recall, Brief Stimuli, and Little Expression categories on the WISC-R significantly predicted WRAT Reading, Spelling, and Arithmetic subtest scores in a group of referred children (n = 105). The authors also indicated that consistent with the pattern of results found in a group of children with conduct disorders (Paget, 1982), the children comprising their clinical sample appeared to perform better on tasks requiring reasoning and problem-solving skills in comparison to those tasks requiring the recall or retrieval of information. However, due to the nature of their sample (referral population in the public schools), the limited generalizability of their findings was acknowledged (Grossman & Galvin).

Bannatyne (1974, 1979), whose interpretive system dominated much of the earlier profile analytic movement, proposed a recategorization of WISC subtests to aid in the identification and evaluation of learning disabled children (Kaufman, 1994; Sattler, 1992). His approach evolved from the clinical application of the Wechsler scales and was later revised based on factor analytic research (Kaufman; McKay, Neale, & Thompson, 1985). His labels explicitly convey the proposed abilities subserved by each of his categories, which are as follows: Verbal Conceptualization Ability (Similarities, Vocabulary, Comprehension); Acquired Knowledge (Information, Arithmetic, Vocabulary); Spatial Ability (Picture Completion, Block Design, Object Assembly); and Sequencing Ability (Arithmetic, Digit Span, Coding) (Kaufman).

Within Bannatyne's tripartite model, a pattern of Spatial > Conceptual > Sequential was hypothesized to identify children with learning disability. Although this factor structure has been supported in learning disability samples using the WISC and WISC-R (Grossman & Galvin, 1987; McKay et al., 1985; Rugel, 1974; Saklofske, Schmidt, & Yackulic, 1984), research has also indicated that this pattern of performance is not stable over time (Saklofske et al.), does not characterize many individual children (Kavale & Forness, 1984), and has not been successful in the differential diagnosis of learning difficulties in comparison to other special education categories (Kavale & Forness). However, it is important to note that Bannatyne's Sequential category bears a resemblance to the WISC-III FDI (with only the addition of Arithmetic) and includes all three of the subtests which comprise the WISC-R FDI. The relevance of performance on these subtests in learning disability populations is also consistent with recent factor analytic studies of the WISC-III (Keith & Witta, 1997; Reinecke et al., 1999).

Another profile, the ACID pattern, that denotes those WISC profiles in which Arithmetic, Coding, Information, and Digit Span subtest scores are depressed relative to the remaining subtests, has been fairly consistently identified in children with learning disabilities (Kaufman, 1994; Reynolds & Kaufman, 1990; Ward, Ward, Hatt, Young, & Mollner, 1995). This profile has been associated with poor academic performance in reading, spelling, and arithmetic, and significant early reading problems (Joschko & Rourke, 1985). Similar to Bannatyne's Sequential category, three-quarters of this profile is comprised of subtests from the WISC-R FDI (and two of the four subtests are from the WISC-III FDI). Accordingly, some researchers have suggested that special attention be paid to the interpretation of performance on the FDI within the learning disabled population (Reynolds & Kaufman).

Apart from the clinical utility of this profile, a series of studies designed to examine the heterogeneity of children exhibiting the ACID profile have indicated that there are at least two reliable subtypes of learning-disabled children within two distinct age groups (5 to 8 years and 9 to 15 years) who exhibit the ACID pattern (Joschko & Rourke, 1985). Following their demonstration of internal validity, Joschko and Rourke examined the external validity of the derived classification of children exhibiting the ACID pattern. They found that performance on the WRAT Arithmetic and Reading subtests was significantly lower in the younger and older ACID groups, respectively, in comparison to matched learning-disabled controls who did not exhibit the ACID profile. In analyzing patterns of neuropsychological test performance, they further demonstrated that each of the subgroups of children who exhibited the ACID profile had a qualitatively different ability profile of potential clinical significance in educational and remedial programming.

As previously indicated, the majority of research examining the ACID profile has been conducted using the WISC-R. Much of this earlier research was based on group means rather than reflecting the proportion of individuals who presented with the ACID profile and it had neglected to explore the incidence of the ACID profile within clinical samples (Daley & Nagle, 1995). However, since the publication of the WISC-III manual (Wechsler, 1991), some additional research examining the utility and incidence of the ACID profile has been conducted.

In addition to the data presented in the WISC-III manual (Wechsler, 1991) indicating that children with learning difficulties or ADHD earned lower scores on the ACID grouping than the FDI or PSI, current research has reported a greater incidence of the ACID profile among children with ADHD and learning disability in comparison to the standardization sample (Prifitera & Dersh, 1993; Ward et al., 1995). It is important to note that in each of the studies where a greater incidence of the ACID pattern was observed among clinical samples, a higher proportion of children exhibited only a partial ACID profile (lowered scores on any three of the four subtests), and of those studied the partial ACID profile occurred rather infrequently (between 1% and 5.1%).

Despite its apparent utility in identifying children with learning disability, Kaufman (1994) has suggested the replacement of the ACID profile with the SCAD profile (Symbol Search, Coding, Arithmetic, and Digit Span) for use with the WISC-III. His rationale stems in part from the findings of Prifitera and Dersh (1993) who found that the learning disabled children included in their research earned their lowest mean scores on the subtests that collectively comprise the FDI and PSI. Performance on the Information subtest did not reflect a relative weakness for their learning disability group. In contrast, Snow and Sapp (2000) found evidence of lowered scores on the ACID, Bannatyne Sequential, and SCAD profiles in two independent groups of children with ADHD in comparison to the WISC-III standardization sample. They indicated that their results would suggest that children with attentional difficulties exhibit similar patterns of performance on the WISC-III, which may prove valuable in diagnosis and the formulation of interventions (Snow & Sapp).

Research on the validity and reliability of the SCAD profile has been limited, and of the few published studies, there seems to be little support for its utility in the identification of learning-disabled children. Watkins et al. (1997) examined the prevalence and diagnostic utility of the SCAD profile among learning disabled children (n = 332), children with emotional disorders (n = 31), and those children from the WISC-III standardization sample who obtained FSIQ scores greater than 70 (n = 2158).

Following the procedure recommended by Kaufman (1994), SCAD indexes were obtained by summing the scaled scores for each of the four subtests, the total of which was subtracted from the sum of the Picture Completion, Picture Arrangement, Block Design, and Object Assembly scaled scores. The authors found that the SCAD profile did not accurately classify children into disabled and non-disabled groups nor was it predictive of children's academic performance on selected subtests from the Woodcock Johnson-Revised, Tests of Achievement (Reading, Math and Written Expression). Based on their findings, the authors discouraged the use of the SCAD index or profile analysis in general for either diagnostic or clinical hypothesis testing endeavours.

The conclusions of Watkins et al. (1997) are consistent with the findings of Ward et al. (1995) who did not find a greater incidence of the SCAD profile across their groups of children referred for psychoeducational assessment when compared to the standardization sample. Indeed, Ward et al. found that the SCAD profile represented the most prevalent configuration in the largest proportion of children in both the clinical and standardization samples. These researchers also examined the conditional probabilities and incremental gains in diagnostic classification, and found that both the ACID and SCAD profiles may be more useful in the differential diagnosis of learning disability in comparison to other disabilities. However, they noted that because students with ADHD, who as a group have also been shown to exhibit a higher incidence of ACID and SCAD profiles, were not classified into a distinct category, their possible inclusion in the learning disability group would have increased the incidence of these profiles. Accordingly, upon separation of the ADHD children from the learning disability group, the utility of both profiles would be expected to decline.

In a comprehensive evaluation of the clinical utility of the WISC-III in the identification of learning disabilities and learning disability subtypes, Daley and Nagle

(1995) examined the WISC-III performance (omitting Mazes) of 308 children who qualified for special education services. Among their analyses was the extent to which study participants exhibited VIQ-PIQ discrepancies, subtest scatter, ACID and SCAD profiles, and Bannatyne recategroziation hierarchies.

Daley and Nagle (1995) found no greater subtest variability in their learning disabled sample than that reported for the standardization sample. Similarly, although the mean VIQ-PIQ discrepancy for their learning disabled sample was significantly different than the mean discrepancy reported for the standardization sample, the cumulative frequency of this discrepancy was approximately 27%, which also occurred in 21.6% of the standardization sample. Although the Bannatyne profile (Spatial > Conceptual > Sequential) was obtained at the group level, only 26% of the children exhibited this profile. Similarly, in their evaluation of ACID and SCAD profiles, Daly and Nagle found that only 1% and 2% of the subjects exhibited the complete ACID and SCAD profiles, respectively. On the other hand, 12% of the learning disabled sample exhibited a partial ACID profile, which the authors acknowledged may be considered to provide limited support for the partial ACID profile as diagnostic of learning disability.

Results of factor analysis also conducted by Daly and Nagle (1995), which accounted for 73% of the variance, revealed a three-factor solution that was consistent with the factor structure of the WISC-R. Using the factor scores, a two-cluster solution was generated through cluster analysis in which the two groups were differentiated across all three factors. That is, one group was characterized by uniformly below average abilities while the second group was defined by average abilities. Based on their findings, the authors concluded that Wechsler subtest patterns and recategorization have limited utility in the differential diagnosis of learning difficulties. However, it is important to bear in mind that the use of factor scores in cluster analysis may reduce the distinctiveness of the groups due to their normalized distribution (Aldenderfer & Blashfield, 1984).

Despite the popularity of profile analysis, many vocal opponents share the sentiments of Daley and Nagle (1995). Frank (1983), in his comprehensive review of the Wechsler series, cites results of factor analytic research as empirical evidence against profile analysis. He explains that in the absence of data supporting the psychometric independence of subtests, and hence the degree to which they actually reflect independent cognitive processes, pattern analysis should be avoided. Similar observations have been made by several researchers (Kamphaus, 1993) while others have taken a more extreme stance and have admonished the use of IQ testing in the educational system altogether (Macmann & Barnett, 1997a, 1997b; Siegel, 1989).

In summary, critics of profile analytic approaches have strongly opposed its practice, denouncing it from a number of perspectives. There is little consistency across studies that have provided statistical support for the diagnostic utility of pattern analysis (Anastasi & Urbina, 1997). However, it is important to note that reorganization, conceptualization, and analysis at the level of subtest scores have inherent value in the documentation of children's cognitive strengths and weaknesses. Kaufman (1994) indicated that reorganization approaches are typically designed with the intent of providing the clinician "with information not readily apparent from the IQs or standard scores yielded by the instrument to offer better understanding of the cognitive capabilities of individual children or well-defined groups" (p. 56). Factor analytic studies have further provided evidence of the reliability of multiple-factor interpretations of the WISC-III, suggesting that additional research would be helpful in the clarification of the ambiguous results to date (Ward et al., 1995). Furthermore, given the popularity of the Wechsler scales, and the resources designed to systematize, facilitate, and improve utilization of profile analysis, it is clear that this technique is a preferred approach in the interpretation of intellectual test data (Anastasi & Urbina, 1997), warranting greater attention in the scientific literature.

The WIAT

The WIAT is an individually administered test battery used in the assessment of academic skills and achievement in children aged 5 through 19 years. The WIAT was developed to address three primary goals: the creation of a test relevant to educational curriculum trends, one that could be linked with the Wechsler scales to enable meaningful ability-achievement discrepancies, and one that would include subtests that parallel the seven areas of achievement specified in Public Law 94-142 (Wechsler, 1992). Specifically, these latter areas consist of: oral expression, listening comprehension, written expression, basic reading skill, reading comprehension, mathematical calculations, and mathematical reasoning.

Following the Wechsler tradition, the WIAT is comprised of individual subtests (eight) which combine to yield four academic composite scores (Reading, Mathematics, Language, and Writing) and a total summary measure (Total Composite score). Similar ~~to that of the WISC-III, the WIAT manual devotes a chapter to the coverage of reliability~~ and validity, which reveals stability coefficients for composite and total scores across all five grade levels as follows: Reading .93, Math .91, Language .78, Writing .94, Screener .95, Total .96.

Three of the WIAT subtests (Basic Reading, Mathematics Reasoning, and Spelling) are published separately as a brief screening measure, making it more comparable to the widely used Wide Range Achievement Test (WRAT). Despite the fact that the WRAT and the WIAT are both measures of academic functioning, they are nonetheless comprised of subtests that are conceptually quite different. For example, while Smith and Smith (1998) found significant and moderate to high correlations between performance on the WRAT-3 and WIAT in a group of learning disabled children, significant mean differences between the WRAT Reading and WIAT Reading Comprehension and between WRAT Arithmetic and WIAT Numerical Operations subtests were noted. The authors indicated that the two measures appear to provide similar results when screening for reading, spelling and arithmetic but that due to differences in specific tasks offered, clinicians should select the subtest that most appropriately evaluates the cognitive domain of interest. Additionally, the WIAT was co-normed with the WISC-III, having the advantage of reliable statistical computation of ability-achievement discrepancies (Flanagan & Alfonso, 1993) and may therefore be a more appropriate choice within a psychoeducational context.

In a recent study, Casey and Yawny (2001) examined the relationship between 12 WISC-III subtests (10 core subtests, Symbol Search and Digit Span) and a selection of subtests from the WIAT (Basic Reading, Mathematics Reasoning, Spelling and Numerical Operations) using principal component factor analysis. The authors were particularly interested in the construct validity of the Mathematics Reasoning subtest, for which task demands are quite distinct from those of the Arithmetic subtest of the WRAT. Specifying a four-factor solution based on findings of a similar study (Yawny, Casey, & King, 2001), Casey and Yawny found that all four WIAT subtests loaded with the WISC-III subtests that comprise the Freedom from Distractibility Index (i.e., Arithmetic and Digit Span). However, of the four WIAT subtests, Mathematics Reasoning demonstrated the lowest loading on this factor, instead, sharing variance in common with Information, Similarities, Arithmetic, Vocabulary, and Comprehension. It is noteworthy that each of these tasks emphasizes verbal processing. Conversely, the Numerical Operations subtest consistently loaded with the subtests from the Freedom from Distractibility Index, suggesting greater emphasis on attention, sequencing, and symbolic processing. Casey and Yawny asserted that the Mathematics Reasoning subtest is multidimensional in nature and comparatively more complex than tasks involving pure arithmetic calculations. Because Mathematics Reasoning appears to emphasize linguistic skills, it may therefore not be an equivalent alternative to the WRAT Arithmetic subtest.

Saunders, Strang, and Jones (2001) examined WIAT patterns of performance in a split-half sample of 298 clinic-referred children. The authors subjected each of the randomly formed groups to three-stage cluster analysis involving *k*-means iterative partitioning analyses, followed by Ward's analysis with Euclidian distance as the similarity measure. To correct for fusion errors and improper initial assignments, a second *k*-means analysis was performed using mean centroids from the Ward's analysis. The resulting four WIAT subtypes were classified according to the following characteristics: 1) global deficits on all WIAT subtests, 2) better performance on WIAT subtests of mathematical ability in comparison to reading and spelling subtests, 3) better performance on WIAT subtests of reading and spelling in comparison to subtests of mathematical ability, and 4) average performance on all four WIAT subtests. The authors indicated that these four subtypes were consistent with previously generated subtypes using the WRAT, which speaks to their validity in clinical populations and the suitability of the WIAT in the identification of specific learning difficulties.

Cluster Analysis

Multivariate techniques can be broadly grouped into two general categories: *R* analyses and *Q* analyses (Glutting, et al., 1994). Whereas *R* analyses are predicated on the assumption that the measures to be compared vary in a linear fashion, *Q* analyses, such as clustering techniques, evaluate test scores as part of an integrated profile and, as such, are generally sensitive to trends in both level and shape (Glutting et al.). The comparison among subtests, factors, or IQ scores is a common practice in psychoeducational assessment. Accordingly, ability-achievement discrepancy evaluation is often performed using *R* analyses due to its being well-suited to the evaluation of linear relationships, such as those between test scores (Glutting et al., 1994). However, Q methodologies have the advantage of simultaneously analyzing level, pattern, and dispersion of scores within a given profile which guides the classification of cases into homogenous groups. For this reason, the use of Q type procedures is best suited for taxonomic research.

Cluster analysis refers to a family of multivariate statistical techniques used in the generation of classification schemes or typologies. That is, cluster analysis is often undertaken when the primary research endeavour is to find similar groups of observations in a sample of data (Aldenderfer & Blashfield, 1984). Many different clustering algorithms are available, each offering a different approach to the clustering of data. Indeed, each hierarchical technique utilizes different criteria for determining whether an individual belongs to a specific cluster based on his or her similarity or dissimilarity to the cluster (Morris et al., 1998). Accordingly, different solutions may be obtained depending on the clustering algorithm used.

Critical to the design of classification research is the application of procedures that assess the stability of the resultant clustering solution (Morris, Blashfield, & Satz, 1981; Morris & Fletcher, 1988), or its internal validity. Such stability analyses are essential to ensure that derived cluster solutions are not randomly generated or method dependent (Morris et al.). Researchers (Morris et al.; Morris & Fletcher) have identified a number of statistical methods that may be used for this purpose, which include cross-sample or split-sample replications, the addition or deletion of random variance via subjects or variables, or subjecting the data to alternate statistical techniques (e.g., a different clustering procedure). In addition, iterative partitioning techniques which allow for the reclassification of individuals to different clusters with each pass through the data are often applied to hierarchical clusters to further clarify and refine initial hierarchical solutions. External validity, or the extent to which the typology is clinically meaningful and discriminates between subgroups on variables not used in their formation, should also be undertaken as a final evaluation of the generated cluster solution (Morris et al.; Morris and Fletcher).

Empirically Derived WISC-III Taxonomies

Relatively few studies have attempted to develop and validate a typology of distinct learning profiles based on the WISC-III. Research utilizing the WISC-III alone or in combination with achievement tests is largely confined to research on the standardization sample, which presumably comprises children who are free of disability or pathology.

Through a series of studies, Konold et al. (1999) sought to identify profile subtypes within the standardization sample of the WISC-III. Scores from the 10 mandatory WISC-III subtests for the 2,200 children and adolescents were sorted according to shape and level using multistage cluster analysis. The resulting subtest taxonomy included eight subtypes, for which subtest and IQ score patterns and qualitative descriptions were provided. The eight core subtypes were classified as follows; high ability, above average ability, above average ability and VIQ > PIQ, average ability and PIQ > VIQ, average ability and VIQ > PIQ, below average ability and PIQ > VIQ, below average ability, and low ability.

To evaluate the external validity of the subtypes, demographic characteristics (e.g., age, gender, ethnicity, education placement) for each group were presented as prevalence percentages, which were compared with expected prevalence rates based on the total WISC-III standardization sample. In six of the eight profile types, there were significant differences between expected and actual prevalence percentages. A multivariate procedure designed to evaluate whether a given WISC-III profile reflects unusual subtest variation not accounted for in one of the eight-core profile types was also illustrated. Konold et al. (1999) cautioned against forming diagnostic hypotheses based on analysis of the WISC-III subtests due to lacking empirical evidence for the utility of subtest analysis in the identification of children with learning problems or psychopathology.

Donders (1996) sought to identify subtypes of WISC-III performance based on index scores. The methodology employed involved multistage cluster analysis on the Index scores of the entire WISC-III standardization sample using squared Euclidean distance as the similarity measure. First, he subjected the entire sample to Ward's analysis, which was later evaluated for reliability through the complete linkage method. Through this procedure, all five clusters of the original Ward's analysis were accurately replicated, suggesting good reliability. *K*-means iterative partitioning was then performed using results from the first-stage cluster analysis as a final-stage cluster procedure.

Donders (1996) went on to find that of the five WISC-III profiles, three could be differentiated based on level of performance. Cluster 2 exhibited below-average scores on all four indexes, Cluster 3 was characterized by above average scores (greater than one standard deviation above the mean) on all four indexes, and Cluster 4 had average scores across all four indexes. The two remaining clusters were characterized by pattern of performance, primarily in the context of scores on the Processing Speed Index. Specifically, average scores on the Verbal Comprehension, Perceptual Organization, and Freedom from Distractibility Indexes with a relative strength on the Processing Speed Index (greater than 13 points higher) characterized Cluster 1. Conversely, Cluster 5 demonstrated a relative weakness on the Processing Speed Index (9-12 points lower than the other three factors). Evaluation of the external validity of the subtypes was carried-out through evaluation of age and parental education for each of the subgroups. Donders' results indicated that parental educational levels co-varied with children's scores on intelligence testing. Specifically, more than 40% of the children in Cluster 2 (i.e., group with lowest overall level of performance) were from families in which the parents had less than 12 years of education, in comparison to the less than 20% reported for the standardization sample. In summary, Donders' results suggest that both level and pattern of WISC-III performance contribute to the differentiation of cognitive subtypes. In keeping with the recommendations of Konold et al. (1999), Donders cautioned against using WISC-III subtests in favor of index scores in the interpretation of WISC-III profiles due to the latter's greater reliability.

In a later study, Donders and Warchausky (1997) evaluated WISC-III index score patterns in a sample of 153 children who had sustained a traumatic head injury (THI). Repeating the same procedure utilized in a previous evaluation of the standardization sample (Donders, 1996), four reliable subgroups were identified, three of which differed primarily by level of performance. In the only group that demonstrated a distinct pattern of performance, scores on the POI and PSI were approximately one standard deviation below scores on the VCI and FDI. Although the WISC-III performance of this unique profile could not be attributed to demographic characteristics, it was found to be related to injury severity (highest proportion of children with severe injuries and longest length in coma), and neuropathology (highest proportion of children with diffuse and focal right hemisphere lesions) (Donders & Warchausky).

Also evaluating data independent of the WISC-III standardization sample, Saunders et al. (2001) performed a three-stage cluster analysis on the 12 WISC-III subtest scores that comprise the factor indexes. Study participants consisted of a heterogeneous sample of children and adolescents (n=343) between the ages of 6 and 16 years, with Full Scale IQ scores ranging from 70 to 130. The first stage involved a series of *k*-means cluster analyses, which indicated a 6 to 8 cluster solution would best fit the data. Hierarchical cluster analysis was performed using Ward's analysis, with squared Euclidean distance as the similarity measure. Additional *k*-means analyses were run to correct for improper initial assignments. These analyses supported a six-cluster solution in which the subgroups were labeled according to their mean profile patterns. The six clusters were as follows: 1. Broad based processing deficiencies; 2. Deficient language abilities; 3. Deficient nonverbal abilities; 4. Deficits consistent with an ACID pattern; 5. Deficient working memory; and 6. Deficits in tasks involving visual sequencing and language abilities (Saunders et al.).

External validation of the subtypes identified by Saunders et al. (2001) was accomplished through comparison of academic performance profiles for each of the six clusters on subtests from the WIAT (Basic Reading, Spelling, Mathematics Reasoning, and Numerical Operations). Results of a MANOVA and separate univariate analyses for each of the four WIAT subtests indicated significant differences in WIAT performance between the identified WISC-III subtypes.

In an attempt to replicate the findings of Saunders et al. (2001) using a different procedure, Waxman, Casey and Fuerst (2003) subjected the WISC-III data of the same 343 subjects to *Q*-factor analysis. The solution which assigned the greatest proportion of subjects was selected as being most representative of the data. This resulted in a six-group solution, in which many of the WISC-III profiles corresponded with those described by Saunders et al. Specifically, the subtypes were labeled based on outstanding group characteristics and consisted of the following: a group with verbal processing deficits; a group demonstrating a pattern of scores consistent with the SCAD profile (lowest scores on Symbol Search, Coding, Arithmetic, Digit Span); a group exhibiting depressed scores on measures of visual sequencing and language; a group with nonverbal processing deficits; a group with average processing speed skills in an otherwise below average profile; and a group with deficits on tasks requiring nonverbal problem solving and working memory. Five of the six profiles were highly similar to those reported by Saunders et al., providing preliminary support for the validity of their cluster solution.

In summary, cluster analytic research with the WISC-III has been fairly limited. Of the few studies that have sought to identify learning disability subgroups based on WISC-III scores, it is apparent that methodological variability and an almost exclusive interest in the standardization sample limit the generalizability of the findings to date. That having been said, there appear to be some common trends worth noting. The most salient finding shared across studies is that relatively homogenous subgroups of children with learning or processing deficits can be classified based on WISC-III performance. Children with primary processing deficiencies characterized by their performance on the PSI were consistently identified. Similarly, there appear to be a number of children who share common WISC-III profiles based on overall level of performance. However, external validation of the derived subgroups across studies has primarily examined academic, demographic, and injury severity/neuropathology variables (Donders, 1996; Donders & Warschausky 1997; Konold et al., 1999; Saunders et al., 2001). Further cluster analytic research with the WISC-III utilizing more varied clinical samples (e.g., children with known or suspected learning difficulties) with more comprehensive external validation procedures would be of great utility.

Empirically Derived WISC-III-WIAT Taxonomies

As mentioned throughout the present review, the demonstration of ability-achievement discrepancies is essential in the diagnosis of learning disability, and consequently, determinations of eligibility for special education services (Gridley & Roid, 1998; Schuerholz, et al., 1995). Accordingly, Glutting et al. (1994) sought to identify the most common ability and achievement profiles for the linking sample of the WISC-III and WIAT. Because these measures utilize different testing procedures based on age group (i.e., some of the WIAT subtests are omitted in the assessment of younger children), study participants consisted of a selection of children from the linking sample (n=824) ranging in age from 8 years, 9 months through 16 years, 11 months. Study participants were further divided into eight random samples, each consisting of 103 children. This was done for the purpose of conducting replication analyses of the cluster solutions.

Glutting et al. (1994) used a three-stage cluster analysis procedure to sort the four WISC-III factor indexes and WIAT composite scores for each group of children. In addition, three agglomerative algorithms (Ward's, Average Linkage, and Sarle's estimated maximum likelihood), and several internal- and external-validity analyses were performed. All statistical criteria (e.g., pseudo F ratio, homogeneity coefficient) were best satisfied through the third stage iterated Ward's six-cluster solution and thus, it was selected as the best taxonomy of WISC-III-WIAT ability and achievement profiles. The resulting six profiles were differentiated predominantly by level of performance and were labeled according to variations in FSIQ scores. The six subtypes were described as follows:

1. High ability and VIQ>PIQ;
2. Above average ability with slightly above average achievement and PIQ>VIQ;
3. Average ability with underachievement in writing;
4. Average ability with over achievement in reading, mathematics, language, and writing;

5. Below average ability with below average achievement; and
6. Low ability with underachievement in reading, mathematics, writing, and PIQ>VIQ.

Of note, Glutting et al. indicated that five of the six subtypes were not “flat” but, instead, were characterized by patterns of strengths and weaknesses. The authors argued that because nearly one half of the most representative profiles are characterized by ability-achievement discrepancies, clinicians and educators should exercise caution in diagnosing learning disabilities based on such “commonly occurring score patterns” (p. 621).

In a later study, Ward, Ward, Glutting, and Hatt (1999) extended the cluster analytic research of Glutting et al. (1994) in their investigation of WISC-III and WIAT composite scores of 201 children identified as learning disabled. The first cluster analysis, performed using Ward’s analysis, revealed a five-cluster solution which was compared against the six core profile types derived from the WISC-III and WIAT linking sample (Glutting et al., 1994). These five clusters were described as being similar in number and pattern to those obtained in previous research, although, three of the five clusters demonstrated ability-achievement patterns that were not suggestive of learning disability. Of the 201 cases entered into the analysis, 70.1% of the cases exhibited ability-achievement profiles that were consistent with one of the six core profiles.

Ward et al. (1999) conducted a second cluster analysis using only the scores of cases that did not match one of the core profiles generated a two-group solution, which was consistent with two of the subtypes identified in previous research. The two groups were differentiated according to WISC-III FSIQ, PIQ, VIQ, and WIAT Reading, Mathematics, Language and Writing Composites. One bore a resemblance to the sixth core cluster identified by Glutting et al. (1994), characterized by low ability with underachievement in reading, mathematics, writing, and PIQ>VIQ, and the other reflected a classic psycholinguistic learning disability. Ward and colleagues concluded that their results provide evidence for significantly fewer distinct learning disability subtypes in comparison to previous studies in which comparison with a normal taxonomy was not conducted.

Both Glutting et al. (1994) and Ward et al. (1999) argue that the taxonomy of common ability and achievement scores represent normally occurring patterns of strengths and weaknesses in the child population and should therefore not be considered reflective of unusual over- or underachievement. However, exceptional or not, patterns of strengths and weaknesses should be considered informative in and of themselves. Indeed, it is through evaluation of strengths and weaknesses that information regarding the learning potential and adaptive abilities of a child may be obtained. Given the growing body of evidence suggesting that ability-achievement discrepancies are likely not the most appropriate or sensitive approach in the identification of learning disabilities (Fletcher et al., 1994), the evaluation of strengths and weaknesses through both quantitative and qualitative evaluation should be considered a fruitful endeavour in hypothesis generation and the ability to provide unique information that can direct future interventions.

In summary, the WISC-III cluster analytic literature may be best characterized as demonstrating considerable variability with respect to methodology (e.g., utilization of subtest or index scores), number of identified subgroups, and researchers' conclusions about the clinical utility of the derived subgroups in the identification of homogenous groups of learning disabilities. It is clear that further identification and evaluation of subgroups of children who exhibit learning and/or processing difficulties based on the WISC-III and WIAT would be a welcome expansion of the existing subtyping literature.

Current Study: Rationale and Summary

Based on the above review of the literature, it is evident that research on the WISC-III and learning disability diagnosis is replete with inconsistencies, investigator biases, and methodological variability. Cluster analytic research utilizing the WISC-III has been largely undertaken in the standardization or WISC-III-WIAT linking samples, which are primarily comprised of normal school-aged children. Based on the WISC-III subtests, the number of identified subgroups have ranged from five (Donders, 1996) to six (Saunders et al. 2001) to eight (Konold et al., 1999), with the latter study excluding all three supplementary subtests. Only two studies have analyzed WISC-III and WIAT scores conjointly (Glutting et al., 1994; Ward et all, 1999), both of which have looked at Index rather than subtest scores.

Attempts to externally validate previously identified typologies have also been somewhat limited. Glutting et al. (1994), Konold et al. (1999), and Donders (1996) analyzed the demographic characteristics of the WISC-III standardization sample in evaluating the external validity of their six, eight, and five cluster solutions, respectively. Although Donders found evidence of external validity in that levels of parental education were found to co-vary with children's performance on the WISC-III, Konold et al. found significant differences between obtained and expected prevalence rates of IQ-achievement discrepancies.

In spite of the inconsistencies and limitations in the learning disability subtyping literature, it is nevertheless abundantly clear that learning disability is a heterogeneous diagnostic category, characterized by a variety of primary and/or secondary processing deficits. Once identified, information about patterns of strengths and weaknesses may be used to identify children at risk for academic underachievement and be of value in the design of appropriate and effective educational interventions. Accordingly, it would seem that a WISC-III-WIAT taxonomy has the potential to serve this purpose, and considerably augment the information provided by previously identified WISC-III profiles.

Within this context, clarification regarding the nature, and stability of specific learning disability subgroups would be of benefit. To this end, the six-cluster solution of Saunders et al. (2001) is unique from other studies in that some of their derived WISC-III profiles have been less frequently identified in the literature, perhaps reflecting subtle processing difficulties within their clinical sample. Saunders et al. also examined the external validity of their cluster solution on four WIAT subtests and found significant differences between the subgroups. These findings provide further support for the clinical meaningfulness of their six-cluster typology.

Thus, the purpose of the current study was to extend the research of Saunders et al. (2001) in an attempt to identify clinically meaningful subgroups of children who demonstrate processing and learning difficulties that can be identified based on performance on the WISC-III and selected subtests from the WIAT. Given the importance of intellectual and academic test data in the identification of learning disabilities, the current research sought to generate an ability-achievement typology based on the test scores of a clinic-referred sample. Due to the range of available hierarchical clustering techniques, which may consequently produce different solutions, the internal validity of the derived subtypes was examined through comparison of solutions obtained through different clustering techniques. Similarly, validation of the typologies was accomplished through multi-stage cluster analysis, a procedure designed to correct for fusion errors and improper initial subject assignments.

Once a representative and replicable solution was obtained, the relationship between subgroup membership and performance on a range of neuropsychological measures was examined in an attempt to establish the external validity of the subtypes.

Demonstrating the external validity of the derived typology should provide important information about the neuropsychological profiles associated with various patterns of processing and learning difficulties, and, thus, the clinical meaningfulness of the subtypes. To date, no study examining learning disability subtypes based on children's performance on the WISC-III and WIAT has included a similar external validation component as that which was undertaken in the present research.

Hypotheses:

Given the utility of empirically driven subtyping endeavours, and the findings of previous research, the following hypotheses were formulated:

Identification and internal validation of a WISC-III-WIAT typology

Based on previous research, particularly the findings of Saunders et al. (2001), it was hypothesized that the present study would yield a valid ability-achievement typology. It was further hypothesized that the derived typology would be similar to the cluster solution generated by Saunders et al.. More specifically, it was predicted that the derived solution would include the following clusters:

(a) i: a subgroup with relative verbal deficiencies (PIQ>VIQ; low scores on WISC-III verbal subtests; low achievement on WIAT reading and spelling tasks),

ii: one with broad-based processing deficiencies (reduced WISC-III IQs, WISC-III and WIAT subtest scores),

iii: another with relative nonverbal deficiencies (VIQ>PIQ; low scores on WISC-III Performance subtests; low achievement on WIAT subtests of mathematical ability), iv: an ACID pattern (lowest WISC-III subtest scores on Arithmetic, Coding, Information and Digit Span; reduced scores on the WIAT, especially on language-based tasks).

External validation of a WISC-III-WIAT typology

It was hypothesized that subgroup membership would be associated with neuropsychological profiles in a manner consistent with neuropsychological theory and previous research findings. Given the expectation that the derived typology would be similar to that of Saunders et al. (2001), the following general hypotheses regarding the neuropsychological characteristics of some clusters were developed:

(b) i: The presence of a cluster that is characterized by relative verbal deficiencies (VIQ<PIQ; low scores on WISC-III verbal subtests; low achievement on WIAT reading and spelling tasks) would be expected to demonstrate lower scores on measures of the auditory-linguistic domain in comparison to clusters characterized by an ACID pattern or relative nonverbal deficiencies.

ii: The presence of a cluster which is characterized by broad-based processing deficiencies (reduced WISC-III IQs, WISC-III and WIAT subtest scores) would be expected to demonstrate lower scores on nonverbal problem solving/concept formation, attention and concentration, and auditory-linguistic measures in comparison to clusters characterized by an ACID pattern, relative nonverbal deficiencies, and relative verbal deficiencies.

iii: The presence of a cluster which is characterized by relative nonverbal deficiencies (VIQ>PIQ; low scores on WISC-III Performance subtests; low achievement on WIAT subtests of mathematical ability), would be expected to demonstrate lower scores on measures of nonverbal problem solving/concept formation in comparison to clusters characterized by an ACID pattern or relative verbal deficiencies. Based on previous research (Rourke & Telegdy, 1971), this group was also predicted to demonstrate weaker psychomotor skills than clusters consistent with an ACID pattern, verbal deficiencies, and broad-based deficiencies.

iv: The presence of a cluster which is characterized by an ACID pattern (lowest WISC-III subtest scores on Arithmetic, Coding, Information and Digit Span; reduced scores on the WIAT, especially on language-based tasks) would be expected to demonstrate lower performance on attention and concentration and auditory-linguistic measures in comparison to the cluster characterized by relative nonverbal deficiencies.

CHAPTER II

METHOD

Participants

Cases to be considered for inclusion in the present research were obtained from an archival database of children referred to an outpatient neuropsychological service within a children's mental health facility in Southwestern Ontario. Participant data were derived from consecutive patient referrals from 1993 to 2002. All subjects were seen due to perceived impairments in academic, intellectual and/or socioemotional functioning with problems being primarily of developmental origin rather than an acquired medical condition. Those children who met the following criteria were retained for inclusion in the present study: WISC-III Full Scale IQ score between 70 and 130, and completion of each of the neuropsychological measures to be analyzed. The age range of selected participants was restricted due to differences in neuropsychological testing materials used with children under 9 and over 14 years of age. Because the present study sought to extend the findings of Saunders et al. (2001), and to include the largest sample available, all subjects who were included in their research, who also had complete neuropsychological data, were retained for analyses in the present study.

The final sample comprised a subset (i.e., 52%) of the Saunders et al. (2001) study participants in addition to three subjects (who met the study criteria) who had since been added to the database, and were included in the current study to maximize the number of participants. The remaining 48% (n=163; 115 males, 48 females) of the Saunders et al. sample did not meet the current study criteria primarily on the basis of age (i.e., almost 60%), and included children from both the lower and upper bounds of the distribution. Specifically, 90 children between the ages of 6 and 8 years, and 6 children between 15 and 16 years were excluded from the current study sample. The remaining participants were excluded because they did not have complete neuropsychological data, with a disproportionately greater number of participants (25 children) missing scores on the Tactual Performance Test. In addition, many participants (24 children) were missing data on multiple measures.

Of the excluded subset of the Saunders et al. (2001) sample, there were 11, 36, 43, 20, 18, 5, 8, 8, 8, 5, and 1 participant at age levels 6 through 16, respectively (*M* = 9.08, *SD* = 2.42). An ANOVA revealed a significant difference in mean age between the excluded subset of the Saunders et al. sample and the current study sample [F(1, 343) = 54.14, *p* < .001]; however, the samples did not differ in terms of mean WISC-III VIQ [F(1, 343) = .975 *p* = .324], PIQ [F(1, 343) = .2.45 *p* = .119], or FSIQ [F(1, 343) = .173 *p* = .678] scores. The overall mean WISC-III VIQ, PIQ, and FSIQ scores for the excluded subset of the Saunders et al. sample were 89.48 (*SD* = 12.34), 90.64 (*SD* = 14.66), and 89.02 (*SD* = 11.72), respectively.

Though the current study sample originally consisted of 183 participants, following initial cluster analyses, one subject was deemed an outlier and excluded from the study. The final sample therefore consisted of 182 children (144 males, 38 females) between the ages of 9 and 14 years (*M* = 10.64, *SD* = 1.42). There were 45, 51, 43, 18, 17, and 8 children at age levels 9 through 14, respectively. Participating children were predominantly right handed (n = 154, 84.6%) with the remaining participants (n = 26, 14.3%) being left-hand dominant, excluding one child who was identified as ambidextrous, and another whose laterality information was not available. The overall mean WISC-III VIQ, PIQ, and FSIQ scores were 88.24 (*SD* = 11.14), 92.97 (*SD* = 12.92), and 89.53 (*SD* = 10.97), respectively.

Measures

Test measures were administered and scored according to standardized procedures by trained examiners. All test scores were obtained within the context of a comprehensive neuropsychological evaluation, which in addition to the WISC-III and WIAT, comprised selected tests from the Halstead-Reitan Neuropsychological Test Battery for Older Children (HRNB-C) and auxiliary neuropsychological measures. For a detailed description of the WISC-III, WIAT, and neuropsychological measures used in the current study, refer to Appendixes I through IV.

The choice of variables to be included in cluster analysis is among the first and most important considerations in the design of subtyping research (Aldenderfer & Blashfield, 1984). WISC-III and WIAT data were specifically selected as the variables on which cluster analyses would be performed for two primary reasons. First, because of the emphasis on standardized ability and achievement scores in diagnosis and subsequent service delivery decisions within the educational system, it seemed fitting that these measures be used in the generation of a typology of learning and/or processing abilities. Indeed, standardization of the WIAT was undertaken conjointly with the WISC-III specifically to enable computation of ability-achievement discrepancies to aid in the identification of learning difficulties. Second, previous cluster analytic research has demonstrated that meaningful subgroups of distinct processing abilities can be identified in the evaluation of WISC-III subtest scores.

Commonly used measures of neuropsychological abilities were used to evaluate the external validity of the WISC-III-WIAT typology. Among the goals of neuropsychological assessment, perhaps the most important is the identification of patterns of strengths and weaknesses, which may be applied toward the design and implementation of individualized interventions. It has been suggested that fundamental to the accomplishment of this goal, a thorough sampling of skills and abilities thought to be subserved by the cerebral hemispheres should be undertaken (Rourke, Bakker, Fisk, & Strang, 1983). In this vein, research has demonstrated that measures of psychomotor abilities, tactile-kinesthetic-perceptual skills, visual-perceptual/visual-spatial skills, language skills, and the integration and synthesis of information are important ability domains in the subtyping of children with learning difficulties (Fisk & Rourke, 1983). Furthermore, the Halstead-Reitan Neuropsychological Test Battery is the most widely used measure from which to infer neuropsychological functioning (Dean, 1985). The factor analytic work of Batchelor, Sowles, Dean, and Fischer (1991) has demonstrated the validity of the HRNB-C in the assessment of learning disabled children. Based on factor analytic research with the HRNB-C and WISC/WISC-R (D'Amato, Gray, & Dean, 1988; Francis, Fletcher, Rourke, & York, 1992; Klonoff, 1971; Krug, Dean, & Anderson, 1995; Livingston, Gray, Haak, & Jenning, 1997) and current conceptualizations about the skills being assessed (Brown, Rourke, & Cicchetti, 1989; Reitan & Wolfson, 1992; Rourke et al.), the following neuropsychological tests, which were categorized into primary ability domains, were selected to evaluate the external validity of the WISC-III-WIAT typology:

i) Auditory-Linguistic Domain (Language)
    Auditory Closure
    Verbal Fluency
    Sentence Memory Test
    Speech-Sounds Perception Test (SSPT)

ii) Attention and Concentration Domain (Attention)
    Trail Making Test, Parts A and B (Trails A, and Trails B)
    Target Test

iii) Psychomotor Speed and Coordination Domain (Motor)
    Finger Tapping Test (Dominant, NonDominant)
    Grooved Pegboard (Dominant, NonDominant)

iv) Nonverbal Problem-Solving/Concept Formation (Problem-Solving)
Tactual Performance Test (TPT; total completion time)
Halstead Category Test (Total error score)

Analyses

Prior to analyses, all neuropsychological test scores were converted to *T*-scores (a mean of 50 and a standard deviation of 10) based on established age referenced normative data (Knights & Norwood, 1980). To minimize the influence of extreme scores on data analyses, those *T*-scores in excess of 4 standard deviations below the mean (i.e., *T*-scores less than 10) were converted to a *T*-score of 10, and retained for further analyses. In addition, to enable cluster analysis of WISC-III and WIAT subtest scores conjointly, the WIAT standard scores were converted to scaled scores ($M = 10$, $SD = 3$) to reflect an equivalent metric.

Following these transformations, WISC-III and WIAT subtest scores were subjected to a two-stage procedure including hierarchical and iterative partitioning cluster analyses in an attempt to identify a clinically meaningful ability-achievement taxonomy. First, a Ward's analysis with squared Euclidian distance as the similarity measure was applied to the data to estimate the number of clusters present in the sample. Ward's analysis was chosen as it has been shown to be among the best performing hierarchical clustering algorithms (Milligan & Cooper, 1987; Overall, Gibson, & Novy, 1993), and it has been used successfully in previous WISC-III taxonomic research (Donders, 1996; Glutting et al., 1994; Saunders et al., 2001). Similarly, squared Euclidian distance was selected because it is known to be sensitive to profile elevation *and* pattern (Aldenderfer & Blashfield, 1984; Donders; Morris & Fletcher, 1988), both of which were important considerations in the present work. Examination of the agglomeration coefficients and dendogram generated by the Ward's analysis strongly suggested the presence of five, six or eight clusters.

Because there exists a wide range of hierarchical clustering procedures and similarity measures which may consequently generate disparate solutions when applied to the same data (Fuerst, Fisk, & Rourke, 1989; Morris et al., 1981), the first step was to establish the internal validity of the WISC-III-WIAT taxonomy through demonstrating the replicability of the derived solutions. As such, three additional hierarchical analyses were performed to enable comparison of solutions derived through different clustering methods. The hierarchical clustering methods used were; Ward's analysis, complete linkage, average linkage – within groups, and average linkage – between groups (UPGMA). Based on the initial Ward's analysis, five, six, and eight cluster solutions were generated for each method, and the resulting mean profiles were examined for interpretability.

Because hierarchical clustering methods cannot reassign subjects who may have been improperly assigned earlier in the analysis, a *k*-means analysis was applied to each of the clustering results to correct for possible fusion errors and allow for the reassignment of subjects to more appropriate clusters. This was accomplished by using cluster centroids from each of the first-stage hierarchical cluster analyses as initial seeds in the *k*-means analyses. After each hierarchical solution had been subjected to *k*-means analysis, the resulting cluster solutions were compared based on the percentage of subjects reassigned to new clusters, and kappa, a nominal measure of association that corrects for chance agreement. Kappa was selected as the primary measure of association due to its having been used in previous taxonomic research. It has also been shown to correlate well with the Rand statistic, with values similarly ranging from 0.00 to 1.00 (Milligan, 1981; Morey & Agresti, 1984). Kappa values were calculated using the SPSS Crosstabs procedure. However, prior to comparing solutions through Crosstabs, cluster labels based on final cluster centres were reassigned such that cluster profiles obtained using one method, matched as closely as possible, profiles obtained using another method (e.g., labels were reassigned such that group two from one solution most closely matched group two in the comparison solution).

As a second measure of reliability, correlations were calculated between the mean WISC-III-WIAT subtypes generated by each clustering method. The four hierarchical solutions were compared separately for five-, six-, and eight-cluster solutions. The most reliable solution was selected as being representative of the data and it was used in all subsequent analyses. The resultant clusters were then visually matched to the closest corresponding clusters of Saunders et al. (2001), and compared through correlation coefficients.

The external validity of the cluster solution was examined using multivariate analysis of variance (MANOVA), through which the clusters were compared on a selection of standardized neuropsychological measures known to be sensitive to cerebral dysfunction. Indeed, to be meaningful, external validity may only be established by demonstrating that the derived subgroups can be differentiated on variables that were not used in the initial formation of the groups (Morris & Fletcher, 1988). These analyses were performed using individual neuropsychological test scores, domain composites (mean *T*-score across measures within each neuropsychological domain), and a global composite of neuropsychological functioning (mean *T*-score across neuropsychological domains). Composite scores were analyzed in addition to individual test scores due to differences in the stability, sensitivity and discriminability among neuropsychological measures (Brown et al., 1989; Davis, Adams, Gates, & Cheramie, 1989; Francis et al., 1992; Leckliter, Forster, Klonoff, & Knights, 1992). In the case of the motor domain, the composite score was obtained by calculating the mean performance of the dominant (Dom) and nondominant (NDom) hands on both the finger tapping and grooved pegboard tests. Bonferroni adjustment for multiple comparisons was applied to post-hoc analyses of group differences. All data screening and statistical analyses were performed using SPSS for Windows, Version 10.0.

RESULTS

Data Screening:

Prior to analyses, test scores for the 183 clinic-referred children were evaluated for accuracy of data entry, missing data, and multivariate outliers. Because the present sample was thought to be characterized by heterogeneity, univariate outliers, often defined as those subjects with *z*-scores in excess of 3.29 ($p<.001$), were considered part of the target population and retained for further analyses. Additionally, the effect of extreme scores on analyses was minimized through *T*-score transformations (see above). Multivariate outliers were evaluated with respect to Mahalanobis distance through modification to the syntax in the SPSS regression module as recommended by Tabachnick and Fidell (1996). Using this procedure, no multivariate outliers were identified. However, upon running the first series of cluster analyses, it became evident that one subject (a 9-year-old, left-handed male), who formed an isolated cluster (n = 1), was likely not part of the target sample. Based on procedures described in previous cluster analytic research (e.g., Morris et al., 1981), this subject was deemed an outlier and excluded from the study. Cluster analyses were rerun following removal of the subject, and all results were based on this smaller sample (n = 182).

Cluster Analyses:

An initial Ward's analysis was run to assess the number of clusters that would best represent the sample. Examination of the dendogram and agglomerative schedule suggested the presence of five, six, or eight clusters. Individual cluster analyses using each of the hierarchical methods (i.e., Ward's, complete linkage, UPGMA and average linkage – within groups) were applied to the data specifying solutions of five, six, and eight clusters. A *k*-means relocation pass was applied to the first stage cluster centroids from each solution. Each of the four hierarchical methods was then compared separately for five-, six-, and eight-cluster solutions.

Internal Validation: Comparison of Cluster Solutions

General comparison of the final *k*-corrected solutions derived through each of the hierarchical methods revealed that Ward's analysis resulted in the fewest number of subjects being reassigned to other clusters (approximately 15 percent across 5-, 6-, and 8-cluster solutions). A greater number of subjects were reassigned with the complete linkage and average linkage – within groups methods (approximately 34 and 33 percent, respectively). The UPGMA method tended to assign a disproportionate number of subjects to a single cluster, with several subjects being reassigned following the *k*-means procedure (approximately 40 percent).

For the five-cluster solution, good agreement was obtained for three of the four hierarchical methods (complete linkage, average linkage – within groups, and UPGMA). The percentage of subjects assigned to the same clusters across methods, and kappa values for each comparison are shown in Table 1. For a six-cluster solution, there was also good agreement between three of the four methods; however, the strength of the associations was slightly weaker than those obtained for the five-cluster solution. These comparisons are shown in Table 2. Finally, for an eight-cluster solution, overall agreement across methods was not obtained. In fact, the cluster solutions derived from each method were so varied that solutions could not be adequately matched for comparison. In addition, many of the clusters were difficult to interpret clinically. Although a subjective task, clinical interpretability is an important consideration in taxonomic research (Kamphaus, Distefano, & Lease, 2003; Morris et al., 1981). Therefore, based on the poor replicability and clinical interpretability of the eight-cluster

Table 1

Kappa Values and Percent Agreement between all Four Hierarchical Methods for a Five-Cluster Solution

Method	Wards	Average Linkage – within groups	Complete Linkage	UPGMA
Wards	---	---	---	---
Average Linkage – Within groups	.436 54.9%	---	---	---
Complete Linkage	.502 60.4%	.883 90.6%	---	---
UPGMA	.443 55.5%	.897 91.7%	.918 93.4%	---

All values are significant at p< .001.

Table 2

Kappa Values and Percent Agreement between all Four Hierarchical Methods for a Six-Cluster Solution

Method	Wards	Average Linkage – within groups	Complete Linkage	UPGMA
Wards	---	---	---	---
Average Linkage – Within groups	.735 78.0%	---	---	---
Complete Linkage	.749 79.1%	.649 70.9%	---	---
UPGMA	.630 69.2%	.636 69.8%	.490 57.7%	---

All values are significant at *p*< .001.

solution, and the previously obtained agreement between three of the four hierarchical methods for both five- and six-cluster solutions, the eight-cluster solution was eliminated from the remainder of the analyses.

As a second measure of agreement between the different cluster solutions, correlations were calculated between mean WISC-III-WIAT profiles for each of the subtypes derived through the four hierarchical methods. This was performed for both five- and six-cluster solutions; the coefficients are presented in Tables 3 and 4, respectively.

For a five-cluster solution, all subtypes derived by the four hierarchical methods were significantly correlated with the exception of the comparison between Ward's and average linkage-within groups on one subtype. Upon closer inspection of the five-cluster solution, the complete linkage method demonstrated the highest correlations with each of the other three hierarchical methods, with correlations ranging from .602 to 1.00. Similarly, all subtypes derived from the six-cluster solution correlated .605 or better, with the exception of the comparison between the complete linkage and UPGMA methods on one subtype. Of the four hierarchical methods within the six-cluster solution, Ward's demonstrated the highest correlations across subtypes ranging from .662 to .972. Collectively, these results indicate that all four hierarchical methods produced subtypes with similar WISC-III-WIAT profiles for both five- and six-cluster solutions.

Ultimately, the complete linkage five-cluster solution was selected as the standard and was used in subsequent analyses. The complete linkage five-cluster solution was chosen over the Ward's six-cluster solution because it demonstrated the greatest correspondence with each of the comparison methods, and the resultant mean WISC-III-WIAT profiles appeared to be clinically meaningful. Furthermore, even though

Table 3

Five-Cluster Solution: Correlations between Mean WISC-III-WIAT Subtypes Derived through Each of the Four Hierarchical Methods

Cluster Number	Method	Wards	Average Linkage – within groups	Complete Linkage	UPGMA
1	Wards	---	---	---	---
2					
3					
4					
5					
1	Average Linkage – Within groups	.430	---	---	---
2		.971**			
3		.739**			
4		.925**			
5		.831**			
1	Complete Linkage	.602*	.939**	---	---
2		.950**	.984**		
3		.783**	.987**		
4		.874**	.967**		
5		.831**	.944**		
1	UPGMA	.602*	.939**	1.00**	---
2		.970**	.951**	.959**	
3		.702*	.984**	.969**	
4		.907**	.957**	.951**	
5		.830**	.916**	.973**	

* $p<.05$; **$p<.01$

Table 4

Six-Cluster Solution: Correlations between Mean WISC-III-WIAT Subtypes Derived through Each of the Four Hierarchical Methods

Cluster Number	Method	Wards	Average Linkage – within groups	Complete Linkage	UPGMA
1	Wards	---	---	---	---
2					
3					
4					
5					
6					
1	Average Linkage – Within groups	.972**	---	---	---
2		.937**			
3		.942**			
4		.829**			
5		.924**			
6		.966**			
1	Complete Linkage	.968**	.971**	---	---
2		.934**	.970**		
3		.908**	.801**		
4		.851**	.608*		
5		.909**	.849**		
6		.982**	.957**		
1	UPGMA	.968**	.971**	1.00**	---
2		.836**	.771**	.768**	
3		.908**	.971**	.768**	
4		.784**	.735**	.605*	
5		.662**	.735**	.481	
6		.967**	.975**	.971**	

* $p<.05$; **$p<.01$

one subtype was not highly correlated between two methods, it has been suggested that while a valid clustering solution should reappear with different clustering methods, the exact results need not be replicated under all types of cluster analyses (Morris et al., 1981; Morris et al., 1998).

WISC-III-WIAT Profiles:

Mean WISC-III-WIAT scores for each of the five complete linkage subtypes were calculated to obtain the profiles presented in Figures 1 through 5. Figure 6 illustrates the similarities and differences among all five profiles. Means and standard deviations of the WISC-III and WIAT subtest and IQ scores for each of the five subtypes are presented in Table 5. In addition, gender and age distributions for each subtype are summarized in Table 6. There were no significant differences in mean age based on cluster membership [$F(4, 177) = .957$, $p = .432$].

Descriptive labels were assigned to the five subgroups based on the most salient features of each profile. The first group (n= 35) was characterized by predominantly below average scores across WISC-III and WIAT subtests, excluding Picture Completion and Objection Assembly from the WISC-III, and was thus called Low Ability. Group two (n=40), the largest of the five subtypes, was defined by performance levels ranging from below average to low average on all WISC-III and WIAT measures, with the exception of those subtests which comprise the WISC-III PSI (i.e., Coding and Symbol Search), which were within the average range of ability. This group was therefore labeled Low Ability with Average Processing Speed (Avg. PS). The third group (n=39) demonstrated broadly average abilities across WIAT subtests and WISC-III verbal and nonverbal measures, with the exception of Coding, which was just below average. Of note, measures of nonverbal abilities were generally lower than measures of verbal ability. This profile was called Low Visual Spatial/Processing Speed (NPD).

Figure 1

Group 1: Low Ability

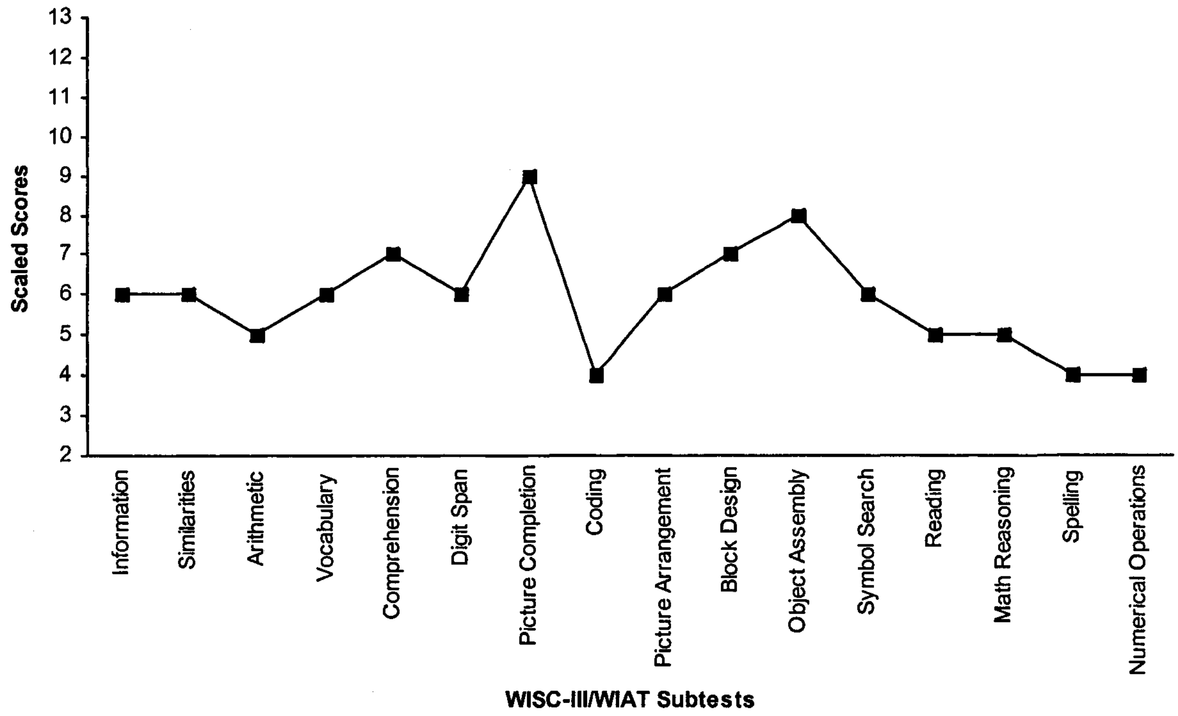


Figure 2

Group 2: Low Ability with Average Processing Speed (Avg. PS)

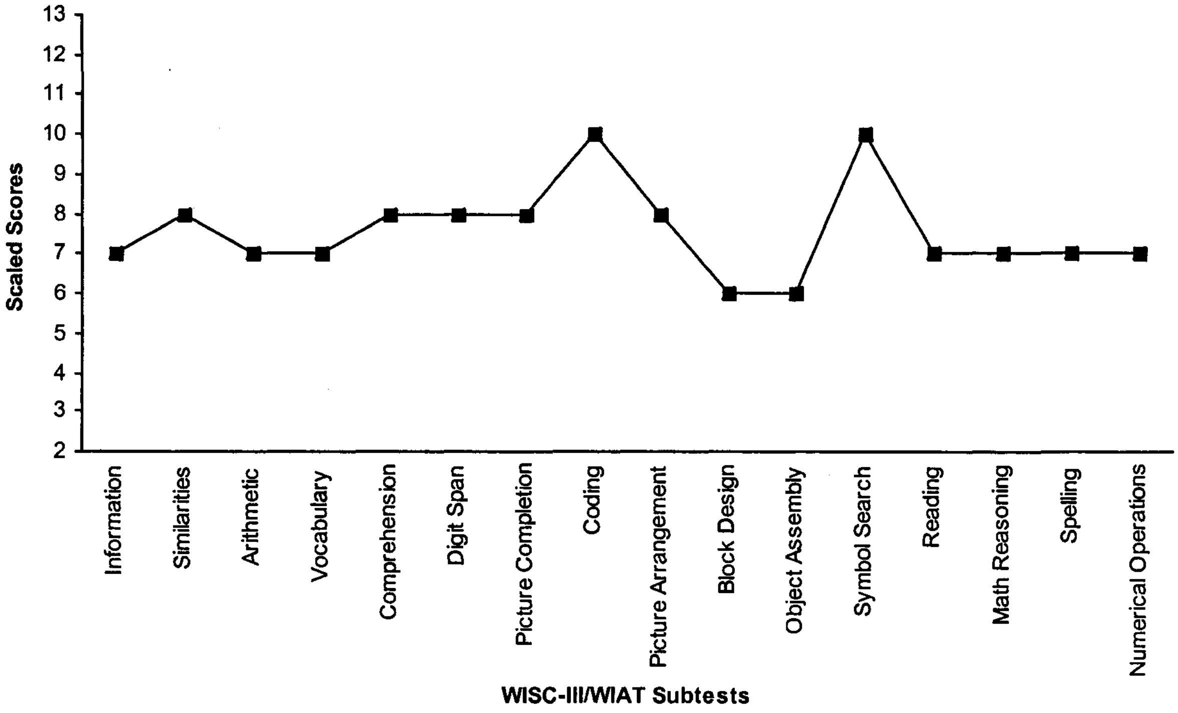


Figure 3

Group 3: Low Visual Spatial/Processing Speed Skills (NPD)

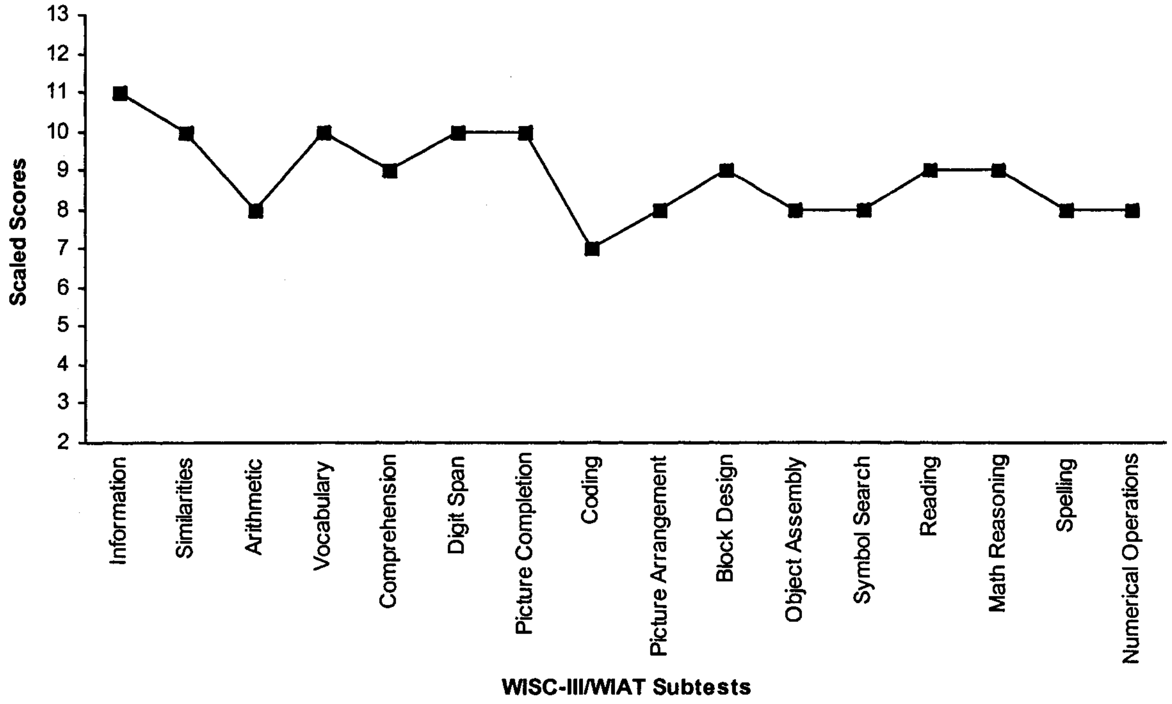


Figure 4

Group 4: ACID

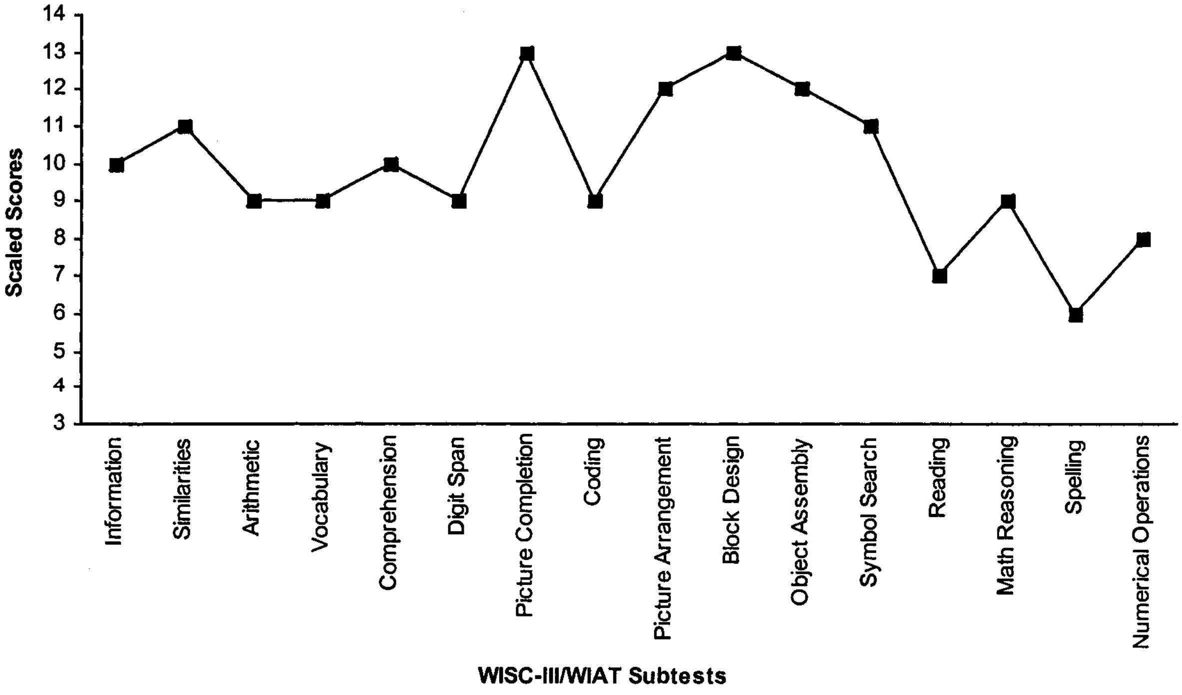


Figure 5

Group 5: Verbal Processing Deficits (VPD)

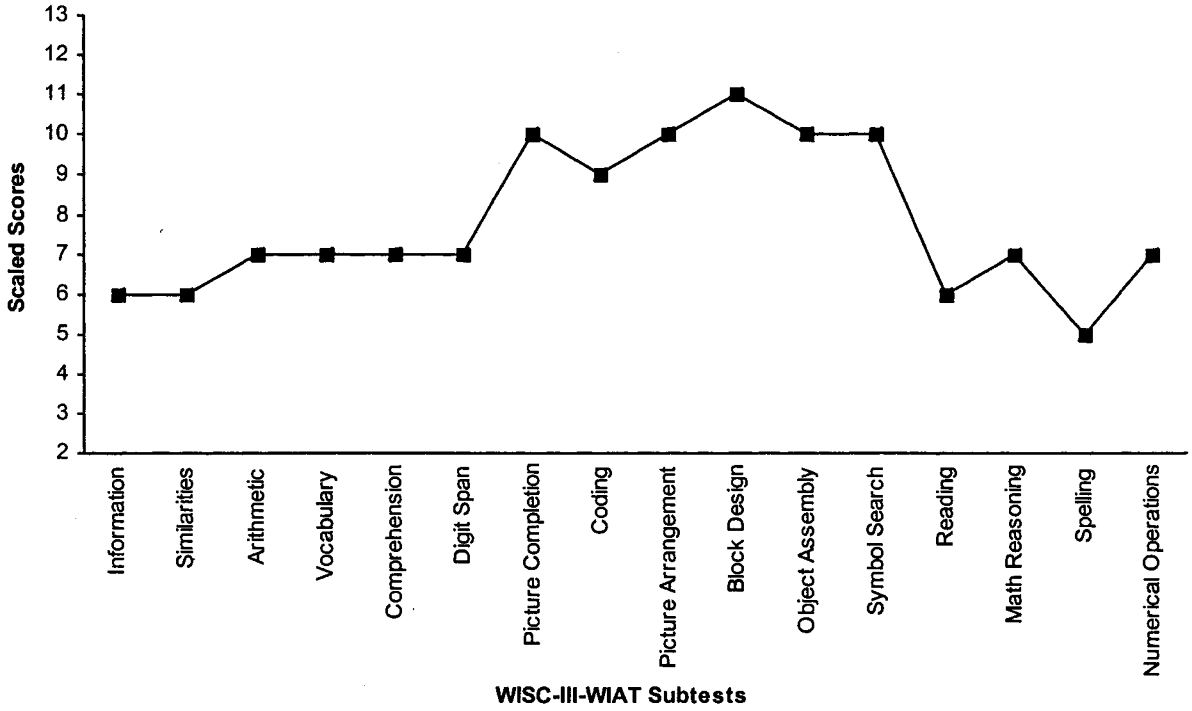


Figure 6

Five-cluster solution

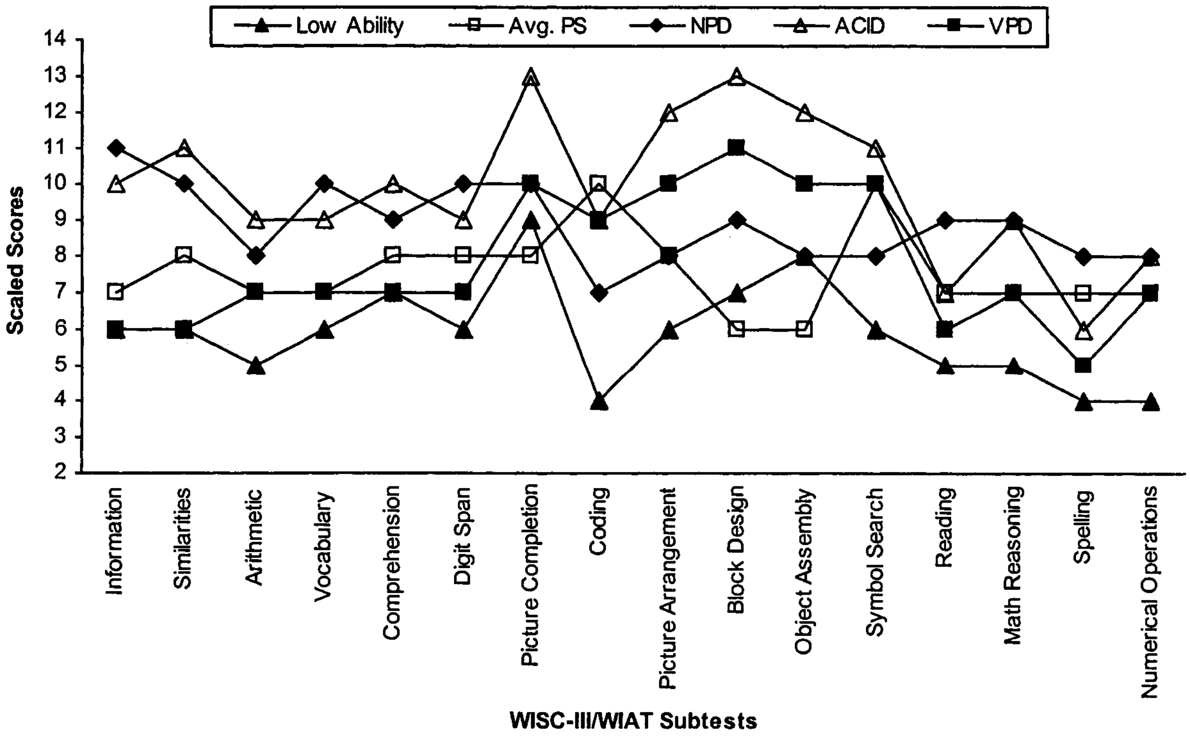


Table 5

WISC-III-WIAT Subtest Means and Standard Deviations for the Complete Linkage Five-Cluster Solution

Subtest	Low Ability (n=35)		Avg. PS (n=40)		NPD (n=39)		ACID (n=31)		VPD (n=37)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
WISC-III										
Information	6.23	(2.18)	6.90	(1.91)	10.67	(2.45)	9.87	(2.19)	6.24	(2.31)
Similarities	6.46	(2.41)	7.62	(1.89)	10.41	(2.42)	10.74	(2.13)	6.22	(1.77)
Arithmetic	5.23	(1.72)	6.83	(1.91)	8.46	(2.44)	8.61	(2.35)	7.24	(1.77)
Vocabulary	6.29	(1.54)	7.10	(2.18)	9.69	(2.19)	9.39	(2.38)	6.54	(1.97)
Comprehension	6.63	(2.18)	7.75	(2.33)	9.28	(1.99)	10.42	(1.88)	7.19	(2.31)
Digit Span	5.89	(2.21)	7.65	(2.23)	9.54	(2.80)	9.03	(2.07)	7.24	(2.55)
Picture Completion	9.09	(2.37)	7.85	(2.32)	9.87	(2.83)	13.39	(2.22)	9.86	(2.19)
Coding	4.46	(1.93)	9.57	(2.43)	6.90	(2.57)	8.77	(2.57)	8.65	(2.43)
Picture Arrangement	6.20	(2.52)	7.58	(3.02)	8.18	(2.82)	11.94	(2.42)	10.16	(2.39)
Block Design	6.80	(2.90)	6.00	(2.24)	9.08	(2.21)	12.52	(2.58)	10.62	(1.66)
Object Assembly	8.31	(2.23)	6.33	(2.29)	8.38	(2.36)	12.32	(2.30)	10.30	(2.13)
Symbol Search	5.71	(2.31)	9.80	(2.04)	8.46	(2.39)	11.42	(2.31)	10.32	(2.24)
VIQ	78.69	(6.56)	84.48	(6.75)	98.51	(8.16)	99.00	(7.69)	81.49	(6.98)
PIQ	81.31	(7.19)	84.20	(7.29)	90.51	(7.67)	112.23	(7.12)	99.92	(5.92)
FSIQ	78.17	(5.82)	82.80	(5.71)	94.13	(6.47)	105.58	(6.91)	89.27	(5.74)
WIAT										
Reading	5.09	(2.09)	6.70	(1.83)	9.05	(2.36)	7.29	(2.18)	5.57	(1.17)
Math Reasoning	5.37	(1.80)	6.55	(1.45)	8.51	(1.99)	9.06	(1.65)	6.78	(1.72)
Spelling	4.37	(1.65)	6.73	(1.91)	8.10	(2.15)	6.29	(2.04)	4.68	(1.16)
Numerical Operations	4.23	(1.37)	6.68	(1.73)	8.23	(2.10)	4.68	(1.16)	6.68	(2.04)

Table 6

Age and Gender Distribution for the Complete Linkage Five-Cluster Solution

	Gender		Age							
	M	F	9	10	11	12	13	14	*M*	*SD*
Subtype										
Low Ability	28	7	7	7	14	3	2	2	10.77	(1.35)
Avg. PS	23	17	12	13	6	5	1	3	10.47	(1.48)
NPD	37	2	7	12	6	4	9	1	10.97	(1.53)
ACID	26	5	10	9	6	1	4	1	10.45	(1.48)
VLD	30	7	9	10	11	5	1	1	10.64	(1.42)
Total	144	38	45	57	43	18	17	8	---	---

Group four (n=31), the smallest of the five groups, demonstrated an ACID profile (comparatively reduced scores on Information, Arithmetic, Digit Span, and Coding) with low scores on all WIAT measures of academic achievement (particularly reading and spelling). This subtype was therefore called ACID. The fifth group (n=37), labeled Verbal Processing Deficits (VPD), exhibited a clear pattern of depressed WISC-III verbal and WIAT academic scores in comparison to uniformly average scores on measures of nonverbal/visual-spatial skills.

Upon closer inspection of Figures 1 through 6 it is evident that each of the subgroups is distinct from one another, some of which appear to be differentiated by pattern and others which are primarily differentiated by elevation. Specifically, at least two clusters appeared to share a similar pattern with different levels of performance discriminating between the groups. To further explore the relationship between the five clusters, correlations were calculated between each of the cluster profiles both with and without the inclusion of WIAT subtest scores, and are presented in Table 6 and Table 7, respectively. The ACID profile was found to be highly positively correlated with both the Low Ability and VPD profiles regardless of whether or not the WIAT subtests were included. Conversely, although the Low Ability and VPD groups were modestly correlated when the WIAT subtests were included, the two groups were not significantly correlated when the WIAT subtests were removed from the calculations.

Relationship to Subtypes of Saunders et al. (2001)

It was anticipated that WISC-III subtypes similar to those identified by Saunders et al. (2001) would be obtained in the present research using WISC-III and WIAT subtests. Visual inspection of Figures 1 through 5 suggested that at least three of the derived subtypes closely resembled subtypes obtained by Saunders et al..

Table 7

Correlations between the Five Subtypes, Including WIAT Subtests

Cluster	Low Ability (n=35)	Avg. PS (n=40)	NPD (n=39)	ACID (n=31)	VPD (n=37)
Low Ability	---	---	---	---	---
Avg. PS	-.159	---	---	---	---
NPD	.424	-.259	---	---	---
ACID	.811**	.032	.174	---	---
VPD	.548*	.176	-.307	.802**	---

* *p*<.05; ***p*<.01

Table 8

Correlations between the Five Subtypes, Excluding WIAT Subtests

Cluster	Low Ability (n=35)	Avg. PS (n=40)	NPD (n=39)	ACID (n=31)	VPD (n=37)
Low Ability	---	---	---	---	---
Avg. PS	-.433	---	---	---	---
NPD	.347	-.351	---	---	---
ACID	.732**	-.272	.000	---	---
VPD	.369	.039	-.537	.714**	---

**p<.01

To further examine the relationship between the present cluster solution and that of Saunders et al., the WISC-III subtest scores of visually corresponding profiles were compared through correlation coefficients.

The Saunders et al. (2001) group labeled Deficits Consistent with an ACID Pattern was replicated with excellent accuracy in the current study (*r* = .95). The mean WISC-III profiles for these two subtypes are presented in Figure 7. The current study also replicated the Saunders et al. Deficient Language Abilities group with good accuracy (*r* = .82). The mean WISC-III profiles for these two subtypes are presented in Figure 8. The current study obtained a subtype that was highly similar to the Broad-Based Processing Deficiencies group of Saunders et al. (*r* = .70). These mean WISC-III profiles are presented in Figure 9. A relatively similar pattern to that of the Deficient Nonverbal Abilities group of Saunders et al. was obtained in the present study (*r* = .61). The mean WISC-III profiles for these two subtypes are presented in Figure 10.

Of the five subtypes identified in the current study, only the Avg. PS group could not be matched to one of the subtypes of Saunders et al (2001). Similarly, not all of the subtypes identified by Saunders et al. were replicated in the present study.

External Validation: Differences between Subtypes on Neuropsychological Measures

The neuropsychological measures used for evaluation of the external validity of the cluster solution reflect a selection of those tests which are commonly utilized in the neuropsychological assessment of children. Neuropsychological test scores were converted to normalized *T*-scores, and all five clusters were compared on individual neuropsychological tests and domain and global composite scores. To examine the relationship among neuropsychological measures, correlations were computed between test scores within each domain and between domain composite scores. These results

Figure 7

Mean WISC-III Profiles for the ACID and Deficits Consistent with an ACID Pattern Subtypes

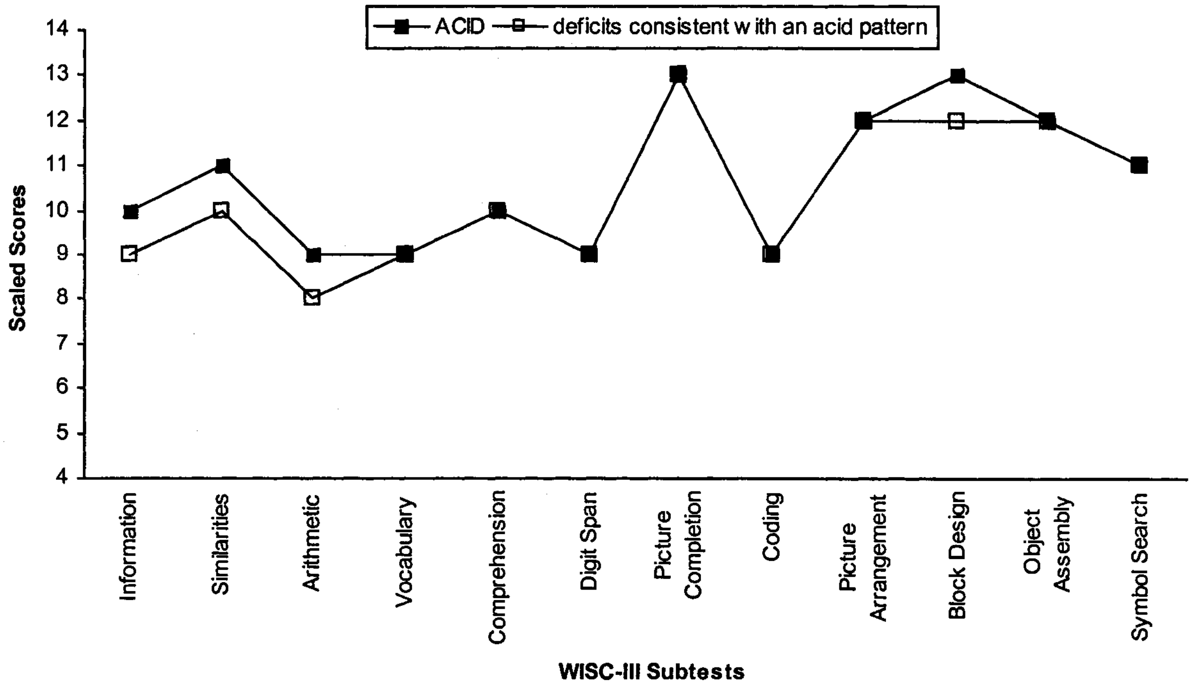


Figure 8

Mean WISC-III Profiles for the Verbal Processing Deficits and Deficient Verbal Abilities Subtypes

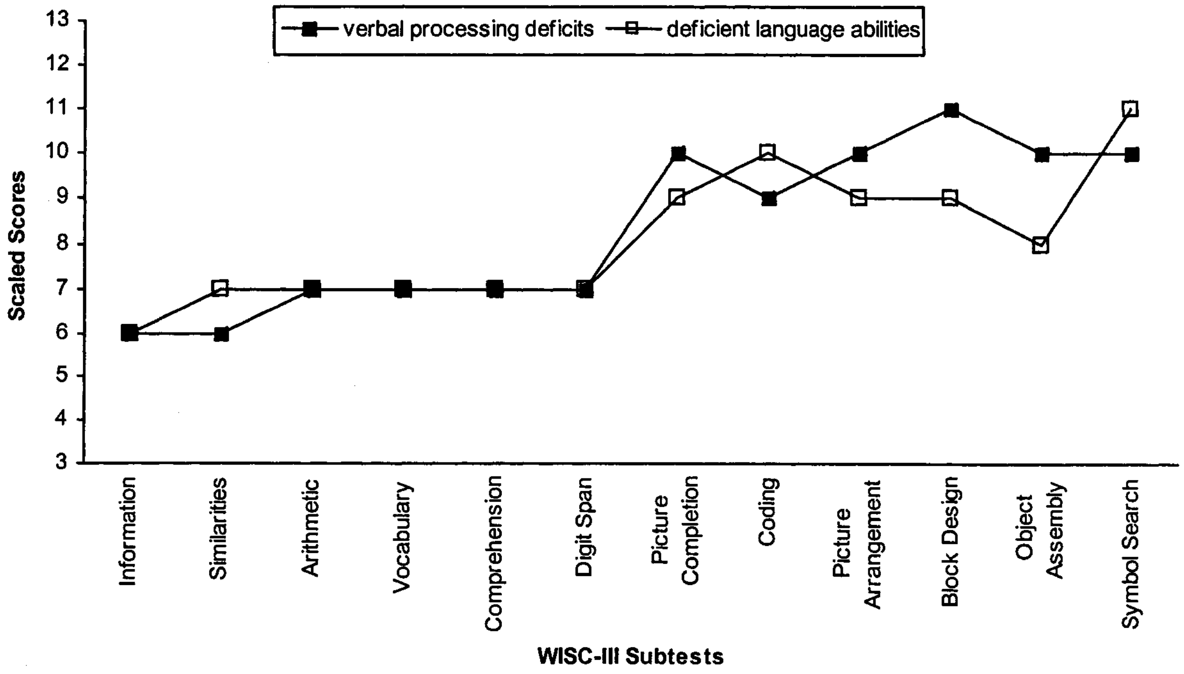


Figure 9

Mean WISC-III Profiles for the Low Ability and Broad-Based Processing Deficiencies Subtypes

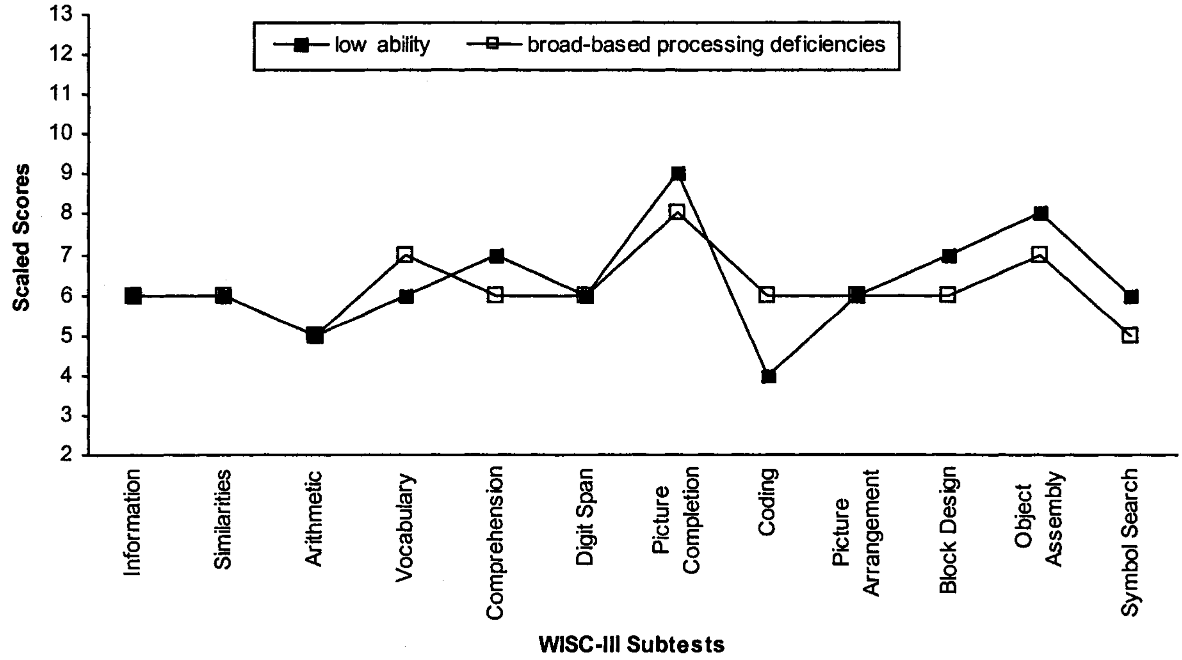
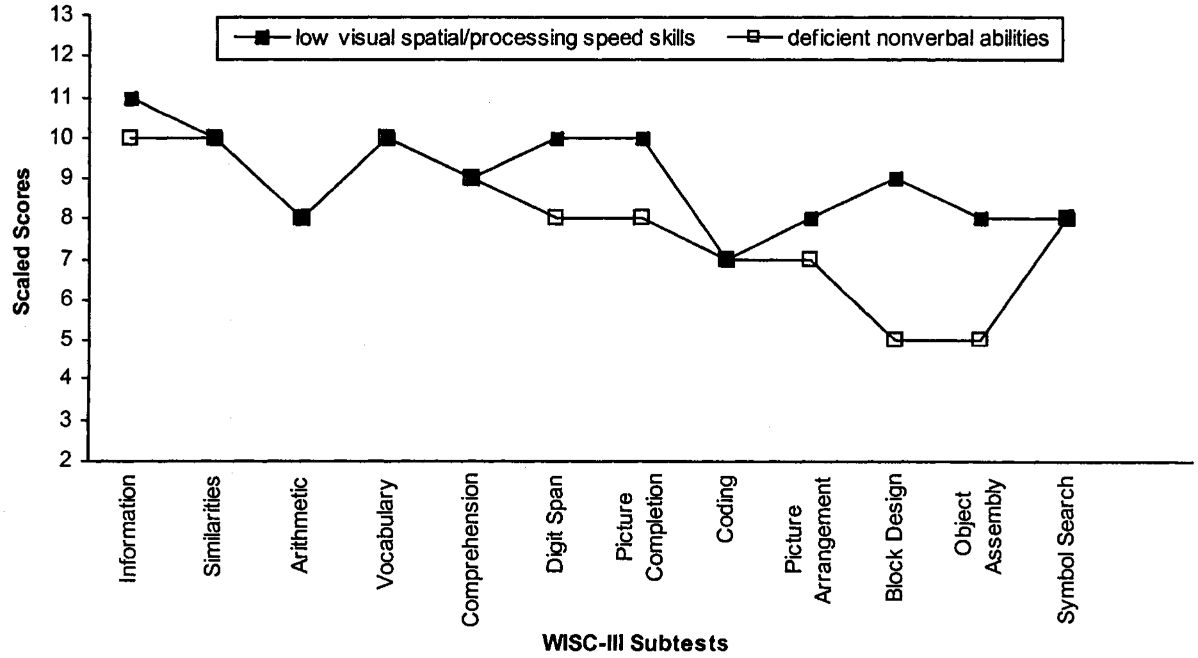


Figure 10

Mean WISC-III Profiles for the Low Visual Spatial/Processing Speed Skills and Deficient Nonverbal Abilities Subtypes



may be found in Tables 9 and 10, respectively. Of note, although each of the comparisons between individual tests and domains revealed statistically significant correlations, in many cases, the proportion of variance accounted for was quite small.

Subtype means and standard deviations for each of the neuropsychological measures, including domains and the global composite, are presented in Table 11. An ANOVA revealed significant differences between the subtypes on the global composite [$F(4, 177) = 27.94$, $p < .001$]. Global composite scores for each subtype are presented in Figure 11. Post-hoc comparisons indicated that the Low Ability group obtained a significantly lower score on the global composite relative to each of the other groups ($p<.05$), and the ACID group obtained a significantly higher score on the global composite relative to each of the other groups ($p<.05$). None of the other groups were significantly different from one another on the global composite.

A MANOVA, evaluating domain composite scores, also indicated significant differences in neuropsychological performance based on subtype membership [$F(16, 532) = 9.60$, $p<.001$]. Separate univariate analyses for each of the neuropsychological measures also indicated the presence of significant differences between the subtypes, with the exception of the Finger Tapping Test (dominant and nondominant hands). Corresponding *F* values and a summary of these results are presented in Tables 12 and 13, respectively. Similarities and differences between mean neuropsychological test, and domain composite scores for each of the WISC-III-WIAT subtypes are illustrated in Figures 12 through 16.

Post-hoc analyses, using a Bonferroni adjustment for multiple comparisons, revealed significant differences between many of the clusters on both individual test and domain composite measures. Overall, the Low Ability group demonstrated the lowest mean performance across most measures of neuropsychological functioning.

Table 9

Correlations between Measures within Neuropsycological Domains

	Measures												
Measures	1	2	3	4	5	6	7	8	9	10	11	12	13
Language Domain													
1. AUDCLO	1.00	.42**	.44**	.47**	.12	.15*	.06	.11	.13	-.01	.07	.20*	.08
2. FLUENCY	-	1.00	.18*	.27**	.27**	.20**	.15*	.09	.08	.07	.07	.09	.12
3. SENT MEM	-	-	1.00	.32**	.42**	.37**	.14	-.04	-.16	.01	.02	.13	.11
4. SSPT	-	-	-	1.00	.30**	.19**	.26**	.12	.20**	.16*	.23**	.32**	.12
Attention Domain													
5. TRAILS A	-	-	-	-	1.00	.41**	.16*	.19*	.26**	.20**	.15*	.15*	.11
6. TRAILS B	-	-	-	-	-	1.00	.30*	.28**	.22**	.21**	.20**	.24**	.29**
7. TARGET	-	-	-	-	-	-	1.00	.28**	.26**	.18**	.21**	.24**	.18*
Motor Domain													
8. FTT, Dom	-	-	-	-	-	-	-	1.00	.69**	.17*	.23**	.10	.14
9. FTT, NDom	-	-	-	-	-	-	-	-	1.00	.15*	.28**	.16*	.08
10. PEG, Dom	-	-	-	-	-	-	-	-	-	1.00	.74**	.24**	.30**
11. PEG, NDom	-	-	-	-	-	-	-	-	-	-	1.00	.19**	.21**
Problem Solving Domain													
12. CATEGORY	-	-	-	-	-	-	-	-	-	-	-	1.00	.25**
13. TPT	-	-	-	-	-	-	-	-	-	-	-	-	1.00

*$p < .05$; **$p < .01$

Table 9 (cont.)

Correlations between Neuropsychological Measures

Note: AUDCLO=Auditory Closure; FLUENCY= Verbal Fluency; SENT MEM=Sentence Memory Test; SSPT=Speech-Sounds Perception Test; TRAILS A=Trail Making Test, Part A; TRAILS B=Trail Making Test, Part B; TARGET=Target Test; FTT Dom=Finger Tapping Test, Dominant hand; FTT NDom=Finger Tapping Test, Nondominant hand; PEG Dom=Grooved Pegboard, Dominant hand; PEG NDom=Grooved Pegboard, Nondominant hand; CATEGORY=Category Test; TPT=Tactual Performance Test.

Table 10

Correlations between Neuropsychological Domains

	1	2	3	4
1. Language Domain	1.00	---	---	---
2. Attention Domain	.32**	1.00	---	---
3. Motor Domain	.17*	.32**	1.00	---
4. Problem Solving Domain	.26**	.35**	.31**	1.00

* *p*<.05; ** *p*<.01

Table 11

Means and Standard Deviations for each WISC-III-WIAT Subtype on Neuropsychological Measures

Measure	Low ability (n=35)		Avg. PS (n=40)		NPD (n=39)		ACID (n=31)		VPD (n=37)	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Language Domain										
Auditory Closure	38.37	(8.72)	43.03	(9.77)	45.92	(9.20)	48.03	(10.96)	39.78	(10.87)
Verbal Fluency	35.11	(9.98)	41.60	(11.39)	42.64	(10.48)	45.61	(12.34)	37.41	(8.98)
Sentence Memory	28.23	(12.78)	28.40	(12.83)	36.03	(12.27)	37.61	(14.13)	28.03	(10.22)
SSPT	22.89	(16.07)	30.05	(15.22)	41.64	(17.32)	44.94	(13.62)	32.43	(15.67)
Composite Score	31.15	(7.92)	35.77	(8.62)	41.56	(8.48)	44.05	(7.60)	34.41	(8.13)
Attention Domain										
Trails A	39.86	(12.01)	50.63	(7.22)	46.00	(11.76)	52.06	(8.63)	46.89	(9.90)
Trails B	33.94	(15.15)	44.33	(12.24)	46.33	(12.06)	50.65	(9.74)	46.97	(11.09)
Target Test	33.60	(13.49)	40.80	(12.66)	45.95	(12.13)	48.19	(9.74)	44.92	(12.23)
Composite Score	35.80	(10.11)	45.25	(7.39)	46.09	(7.67)	50.30	(7.41)	46.26	(5.98)
Motor Domain										
Finger Tapping Dom	49.11	(14.50)	56.40	(14.15)	50.49	(14.17)	55.71	(11.96)	53.76	(15.66)
Finger Tapping NDom	47.74	(12.02)	53.98	(13.29)	48.97	(11.29)	53.10	(9.34)	54.81	(13.26)
Pegboard Dom	39.69	(13.24)	44.63	(14.03)	44.10	(14.73)	55.77	(9.25)	52.16	(11.89)
Pegboard NDom	33.97	(17.49)	39.38	(15.31)	39.36	(15.82)	53.77	(9.41)	50.81	(12.64)
Composite Score	42.63	(10.47)	48.59	(9.94)	45.73	(10.89)	54.59	(6.37)	52.88	(8.90)
Problem Solving Domain										
TPT Total	39.71	(15.23)	41.50	(13.34)	46.59	(11.15)	53.35	(6.86)	49.78	(10.09)
Category Total	46.46	(9.08)	45.75	(8.97)	53.92	(9.04)	55.61	(8.58)	53.46	(9.69)
Composite Score	43.09	(8.73)	43.62	(8.45)	50.26	(7.95)	54.48	(5.63)	51.62	(7.81)
Global Composite	38.17	(5.32)	43.31	(4.81)	45.91	(5.96)	50.85	(3.81)	46.29	(5.09)

Figure 11

Mean Performance by Subtype on Global Composite

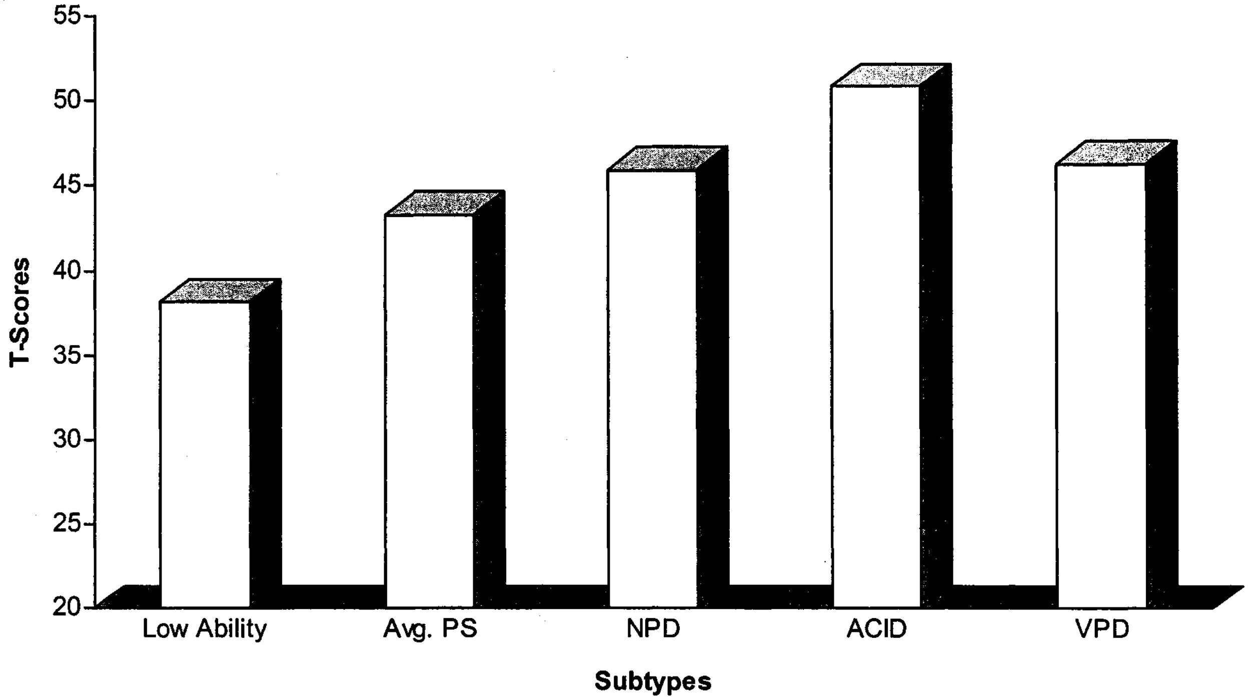


Table 12

F Values for Univariate Comparisons between the Five Subtypes on Neuropsychological Measures

Measure	F	p
Language Domain		
Auditory Closure	5.73	<.001
Verbal Fluency	5.30	<.001
Sentence Memory	5.10	.001
SSPT	11.18	<.001
Composite Score	14.32	<.001
Attention Domain		
Trails A	7.75	<.001
Trails B	9.07	<.001
Target Test	7.70	<.001
Composite Score	16.11	<.001
Motor Domain		
Finger Tapping Dom	1.84	.123
Finger Tapping NDom	2.53	.042
Pegboard Dom	8.68	<.001
Pegboard NDom	11.52	<.001
Composite Score	9.20	<.001
Problem Solving Domain		
TPT Total	7.93	<.001
Category Total	9.18	<.001
Composite Score	14.44	<.001
Global Composite	27.94	<.001

Table 11

Comparisons among the Five Subtypes on Neuropsychological Measures

Neuropsychological Domain	**Low Ability**	**Avg. PS**	**NPD**	**ACID**	**VPD**
Language Domain					
Auditory Closure	NPD *; ACID *	NS	Low Ability	Low Ability ; VPD	ACID *
Verbal Fluency	NPD *; ACID *	NS	Low Ability	Low Ability ; VPD	ACID *
Sentence Memory	ACID *	ACID	NS	Low Ability; Avg. PS; VPD	ACID *
SSPT	NPD *; ACID *	NPD *; ACID *	Low Ability ; Avg. PS	Low ability; Avg. PS; VPD	ACID *
Composite	NPD *; ACID *	NPD *; ACID *	Low Ability ; Avg. PS; VPD	Low ability; Avg. PS; VPD	NPD *; ACID *
Attention Domain					
Trails A	Avg. PS *; ACID *; VPD *	Low Ability	NS	Low Ability	Low Ability
Trails B	all subtypes *	Low Ability	Low Ability	Low Ability	Low Ability
Target	NPD *; ACID *; VPD *	NS	Low Ability	Low Ability	Low Ability
Composite	all subtypes *	Low Ability	Low Ability	Low Ability	Low Ability
Motor Domain					
Finger Tapping (Dom)	NS	NS	NS	NS	NS
Finger Tapping (Ndom)	NS	NS	NS	NS	NS
Pegboard (Dom)	ACID *; VPD *	ACID *	ACID *	Low Ability ; Avg. PS; NPD	Low Ability
Pegboard (Ndom)	ACID *; VPD *	ACID *; VPD *	ACID *; VPD *	Low Ability ; Avg. PS; NPD	Low Ability ; Avg. PS; NPD
Composite	ACID *; VPD *	NS	ACID *; VPD *	Low Ability ; NPD	Low Ability; NPD
Problem Solving Domain					
TPT(total time)	NPD *; ACID *; VPD *	ACID *; VPD *	NS	Low Ability ; Avg. PS	Low Ability ; Avg. PS
Category Test (total)	NPD *; ACID *; VPD *	NPD *; ACID *; VPD *	Low Ability ; Avg. PS	Low Ability ; Avg. PS	Low Ability ; Avg. PS
Composite	NPD *; ACID *; VPD *	NPD *; ACID *; VPD *	Low Ability ; Avg. PS	Low Ability ; Avg. PS	Low Ability ; Avg. PS
Global Composite	all subtypes *	Low Ability ; ACID *	Low Ability ; ACID *	all subtypes *	low ability; ACID *

* Denotes significantly better score relative to the comparison group; NS = Non significant

Figure 12.

Mean Performance by Subtype on Neuropsychological Domains and Global Composite

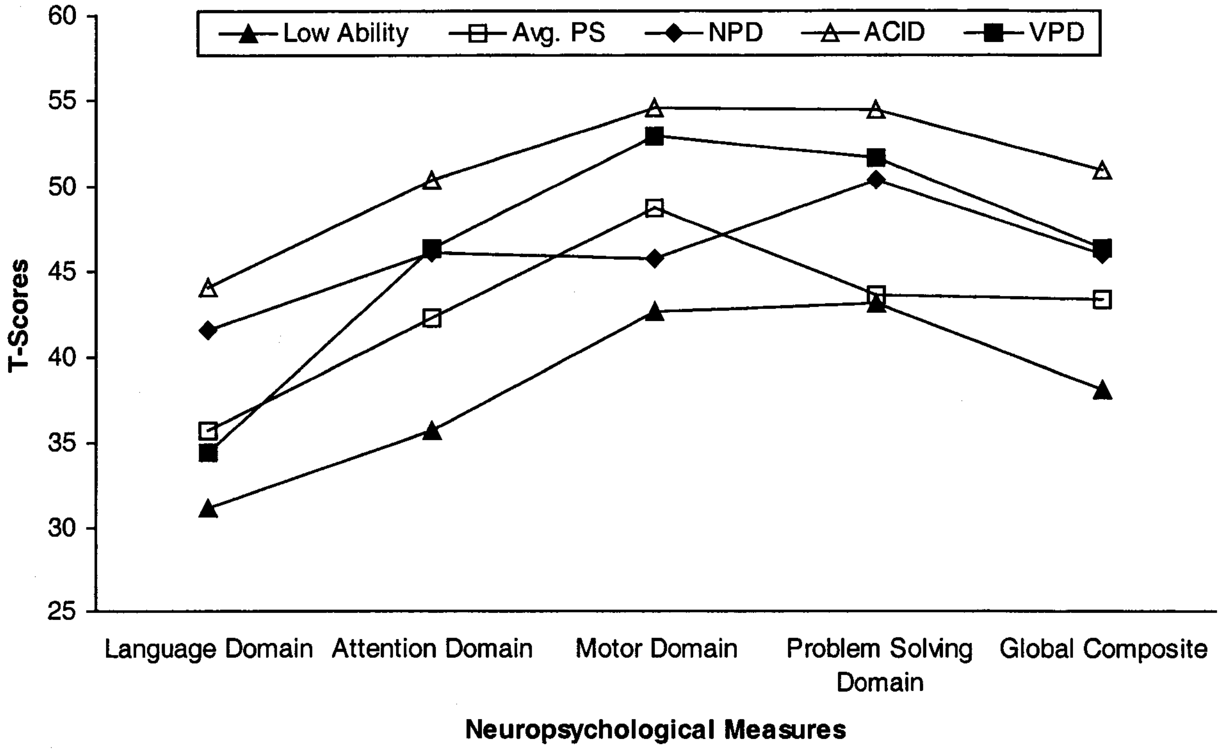


Figure 13

Mean Performance by Subtype on Measures within the Language Domain

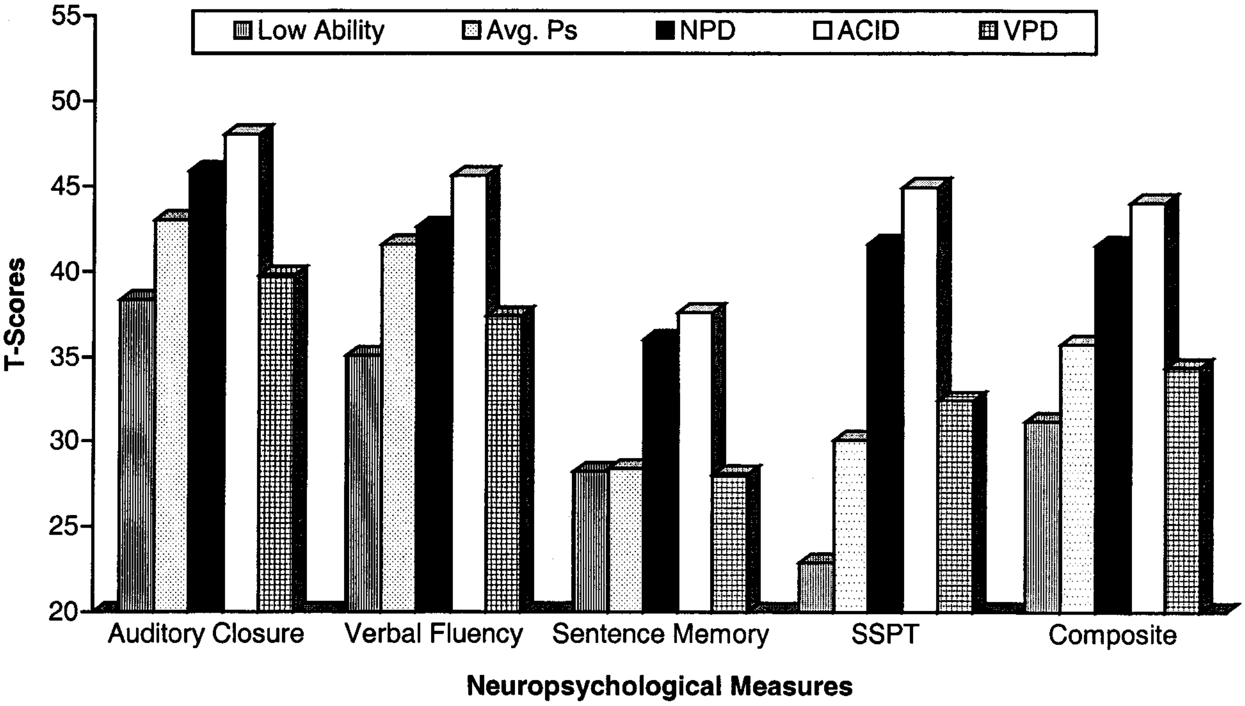


Figure 14.

Mean Performance by Subtype on Measures within the Attention Domain

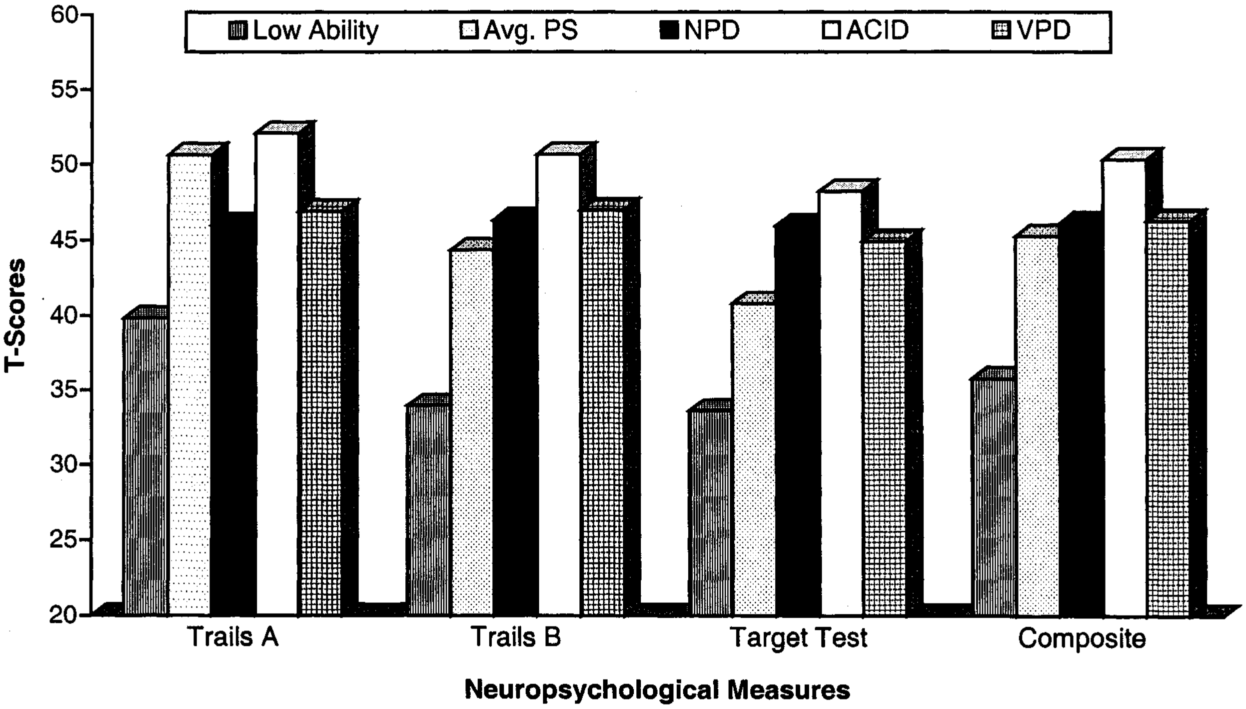


Figure 15.

Mean Performance by Subtype on Measures within the Motor Domain

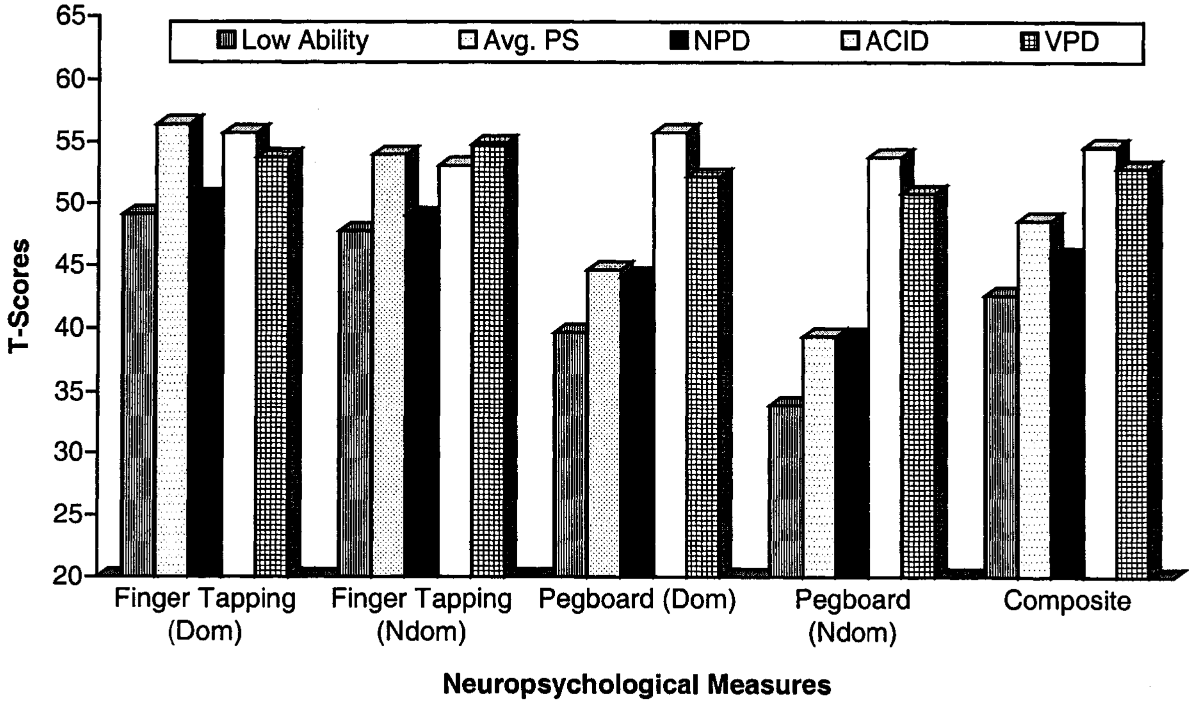
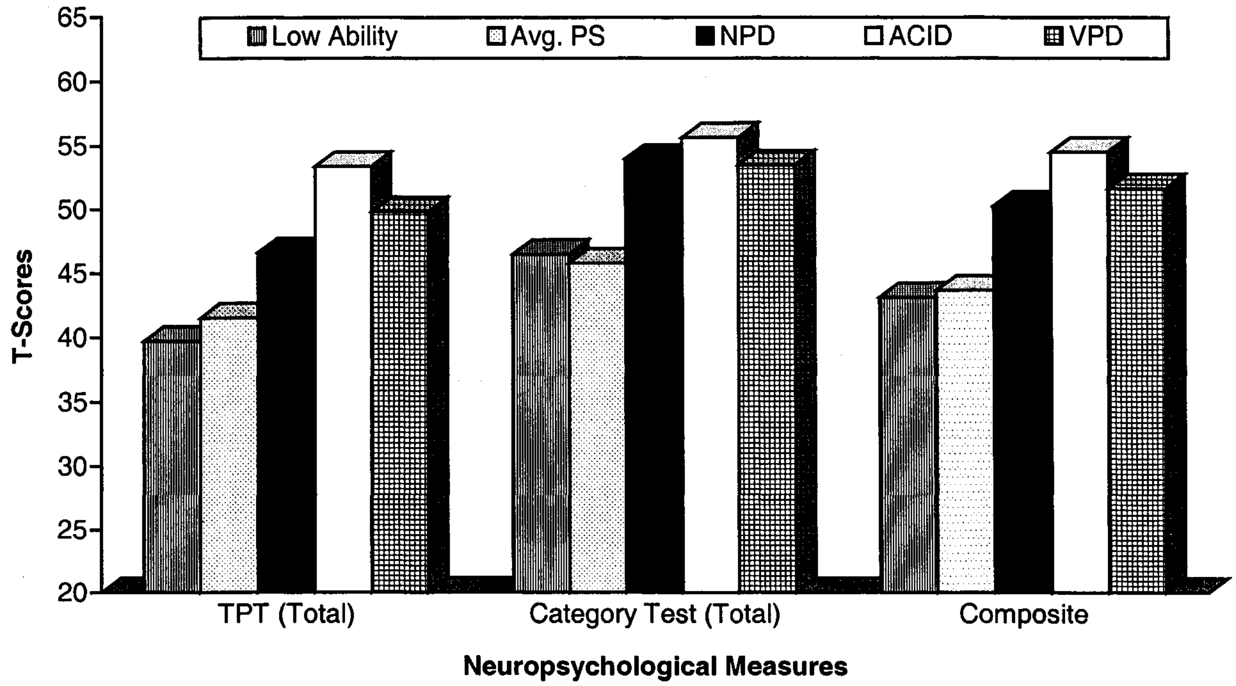


Figure 16.

Mean Performance by Subtype on measures within the Problem Solving Domain



Indeed, it was the only subtype that was significantly different than any of the other subtypes on measures within the attention domain.

With respect to the language domain, the Low Ability group performed significantly lower than the ACID and NPD groups on Verbal Fluency ($p<.05$), Auditory Closure, Speech Sounds Perception Test, and the language composite ($p<.01$). The Low Ability group also performed lower on the Sentence Memory Test in comparison to the ACID group ($p<.05$). Curiously, the NPD group did not differ significantly from the ACID group on any of the measures within the language domain, and was only differentiated from (i.e., significantly higher than) the VPD group by the language composite ($p<.01$). Closer inspection of individual post-hoc comparisons revealed that differences between the Low Ability and VPD groups on the Sentence Memory Test closely approached statistical significance. However, the NPD group did perform significantly better than the Avg. PS group on some measures (Speech Sounds Perception Test and language composite). In addition, as a group, the VPD subtype performed significantly lower than the ACID subtype on all measures within the language domain, including the composite ($p<.05$). With the exception of the Auditory Closure and Verbal Fluency Tests, the Avg. PS group performed lower than the ACID group on each of the other language measures ($p<.05$).

There was also some variability in the degree to which individual measures differentiated subtypes within the Attention Domain. While the Low Ability group performed significantly lower than each of the other subtypes on the Trail Making Test, Part B and the attention composite ($p<.001$), the Trail Making Test, Part A and the Target Test less consistently differentiated between the groups. The Low Ability group achieved significantly lower scores than the Avg. PS, ACID, and VPD groups on the Trail Making Test, Part A, and it obtained lower scores than the NPD, ACID, and VPD groups on the Target Test. None of the other groups differed from one another on attentional measures.

Of the measures from the motor domain, the Finger Tapping test contributed least to the differentiation between WISC-III-WIAT subtypes. That is, none of the groups differed from one another on this measure (on dominant, nondominant, or both hands together), making it the least sensitive to group differences of those measures included in the present work. The Avg. PS, and NPD groups were outperformed by the ACID group on the Grooved Pegboard Test on the dominant hand and were outperformed by both the ACID and VPD groups on the nondominant hand ($p<.01$). Similarly, the Low Ability group obtained a significantly lower score on both trials (i.e., dominant and nondominant hands) of the Grooved Pegboard in comparison to the ACID and VPD groups. However, only the Low Ability and NPD groups were significantly different from the ACID and VPD groups on the composite measure ($p<.01$). There were no significant differences between the ACID and VPD groups on measures of motor functioning (individual and composite scores).

Finally, on measures of nonverbal problem solving/concept formation, the Low Ability and Avg. PS groups were differentiated from the NPD, ACID and VPD groups. Specifically, the Low Ability and Avg. PS groups both obtained significantly lower scores than the NPD, ACID, and VPD groups on the Category Test and problem solving composite ($p<.01$). The Tactual Performance Test, total score differentiated between the Low Ability and Avg. PS groups in comparison to the ACID and VPD groups, with the former demonstrating significantly lower scores ($p<.01$ and $p<.02$, respectively).

Summary of Results

In the following section, the results will be discussed in the context of specific hypotheses proposed at the outset of the current study.

Identification and Internal Validation of a WISC-III-WIAT Typology

In general, it was expected that the present work would identify a valid ability-achievement typology, which would consist of subtypes similar to those generated by Saunders et al., (2001). This hypothesis was generally supported in that; a) the current typology was reasonably well replicated across three of the four hierarchical clustering algorithms, and b) at least three of the five subtypes were found to be highly similar to subtypes described by Saunders et al.. However, the cluster solutions derived did not correspond with the clusters of Saunders et al. as well as initially anticipated. In addition, some of the WISC-III-WIAT clusters were correlated with one another.

Hypothesis (a) i:

It was expected that that the current study would generate a cluster solution that would include a subtype defined by relative verbal deficiencies (i.e., PIQ>VIQ; low scores on WISC-III verbal subtests; low achievement on WIAT reading and spelling tasks). This hypothesis was supported by the data. Indeed, the VPD group was characterized by greater PIQ than VIQ (differing by more than one standard deviation), consistently lower scores on WISC-III verbal subtests, and below average scores on WIAT measures of academic achievement (lower scores on WIAT Word Reading and Spelling in comparison to math-based tasks).

Hypothesis (a) ii:

It was predicted that similar to the results of Saunders et al. (2001), a cluster with broad-based processing deficiencies (reduced IQs, WISC-III and WIAT subtest scores) would be found. This hypothesis was supported. The present cluster solution included a group that was labeled Low Ability, which was defined by below average scores across all WISC-III and WIAT subtests, with the exception of broadly average skills on two WISC-III nonverbal measures (Picture Completion and Object Assembly). It is noteworthy that the Saunders et al. broad-based processing deficiencies group also exhibited peak performances on these two WISC-III subtests.

Hypothesis (a) iii:

It was expected that that the current study would find a subtype characterized by relative nonverbal deficiencies (i.e., VIQ>PIQ; low scores on WISC-III performance subtests; low achievement on WIAT subtests of mathematical ability). The NPD group exhibited many of these features (e.g., lower scores on WISC-III Performance subtests in comparison to scores on verbal subtests, low achievement on WIAT subtests of mathematical ability).

Hypothesis (a) iv:

It was predicted that a group consistent with an ACID pattern (lowest scores on WISC-III Arithmetic, Coding, Information and Digit Span) would be among the clusters identified in the present study. This hypothesis was supported by the data. Consistent with an ACID pattern, this group also demonstrated the characteristically reduced scores on measures of academic achievement (with slightly weaker scores on language-based tasks of the WIAT).

External validation of a WISC-III-WIAT typology

It was hypothesized that subgroup membership would be associated with neuropsychological profiles in a manner that was consistent with neuropsychological theory and previous research findings. While the cluster solution, and hence, the neuropsychological profiles associated with the derived clusters were more complicated than anticipated, many of the neuropsychological measures consistently differentiated between the groups within and across neuropsychological ability domains. For example, although each measure selected to represent the various neuropsychological domains did not equally contribute to the differentiation of the clusters, measures within each domain tended to collectively differentiate between the same clusters. The only exception to this was the Finger Tapping Test, which did not discriminate between any of the subtypes.

The following specific predictions regarding neuropsychological performance were made at the outset of the current research:

Hypothesis (b) i:

It was predicted that a cluster characterized by relative verbal deficiencies (VIQ<PIQ; low scores on WISC-III verbal subtests; low achievement on WIAT reading and spelling tasks) would be found in the present research, and that this cluster would demonstrate lower scores on measures of the language domain in comparison to clusters characterized by an ACID pattern or relative nonverbal deficiencies. This hypothesis was generally supported by the results in that the VPD group performed significantly lower than the ACID group on all measures within the language domain, and was significantly weaker on the language composite in comparison to the NPD group. However, none of the individual subtests differentiated between the VPD and NPD groups.

Hypothesis (b) ii:

It was expected that a cluster defined by broad-based processing deficiencies (reduced WISC-III IQs, WISCIII and WIAT subtest scores) would demonstrate lower scores on measures of the problem solving, attention and language domains in comparison to clusters characterized by an ACID pattern, relative nonverbal deficiencies, and relative verbal deficiencies. While the Low Ability group of the present study was significantly different from most of the clusters across each of the measures within the attention domain, including the composite, the degree to which it differed from the NPD, ACID, and VPD clusters on measures within the language and problem solving domains varied. Specifically, the Low Ability group differed from the NPD, ACID, and VPD groups on the Category Test, and the problem solving composite. It also differed significantly from the ACID and VPD groups on the Tactual Performance Test. Similarly, the Low Ability group obtained lower scores than the ACID and VPD groups on each of the measures within the language domain and on the language composite, with the exception of the Sentence Memory Test.

Hypothesis (b) iii:

It was expected that that the current study would generate a cluster solution that would include a subtype defined by relative nonverbal deficiencies (VIQ>PIQ; low scores on WISC-III performance subtests; low achievement on WIAT subtests of mathematical ability) that would demonstrate lower scores on measures of the problem solving domain in comparison to clusters characterized by an ACID pattern or relative verbal deficiencies. This group was also expected to demonstrate weaker scores on measures of the motor domain than clusters characterized by an ACID pattern, relative verbal deficiencies, and broad-based deficiencies. The NPD subtype was not differentiated from the ACID or VPD subtypes on any of the measures within the problem solving domain. However, it did demonstrate significantly lower scores than the ACID subtype on the Grooved Pegboard (both hands) and problem solving composite measure. It also demonstrated weaker performance on the Grooved Pegboard (Nondominant hand) and problem solving composite measure than the VPD group.

Hypothesis (b) iv:

It was predicted that similar to the results of Saunders et al. (2001), a cluster characterized by an ACID pattern (lowest performance on WISC-III Arithmetic, Coding, Information and Digit Span; low achievement on WIAT subtests) would demonstrate lower scores on measures of the attention and language domains in comparison to the cluster characterized by relative nonverbal deficiencies. This hypothesis was not supported. The ACID subtype of the current study did not differ from the NPD subtype on any of the attention or language measures.

CHAPTER IV

DISCUSSION

The purpose of the present study was two-fold. The primary objective was to identify a meaningful ability-achievement typology using the WISC-III and WIAT, based on the research of Saunders et al. (2001). Second, was to examine the neuropsychological profiles associated with subtype membership as a means of demonstrating the external validity of the derived cluster solution.

Comparison of results obtained using several two-stage cluster analyses suggested the presence of five distinct ability-achievement subtypes. Of the five subtypes, three bore a considerable resemblance to profiles generated by Saunders et al. (2001), and all five were similar to subtypes that have been identified in previous research, confirming current theoretical perspectives on the heterogeneity of learning disability. Furthermore, the five subtypes exhibited distinct patterns of performance on neuropsychological domains, suggesting that the ability-achievement profiles are clinically meaningful and may be used to assist in designing appropriately tailored educational programs and interventions. The discussion that follows will be broadly categorized according to the two primary hypotheses of the present work.

Hypothesis (a): Identification and Internal Validation of a WISC-III-WIAT Typology

Based on previous research indicating that reliable subtypes of processing abilities can be identified through cluster analysis of the WISC-III, it was hypothesized that the present study would generate a reliable ability-achievement typology. Reliability was assessed though comparison of cluster solutions derived using four different hierarchical clustering algorithms.

Internal Validity of the WISC-III-WIAT Typology

Although replication is a commonly used approach by which a cluster solution may be deemed valid, there are a number of possible techniques which include examination of the effects of additional subjects, using highly correlated measurement instruments, dividing the sample in half, or subjecting the data to alternative statistical techniques (Morris & Fletcher, 1988). Like previous taxonomic research using the WISC-III, the current study compared solutions obtained using different clustering algorithms. Good agreement was obtained between three of the four clustering methods for both a five- and six-cluster solution. However, the specific clustering method that demonstrated the weakest agreement with the other methods was not consistent across solutions. That is, while Ward's method demonstrated the poorest correspondence with the other three hierarchical methods for a five-cluster solution, the method with the poorest correspondence for the six-cluster solution was UPGMA.

Within this context, it is important to consider that even though each of the hierarchical clustering algorithms is designed to classify individuals into relatively homogenous clusters, the manner by which each method evaluates the similarities and differences both within and between clusters varies considerably (Lange et al., 2002). For example, Ward's method seeks to minimize the within cluster variance by computing the sum of squares between two clusters across all variables (Lange et al). In contrast, the complete linkage method computes the distance between clusters as the maximum distance between any two members of the clusters. Therefore, while one might expect well-defined subtypes to emerge regardless of the clustering technique used, variation between profiles derived with different methods should be anticipated (Morris et al., 1981; Morris et al., 1998). Ultimately, the complete linkage five-cluster solution demonstrated the highest kappa values, was clinically meaningful and consistent with previously identified subtypes, and was therefore selected as being most representative of the data. The good agreement between three of the four clustering methods was taken to suggest that the current five-cluster solution was reliable.

Description of the WISC-III-WIAT Typology

Each of the WISC-III-WIAT subtypes was assigned a descriptive label based on the most salient features of each group. Accordingly, the Low Ability subtype was characterized by deficient performance on all but two subtests from the WISC-III (i.e., Picture Completion and Object Assembly). As a group, this subtype also obtained the lowest IQ scores in comparison to the other four subtypes. This subtype included 19% of the sample and was comprised of a significantly greater proportion of males than females (28 and 1, respectively).

The Avg. PS subtype demonstrated low to below average scores on each of the WISC-III and WIAT subtests, with the exception of those subtests that comprise the PSI. Curiously, this was the only group of the five subtypes that was comprised of an almost equal proportion of males and females (23 and 17, respectively). For the remaining subtypes, the ratio of males to females was otherwise biased in favor of the former, in keeping with other reported samples of learning disabled children (Fuerst & Rourke, 1993; Guerin, Griffin, Gottfried, & Christenson, 1993; Ward et al., 1999). However, many researchers have suggested that the finding of gender differences in learning disabled samples reflects a referral bias rather than a genuine disparity between male and female ability-achievement patterns (Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). Learning disabilities are more frequently identified in males than females, making it difficult to determine to what extent the gender configuration impacted upon the pattern of strengths and weaknesses of this cluster. Indeed, the Avg. PS group performed very similarly to the Low Ability group with the exception of its relatively well-developed performance on measures of processing speed. Twenty-two percent of the sample was assigned to this subtype.

The NPD subtype was characterized by relatively lower scores on the WISC-III Performance subtests in comparison to the WISC-III Verbal subtests and the WIAT subtests. There was a mean VIQ-PIQ difference; however, the discrepancy was rather small (i.e., eight standard score points). Similarly, the academic performance of this subtype was generally commensurate with ability levels on the WISC-III, which belies a discrepancy diagnosis of learning disability. This subtype included 21% of the sample, with 37 males and 2 females.

The ACID subtype demonstrated a mean WISC-III-WIAT profile characterized by relatively lower scores on the WISC-III Arithmetic, Coding, Information, and Digit Span, subtests and all four WIAT subtests, with notably weaker performance on the Spelling and Numerical Operations subtests. This group also exhibited a significant discrepancy between Verbal and Performance IQ scores (in favor of the latter) and underachievement on academic testing, fulfilling, at least by definition, the basic diagnostic criterion for a learning disability. However, this group also obtained the highest IQ scores of all the five subtypes. Seventeen percent of the sample was assigned to this subtype (26 males and 5 females).

The VPD subtype was characterized by relatively deficient scores on the WISC-III verbal subtests and WIAT subtests in comparison to the WISC-III Performance subtests. This subtype, like the ACID subtype, meets criteria for a learning disability based on a fairly large VIQ-PIQ discrepancy (i.e., a difference of 18 standard score points), and underachievement on all academic subtests. Twenty percent of the sample was assigned to this subtype, of which 30 were males and 7 were females.

To further examine the uniqueness of the current cluster solution, the relationship between the derived clusters was examined through correlation coefficients. The ACID, Low Ability, and VLD subtypes were found to be correlated with one another when the WIAT subtests were included, but only the ACID group was significantly correlated with other two groups when the WIAT subtests were excluded. This would suggest that the ACID, Low Ability and VLD groups share similar patterns of performance on the WISC-III and WIAT, and that similarities on the WIAT subtests in particular are responsible for the relationship between the Low Ability and VLD groups. It is not surprising that the ACID, Low Ability, and VLD groups would exhibit similarities in their WISC-III-WIAT profiles, given that each is characterized by relatively low scores on verbal and academic measures. That is, the ACID, Low Ability, and VLD groups each exhibited lower scores on the WIAT and WISC-III Verbal subtests relative to the WISC-III Performance subtests. However, upon examination of the obtained pattern of scores on the neuropsychological domains, it becomes abundantly clear that these three subtypes are distinct from one another and would therefore differ with respect to the interventions and services they require.

The Low Ability group was consistently the lowest performing group and differed significantly from the ACID group on each of the neuropsychological domain composite scores. The Low Ability group also obtained a significantly lower score than the VPD group on the motor and problem solving domain composite scores. With such globally depressed performance on ability, achievement, and neuropsychological testing, academic accommodations prescribed for children with the Low Ability profile would likely be less geared toward ameliorating specific weaknesses but may instead encourage the use of concrete materials in teaching, and organizing information into simple, brief units. In contrast, the ACID group obtained the highest score on each of the neuropsychological domain composite scores, and performed better than the VPD group on the language and global composite score. Based on this pattern of abilities, children within the ACID group would likely benefit from remediation of specific skills related to their underlying processing deficit(s), and educational accommodations that would address the primary academic deficits. On the other hand, the VPD group demonstrated difficulties circumscribed to the language domain, which could best be managed through specific instruction in phonetic and/or auditory processing skills, speech-language therapy, and/or remediation in reading, in combination with accommodations that would at least partially compensate for such deficits. In addition, analysis of the qualitative aspects of one's performance is recommended to further enhance an understanding of the specific processing deficits unique to the child.

Relationship to Known Subtypes

In addition to attempts at replicating the five-cluster solution across different hierarchical methods, the reliability of the present cluster solution was evaluated through examination of the relationship between the current five subtypes and those identified in previous research. That a full complement of distinct ability-achievement subtypes was generated, and that each of the ability profiles have been identified in previous taxonomic research, confirms the well-accepted view that learning disability and processing abilities are best viewed as heterogeneous diagnostic categories. It also suggests that the ability profiles identified in the current research reflect valid and reliable subtypes, some of which may be seen among the general population and others which appear to be more characteristic of clinic-referred samples.

Consistent with initial hypotheses, the current work identified a number of clusters that were similarly found by Saunders et al. (2001). Indeed, all four of the ability profiles that generally matched the profiles generated by Saunders et al. were those which were hypothesized to be present in the current study. That not all of the Saunders et al. subtypes were replicated not only corresponds with the original hypotheses but is also reasonable given that the present study was designed to extend their research. As such, differences in the methodology and sample should be considered.

First, and probably most significant, is that the methodology adopted in the current work was considerably different than that used in the research of Saunders et al. (2001). Rather than subject WISC-III scores to cluster analysis followed by an evaluation of differences in mean academic profiles, the current work included both ability and achievement scores in the development of the five-cluster typology. Given the significance of ability and achievement testing in the context of psychoeducational assessments, it seemed fitting that the relationship between patterns of ability and achievement be explored through cluster analysis. The five ability-achievement subtypes were then examined on the basis of neuropsychological patterns of strengths and weaknesses. In addition, Saunders et al. compared their final solution, derived using a single hierarchical clustering method (with *k*-means correction), to that of a separate *k*-means analysis. In contrast, the current study applied four different hierarchical clustering algorithms (with *K*-means correction) to the data, from which the most reliable solution was selected for further evaluation. Taken together, these methodological variations would certainly have contributed to the observed differences between obtained results.

Second, the sample of Saunders et al. (2001) differed from the current study sample. Many of the participants of the Saunders et al. study did not have complete neuropsychological test data. While the reason for this missing data is not entirely clear (i.e., whether due to random or systematic factors), the most likely explanation is that these participants were not administered selected tests during their neuropsychological examination. At the facility from which the data were gathered, the most common reason for curtailing the number of tests administered during a neuropsychological assessment was due to the child's inability to complete the entire examination within a designated period of time. That is, those children who worked more slowly tended to be evaluated on fewer measures. Of those measures to be eliminated, among the most common was the Tactual Performance Test, which was among the measures included in the present study. Furthermore, there was a statistically significant difference in the mean ages of the two samples, with the current study sample being just over a year and a half older than the excluded Saunders et al. participants. Therefore, in addition to the current sample being comprised of considerably fewer subjects, the excluded participants of Saunders et al. may be qualitatively different from the retained participants.

Furthermore, both of the Saunders et al. (2001) subtypes that were not generated in the current work (i.e., deficient working memory and deficits in tasks involving visual sequencing and language abilities) were unique to their study (i.e., they have also been absent from previous cluster analytic research), with the exception of the research of Waxman et al. (2003), which utilized the same sample. That is, even though these profiles may be encountered in clinical practice, they have yet to be empirically validated among children who were not included in the sample of Saunders and colleagues.

Despite the differences in methodology and sample selection, the current cluster solution and that of Saunders et al. (2001) are remarkably similar. Furthermore, the five subtypes identified in the current study share several characteristics in common with previously identified WISC-III subtypes.

The current five-cluster solution included a subtype with average performance on the Picture Completion subtest in the context of an otherwise below average

WISC-III-WIAT profile. Saunders et al. (2001) similarly identified a subtype characterized by deficits on each of the WISC-III subtests, with the exception of an average score on Picture Completion. Indeed, variations of this profile have been consistently reported in the WISC-III subtyping literature. For example, in evaluating WISC-III patterns of performance in the standardization sample, Donders (1996) identified a subtype that he described as exhibiting below average scores on all four factor indexes. Examining WISC-III subtest score patterns (excluding Mazes and Symbol Search) in the standardization sample, Konold et al. (1999) generated an eight-cluster solution which included a subtype characterized by subaverage performance on each of the ten core subtests. A subtype of below average WISC-III and WIAT Index scores in the WISC-III-WIAT linking sample (Glutting et al., 1994) and among children diagnosed with a learning disability (Ward et al., 1999) was also reported.

While the Saunders et al. (2001) study did not identify a subtype characterized by subaverage ability with relative strengths on the subtests comprising the PSI, this group has been previously described in the cluster analytic literature. Waxman et al. (2003) identified a subtype characterized by below average ability overall with average scores on both subtests of the PSI. Similarly, Glutting et al. (1994) identified a subtype characterized by below average ability and achievement with a relative strength on the PSI. Donders' (1996) cluster analysis of the WISC-III Index scores in the standardization sample generated a five-cluster solution with two of the five clusters differing in terms of performance on the PSI; one of these groups was described as exhibiting a relative strength on the PSI, while the other group demonstrated a relative weakness on the PSI. The present findings provide further support for the validity of this subtype, suggesting that performance on the PSI subtests may be particularly meaningful in the context of children's cognitive functioning and the identification of primary processing assets and deficits.

Based on the findings of Saunders et al. (2001) and a large literature which has identified and researched the Syndrome of Nonverbal Learning Disabilities (NLD; for a review, see Harnadek & Rourke, 1994, Rourke, 1989; Rourke et al., 1983), it was hypothesized that the present cluster analysis would identify a subtype characterized by relatively deficient nonverbal and mathematical abilities which would be accompanied by specific neuropsychological weaknesses (i.e., relatively deficient performance on the motor and problem solving domains). While this hypothesis was generally borne out in that the NPD group exhibited lower scores on WISC-III Performance subtests relative to the Verbal subtests, Low Average scores on both of the WIAT math subtests, and relatively reduced scores on the motor domain, there is an important outstanding characteristic that is worth noting. Namely, that the discrepancy between VIQ and PIQ was rather small (i.e., eight standard score points) in comparison to the discrepancies identified in previous studies which have found a similar subtype (e.g., Saunders et al.). Indeed, the Saunders et al. deficient nonverbal subtype demonstrated a VIQ-PIQ discrepancy of almost 18 standard score points. Similarly, Waxman et al. (2003) reported a 27 point split between VIQ and PIQ in their subtype which was described as exhibiting nonverbal processing deficits.

Donders and Warchausky (1997) identified a subtype with greater VIQ than PIQ (18 standard score point difference) in their study of head injured children in which subtype membership was also found to be associated with injury severity (severe brain injuries) and pathology (right cerebral) characteristics. In the standardization data, Konold et al. (1999) identified two subtypes with higher than expected rates of VIQ-PIQ discrepancies (in favor of the former); one which demonstrated average ability levels and another which demonstrated above average ability. However, these researchers did not examine the academic profiles associated with their eight-cluster typology. Thus, the lack of a significantly large VIQ-PIQ discrepancy in the current NPD cluster may partially explain why this group did not exhibit some features consistent with NLD, and therefore did not fulfill some of the predictions regarding neuropsychological functioning proposed at the outset of the current study. For example, the NPD group was expected to perform more poorly on the problem solving domain in comparison to the VPD and ACID groups; however, it did not perform significantly different from either group on any of the measures which comprise the problem solving composite. Similarly, of the four subtypes that overlapped with subtypes of Saunders et al., the NPD group demonstrated the weakest correspondence, which may suggest that this group is different from previously identified groups with NLD features.

With respect to the ACID subtype generated in the current study, the Arithmetic, Coding, Information, and Digit Span were among the lowest obtained WISC-III scores with a 13 point VIQ-PIQ discrepancy. This pattern of performance is consistent with that obtained by the deficits consistent with an ACID pattern subtype of Saunders et al. (2001). Others have also reported on the ACID or partial ACID pattern in their research; however, the criteria used to identify the ACID profile has not been consistent across studies (Daly & Nagle, 1996; Prifitera & Dersh, 1993). While previous research has reported a greater preponderance of the ACID pattern among clinical samples relative to healthy comparison groups, several researchers have found that relatively few individual children actually exhibit the ACID profile (Daly & Nagle; Joschko & Rourke, 1985; Prifitera & Dersh; Ward et al., 1995). Accordingly, it would be instructive to examine the individual profiles that comprise the current ACID subtype, both with respect to the degree of variability in the children's learning needs (e.g., potentially different ACID subtypes), and the proportion of children who, at the individual level, present with the ACID pattern. Such information would lead to a better understanding of the strengths and weaknesses obtained by the ACID group identified in the current study. It would also shed light on the underlying abilities that unite this potentially heterogeneous group of children.

The last of the five subtypes was characterized by a pattern of weak verbal and academic skills. As a group, the VPD subtype demonstrated a VIQ-PIQ discrepancy of 18 standard score points (in favor of the latter) with underachievement on all academic measures, meeting definitional criteria for a learning disability. The ability profile of the VPD group was consistent with the mean WISC-III profile of the deficient language abilities group of Saunders et al. (2001), and supported the original hypothesis that such a subtype would be present in the current work. Waxman et al. (2003) similarly identified a subtype characterized by verbal processing deficits which was almost identical to the deficient language abilities group of Saunders and colleagues. However, although, numerous reports of language-based learning disabilities can be found in the literature (e.g., Boder, 1973; Guerin et al., 1993; Harnadek & Rourke, 1994; Shaywitz et al., 1990), not all cluster analytic research with the WISC-III has identified a language disordered subtype.

In his cluster analysis of the WISC-III standardization data, Donders (1996) did not find a subtype with WISC-III Index score patterns which would suggest significant impairment in either verbal or nonverbal skills. Similarly, Glutting et al. (1994) failed to identify a subtype with discrepant VIQ-PIQ scores in their research on the WISC-III-WIAT linking sample.

In contrast, Konold et al. (1999) identified a subtype characterized by a greater than expected proportion of VIQ-PIQ discrepancies (in favor of the latter) in the standardization sample; however, these researchers did not measure how the eight subtypes varied on academic measures. Ward and colleagues (1999) also identified a language-based learning disabled group among a sample of children who had been earlier diagnosed with a learning disability. Similar to the profile identified in the current study, this subtype demonstrated the characteristic discrepancy between VIQ and PIQ, and between ability and achievement.

While each of the generated WISC-III-WIAT subtypes are consistent with those identified in previous WISC-III or WISC-III-WIAT taxonomic research, unlike most prior research, the current study failed to identify a subtype characterized by a "normal" or average WISC-III-WIAT profile. It is certainly likely that some children within the present sample obtained ability-achievement profiles that fell within the average range of ability. However, even children who might be described as exhibiting average ability and achievement would nevertheless be expected to demonstrate some variation within the average range. Indeed, subtypes were generated based on group means, and as such, children with broadly average WISC-III-WIAT profiles may have been distributed among the various clusters. Furthermore, the uniqueness of the sample itself may account for there being no "normal" subtype. That is, the present sample was comprised of children who were referred based on their meeting criteria to be assessed within a children's mental health centre, and therefore, it is likely that a considerable proportion of the sample experienced behavioural and/or emotional problems in addition to learning or cognitive difficulties. The presence of such psychological difficulties, in combination with cognitive or learning problems, may be associated with clinically complex disorders and syndromes, and has the potential to negatively impact performance on standardized testing. Thus, in contrast to previous research that has evaluated standardization samples or clinical samples with primary medical or neurological disorders (e.g., brain injury), the clinical sample of the current study could be described as consisting of children with primarily developmental disorders who are qualitatively different from those who may be evaluated in a major medical centre.

Hypothesis (b): External Validation of a WISC-III-WIAT Typology

Because cluster analytic statistical methods will impose a classification structure upon any data to which they are applied, it is essential that the internal validity of the obtained typology be demonstrated (Butler, Rourke, Fuerst, & Fisk, 1997; Joschko & Rourke, 1985). However, regardless of the statistical stability of a cluster solution, any typology is only as useful as the practical and clinical outcomes or treatments it predicts or directs (Butler et al., 1997).

To this end, commonly used neuropsychological measures were used to assess the external validity of the current five-cluster solution. Summary scores reflecting general ability domains and a global composite were included in the analyses because of differences in the stability, sensitivity and discriminability of the neuropsychological measures (Brown et al., 1989; Davis et al., 1989; Francis et al., 1992; Leckliter, Forster, et al., 1992). Indeed, of the thirteen neuropsychological variables evaluated, only the Finger Tapping Test failed to differentiate between any of the five clusters. However, this finding is not unexpected and is in fact in keeping with current knowledge about brain-behaviour relationships.

Many of the individual measures used to compute the ability domains were correlated with one another, both within and between ability domains. Not surprisingly, the three tests comprising the attention domain were significantly correlated with all but one of the neuropsychological measures (i.e., Trail Making Test Part A was not significantly correlated with the Tactual Performance Test), reflecting the multidimensional nature of the tasks and the influence of attentional skills on other areas of functioning. Upon closer inspection of the correlation matrix, it is evident that the majority of between-test comparisons, both between and within domains, are characterized by an objectively low correlation coefficient. This suggests that while there is some overlap in the skill demands of various tasks, the measures are indeed assessing distinct ability composites, each of which appears to be contributing a unique dimension to the domain construct to which it was categorized. Accordingly, interpretation of the results was primarily based upon individual test results and domain composites, rather than the global composite measure.

Hypotheses regarding the neuropsychological profiles associated with the five subtypes were generally supported, suggesting that the current five-cluster ability-achievement typology has clinical utility. However, some of the initial predictions were not met; the implications of which warrant further consideration in the present discussion.

The VPD group performed as expected across the neuropsychological domains, and on the language composite in particular, which differentiated the VPD subtype from the NPD and ACID subtypes. However, the constituent language measures only discriminated between the VPD and ACID groups. This is likely due to the relatively conservative criterion adopted to address the multiple comparisons. Indeed, the difference between the VPD and NPD groups on the Sentence Memory Test nearly reached statistical significance.

The Low Ability subtype obtained the lowest scores and was consistently impaired on each of the neuropsychological domains and the global composite. While this was generally in keeping with initial predictions, there are some important findings that deviate from the hypotheses which are worth noting.

The Low Ability and Avg. PS groups exhibited a comparable level of performance on about half of the neuropsychological domains. Both groups were differentiated from the NPD and ACID groups on the language domain, and from the NPD, ACID and VPD groups on the problem solving domain.

The observed association between a globally deficient (or almost globally deficient) ability-achievement profile and weak motor skills was not expected. Specifically, the Low Ability and Avg. PS groups obtained a relatively low score on the Grooved Pegboard test, which was significantly different from the ACID group on both hands, and from the VPD group on the nondominant hand. However, the motor composite only differentiated the Low Ability group from the ACID and VPD groups; the Avg. PS group was not sufficiently differentiated from any of the other groups on this composite. A possible explanation for this finding relates to the fundamental difference between the Low Ability and Avg. PS groups, namely, performance on the PSI. Indeed, the Avg. PS group exhibited a relative strength on the WISC-III subtests designed to measure psychomotor processing speed, while the Low Ability group maintained consistently below average WISC-III scores, with the exception of the Picture Completion subtest. Thus, the ability of the Avg. PS group to perform well on such clerical type tasks which require rapid visual scanning and eye-hand coordination could be thought to extend to the Grooved Pegboard Test, which requires similar skills.

On the attention domain, the Low Ability group obtained significantly lower scores than the Avg. PS group on both parts of the Trail Making Test and the attention composite. However, upon closer examination of the mean scores, it becomes evident that the Avg. PS group obtained below average scores on each of the attentional measures, with the exception of the Trail Making Test, Part A, which like the subtests of the PSI, is a relatively simple task that requires graphomotor speed, visual attention and eye-hand coordination. This finding further supports the uniqueness of the Avg. PS subtype from globally deficient subtypes, although the clinical meaningfulness of this profile requires further research.

Hypotheses regarding the performance of the NPD subtype on the neuropsychological domains were not consistently supported. In keeping with earlier predictions, this group obtained significantly lower scores on the Grooved Pegboard (nondominant hand), and the motor composite than the ACID and VPD groups and was not differentiated from the Low Ability group on either measure with each performing below the average range of ability. However, the NPD group was expected to score significantly lower than the ACID and VPD groups on the problem solving domain, but instead, the only significant difference on measures of problem solving was between the NPD group and the Low Ability and Avg. PS groups on the Category Test and the problem solving composite; the Tactual Performance Test failed to differentiate the NPD group from any of the other groups. One possible explanation may relate specifically to the Category Test. That is, previous research has demonstrated that the Category Test is sensitive to cerebral impairment, but may not be useful in the localization or differentiation between different clinical conditions (Reitan & Wolfson, 1985). Therefore, while the Category Test was sensitive to the large discrepancy between the two lowest performing subtypes (i.e., the Low Ability and Avg. PS subtypes) and the remaining three subtypes, it was not helpful in differentiating between more subtle differences in conceptual problem solving skills. That having been said, the finding of poor differentiation between the NPD group and the ACID and VPD groups was nevertheless surprising given that the Tactual Performance Test has been shown to be among the best measures to discriminate between children with NLD and children with language-based learning disabilities or normally developing children (Rourke & Harnadek, 1994).

Taken together, it seems likely that the current NPD group is distinct from previously identified groups of children who exhibit NLD. While it shares a relative weakness on the WISC-III Performance subtests and WIAT math-based tasks, the NPD group does not exhibit the characteristically large VIQ-PIQ discrepancy that has been described in the literature.

The ACID group obtained the highest WISC-III IQ scores and outperformed each of the other subtypes (i.e., Low Ability, Avg. PS, NPD, & VPD) on all four neuropsychological domain composites and the global composite. Based on previous research, which has demonstrated a higher incidence of the ACID profile among children with ADHD, and/or predominantly language-based learning disabilities (Joschko & Rourke, 1985; Prifitera & Dersh, 1993; Rourke et al., 1983; Ward et al., 1995), it was expected that the ACID subtype of the current study would obtain lower scores on the attention domain in comparison to the NPD subtype.

That the Low Ability group was the only subtype to differ from the other groups on the attention domain was an unexpected finding. One possible explanation may relate to the specific measures used to form the attention domain. Both the Trail Making Test and the Target Test, although seemingly simple tasks of visual attention, require the integration of multiple abilities for successful completion. For example, the Trail Making Test has been demonstrated to be a measure of attention (Klonoff, 1971; Krug et al., 1995; Livingston et al., 1997; O'Donnell, MacGregor, Dabrowski, Oestreicher, & Romero, 1994), higher order sequencing, motor output (Francis et al., 1992), spatial speed of operations, spatial memory, mental flexibility, sustained attention (D'Amato et al., 1988), visual scanning and visual-motor integration (Chittooran, D'Amato, Lassiter, & Dean, 1993). While research has demonstrated that both the Target Test and Trail Making Test are sensitive to the effects of cerebral dysfunction (Crowe, 1998; Davis et al., 1989; Reitan, 1971, 1974), without a qualitative analysis of the individual child's performance, it would be difficult to determine which component ability contributed to the child's difficulty on the task. In addition, performance on the Trail Making Test has been shown to be affected by psychometric intelligence, but only at the lower end of the ability spectrum (Crowe), which could also partially explain the significantly weaker performance of the Low Ability group in comparison to each of the other subtypes.

Furthermore, just as learning disability refers to a number of potential processing deficiencies which result in difficulties acquiring and performing specific academic skills, the presence of an ACID profile may similarly suggest at least two possible cognitive profiles. Indeed, Joschko and Rourke (1985) identified two distinct subtypes of neuropsychological strengths and weaknesses which were derived from a sample of children who individually demonstrated the ACID pattern. While one group was found to demonstrate relatively deficient sequencing skills, another group exhibited deficits which were thought to reflect difficulties in the "revisualization" (Joschko & Rourke, p. 79) of symbols. Furthermore, several researchers have demonstrated that the ACID pattern is quite rare among individuals diagnosed with learning disabilities and that previous research on the ACID profile has been based on group means (Daly & Nagle, 1996; Joschko & Rourke; Prifitera & Dersh, 1993). Therefore, the percentage of subjects who actually exhibit the ACID pattern may be similarly low in the current study.

Clinical Implications

Results of the current study reinforce and highlight several findings of previous research on patterns of ability and achievement in children. Learning disability is indisputably best considered a heterogeneous classification label that refers to underlying processing deficiencies which often manifest in specific academic difficulties. That the present study identified five subtypes with distinct, but relatively commonly identified ability-achievement profiles attests to the robustness of various learning profiles within the population. For example, a group of children with predominantly low to below average scores on the WISC-III has been described in almost all of the cluster analytic studies using the WISC-III alone or in combination with the WIAT (i.e., Donders, 1996; Glutting et al., 1994; Konold et al., 1999; Saunders et al. 2001; Ward et al., 1999). Thus, it appears that many of the patterns of performance identified in the current work are reasonably invariant across samples and methodological applications. In addition, despite having used statistical methods which have previously produced clusters that were predominantly differentiated by level of performance, the current five-cluster solution exhibited variations in both level and pattern of performance across the WISC-III and WIAT subtests.

A primary goal of classification research is to identify patterns of strengths and weaknesses which predict a likely outcome or inform appropriate treatment and interventions (Glutting et al., 1994). That the five ability-achievement profiles exhibited distinct neuropsychological profiles which reasonably corresponded to theoretically and conceptually-based predictions put forward at the outset of the study suggests that the current five-cluster solution is clinically meaningful. For example, in combination with qualitative behavioural observations, a child's pattern of performance on the WISC-III and WIAT may be used to test hypotheses regarding his or her neuropsychological strengths and weaknesses. In addition, the fact that particular ability-achievement profiles have been shown to covary with various conditions and processing deficits, provides further support for their use in psychoeducational assessment procedures. Thus, with this knowledge, remediation or intervention programs can be appropriately tailored to address the specific processing deficits of the child. In addition, there were no differences in the mean age of the clusters, suggesting that the differences between the subtypes were not age-dependent.

In spite of the utility of subtyping endeavours, it is important to consider the uniqueness of the individual child within the psychoeducational context. Children classified into the same subtype certainly share several common characteristics, with respect to the primary grouping variables, which aid in the synthesis of complex traits and information which may then be used for rehabilitation and/or remediation purposes. However, it should not be assumed that all children assigned to the same subtype are identical. Instead, such children may be expected to vary in terms of their development, and early or current environmental experiences which would certainly impact upon the adaptive and psychosocial functioning of each child (Rourke, 1999). Thus, it is important that programs be designed to address both the common or shared and unique characteristics of the child's pattern of strengths and weaknesses (Rourke).

Limitations of the Present Study

A primary limitation of the present study relates to the sample. Due to the selection criteria employed, the results may be thought to be representative of only a subset of referred children who are able to complete a comprehensive neuropsychological test battery in a timely manner.

Although using retrospectively gathered data affords researchers the opportunity to conduct research that may not have otherwise been possible, or at least without the considerable resources of time and funding, such studies also have the drawback of being limited by the data available for research. In the case of the present work, the availability of neuropsychological measures to assess the external validity of the cluster solution was somewhat limited. Despite the fact that the current study is the only one of its kind to examine patterns of neuropsychological functioning in relation to a

WISC-III-WIAT typology, important ability domains remain unexplored. For example, knowledge of the performance of each of the subtypes on a range of measures sensitive to memory (e.g., visual, auditory, immediate, rote) or behavioural functioning would contribute further to our understanding of the external validity of the subtypes. in a similar vein, the present sample was composed of a clinically heterogeneous sample. Although the data for cross referencing purposes was not available, it would have been interesting to know whether some conditions are better represented within the sample than others, particularly with regards to subtype classification.

In addition, it is possible that the two subtypes of Saunders and colleagues (2001) which were not identified in the current study (i.e., deficient working memory and deficits in tasks involving visual sequencing and language abilities) would have been more readily replicated had the entire sample been included. On the other hand, the mean IQ scores of the current sample were almost identical to those reported for the Saunders' et al. sample, indicating that the two participant groups are comparable on the majority of the measures used to classify children into subtypes. Furthermore, each of the subtypes generated in the present study very closely approximate subtypes identified in previous research and, therefore, likely reflect ability-achievement profiles that are reasonably common among the child population.

Other limitations of the present study relate to the statistical methodology employed. Taxonomic research is viewed by some as a promising avenue of inquiry that continues to be hampered by methodological inconsistencies and unresolved questions (Lange et al., 2002). Cluster analytic techniques have only been in use since the 1960s (Morris et al., 1981), and, as an established statistical method, there remains some uncertainty with respect to the degree of confidence researchers can place in their cluster solution (Lange et al.). This is a consequence, at least in part, of the degree of subjectivity involved in conducting cluster analysis. Thus, although efforts were made to ensure that the similarity coefficient, clustering algorithms and measures of association used to demonstrate the internal validity of the resultant cluster solution followed relatively conventional, and empirically devised standards, ultimately, a somewhat subjective decision was required by the present author. Therefore, future research should attempt to validate the current five-cluster solution through replication in an independent sample.

Summary and Future Directions

In summary, the current results provide further support for the conceptualization of learning disabilities as a heterogeneous group of disorders characterized by distinct patterns of academic achievement and processing abilities. Research on the subtyping of learning disabilities has identified several valid and reliable patterns of cognitive strengths and weaknesses, which have important implications, particularly in the school setting.

Despite the fact that only two of the five subtypes would likely meet a discrepancy definition of a learning disability, it is clear that the obtained ability-achievement subtypes have clinical utility in terms of the outcomes they predict. That the subtypes exhibited distinct neuropsychological profiles which differed from one another in a reasonably predictable and meaningful fashion speaks to the validity of the current five-cluster solution. The finding that some of the ability-achievement subtypes exhibited similarities in their patterns of cognitive strengths and weaknesses, which were clearly differentiated by their neuropsychological profiles underscores the importance of neuropsychological testing within a psychoeducational context. Accordingly, it is recommended that for the purposes of school programming, and the design and implementation of treatment interventions, that a complete neuropsychological evaluation, including a full complement of ability domains, be undertaken.

In a similar vein, given the current results, which support the utility of neuropsychological evaluation within the psychoeducational context, examining the predictive value of ability-achievement subtypes on measures of adaptive or everyday functioning would be a worthwhile endeavour. Such research would likely contribute to a more thorough understanding of learning disability subtypes and hence, effective methods of treatment or intervention. In the case of NLD, investigations of adaptive and psychosocial functioning provided a fruitful avenue of research, which ultimately resulted in its identification and recognition as a unique constellation of cognitive, academic, and adaptive strengths and weaknesses. Indeed, many researchers have indicated that the current discrepancy-based diagnostic criteria are flawed, suggesting instead that learning disabilities be defined solely on the basis of deficiencies in specific skills or ability domains (e.g., Siegel, 1989). As such, examination of daily functioning would shift learning disability identification from statistically driven formulae to the assessment of those skills and abilities which may be most impacted by the underlying processing deficiencies.

The current study was limited by the availability of measures by which to evaluate the external validity of the cluster solution. It would therefore be useful to examine ability-achievement subtype profiles on a wider range of neuropsychological, adaptive, and behavioural domains, such as memory and psychosocial functioning. Access to such information in the current study could have facilitated a better understanding of the five subtypes, and in particular, the NPD group, which appears to share only a subset of those skills and abilities which are typically impaired in children with NLD. Indeed, NLD is thought to be characterized by primary weaknesses in visual perception, tactile perception, complex psychomotor skills and the ability to deal with novel material (Harnadek & Rourke, 1994). Performance of the present clinic-referred sample on measures of memory and psychosocial functioning would help address whether the NPD group is indeed significantly different from the NLD groups described in the literature. Thus, the initiation of prospectively designed research to assess the reliability and validity of the present five-cluster solution is recommended.

Similarly, unlike the WISC-III, for which the present sample had complete data (excluding the Mazes subtest), only four WIAT subtests were available for use in the development of the current ability-achievement typology. While it is difficult to predict the effects of including each of the WIAT subtests, as previous cluster analytic research using both the WISC-III and WIAT has been limited to the examination of factor scores, the inclusion of a greater range of academic skills would likely have resulted in the identification of additional subtypes or a better delineation of the current five subtypes. For example, mean performances on the Reading Comprehension and Pseudoword Decoding subtests could have potentially contributed significantly to a better understanding of the subtypes, particularly the VPD and NPD groups, and may have resulted in a subtype with an ability-achievement pattern more consistent with those reported for groups of children with NLD.

In addition, although not possible within the limitations of the present study, it would be instructive to compare the current five subtypes to children with clinically derived diagnoses, or learning disability subtypes. Such research would further validate the current five-cluster typology and would also have the potential to provide important information about whether or not, at the individual level, the five subtypes of the current research are qualitatively different from children who meet discrepancy-based diagnostic criteria.

Finally, the WISC-III has recently undergone a major revision which has culminated in the very recent publication of the WISC-IV. This latest addition to the Wechsler series, which will undoubtedly come to replace the WISC-III in common practice, not only includes new subtests, but is also structurally distinct from previous versions. As such, cluster analysis of the WISC-IV should provide an ideal opportunity to examine the robustness of previously identified ability and ability-achievement typologies.

APPENDIX I

STRUCTURE OF THE WISC-III

Verbal Intelligence Quotient (VIQ)
- Information
- Similarities
- Arithmetic
- Vocabulary
- Comprehension

Performance Intelligence Quotient (PIQ)
- Picture Completion
- Coding
- Picture Arrangement
- Block Design
- Object Assembly

Full Scale Intelligence Quotient (FSIQ)
- (Summary of core Verbal and Performance subtests)

Verbal Comprehension Index (VCI)
- Information
- Similarities
- Vocabulary
- Comprehension

Perceptual Organization Index (POI)
- Picture Completion
- Picture Arrangement
- Block Design
- Object Assembly

Freedom from Distractibility Index (FDI)
- Arithmetic
- Digit Span

Processing Speed Index (PSI)
- Coding
- Symbol Search

APPENDIX II

DESCRIPTION OF WISC-III SUBTESTS[1,2]

WISC-III Verbal Subtests:

WISC-III Information (Wechsler, 1991)

Involves elementary factual knowledge of history, geography, current events, literature, and general science. *Score*: number of correct items. *Task requirements:* retrieval of acquired verbal information. *Stimulus:* spoken questions. *Response:* spoken answer.

WISC-III Similarities (Wechsler, 1991)

Requires the test-taker to identify the most essential semantically common characteristics of each word pair. *Score*: number of correct responses (scoring ranges from 0 to 2). *Task requirements*: verbal abstraction. *Stimulus*: spoken question. *Response*: spoken answer.

WISC-III Arithmetic (Wechsler, 1991)

Requires the test-taker to solve a series of increasingly difficult arithmetic problems within a designated time period. *Score*: number of correct responses. *Task requirements*: working memory, attention, calculation skills, math reasoning *Stimulus*: spoken question. *Response*: spoken answer.

WISC-III Vocabulary (Wechsler, 1991)

Requires oral definition of words. *Score*: number of correct words (scoring ranges from 0 to 2). *Task requirements*: verbal definition. *Stimulus*: spoken word, with simultaneous presentation of written word. *Response*: spoken definition.

WISC-III Comprehension (Wechsler, 1991)

Involves everyday social judgement and common sense. *Score*: number of correct responses (scoring ranges from 0 to 2). *Task requirements*: knowledge of conventional standards of behaviour *Stimulus*: spoken question. *Response*: spoken answer.

WISC-III Digit Span (Wechsler, 1991)

Requires the test-taker to repeat series of orally presented digits, forwards and backwards. *Score*: number of correct responses. *Task requirements*: rote memory, attention and concentration. *Stimulus*: orally presented digit sequences. *Response:* oral repetition of digit sequences.

---

[1] Some descriptions of WISC-III measures were adapted from Rourke, Bakker, Fisk, & Strang (1983). *Child Neuropsychology An introduction to theory, research and clinical practice*. New York: Guilford Press.

[2] Some descriptions of WISC-III measures were adapted from Sattler, J. (1992). *Assessment of Children. Revised and updated Third Edition*. San Diego: Jerome M. Sattler, Publisher, Inc.

WISC-III Performance Subtests:

WISC-III Picture Completion (Wechsler, 1991)

The test-taker is to identify the most important missing detail from pictures of everyday objects and scenes. *Score*: number of correct responses. *Task requirements*: visual perception, differentiation of essential from nonessential details, alertness to detail. *Stimulus*: visually presented pictures. *Response*: gestured or spoken identification of missing detail.

WISC-III Coding (Wechsler, 1991)

The test-taker is presented with a code consisting of digits paired with symbols. He or she is also presented with a series of random digits and is asked to fill-in the missing symbol below the digit as rapidly as possible. *Score*: number of correct responses. *Task requirements*: eye-hand coordination, attention, short-term memory, visual perception. *Stimulus*: printed digits and symbols. *Response*: rapid coordination of visual identification and written symbol.

WISC-III Picture Arrangement (Wechsler, 1991)

Requires the test-taker to arrange a random sequence of pictures into a meaningful story as quickly as possible. *Score*: total score for speed and accuracy of arrangements. *Task requirements*: visual sequencing, attention to details, planning ability. *Stimulus*: picture cards. *Response*: placement of picture cards.

WISC-III Block Design (Wechsler, 1991)

Involves the arrangement of coloured blocks to form designs which match those presented on a printed card as quickly as possible. *Score:* total score for speed and accuracy of block placement. *Task requirements*: visual-spatial organization, psychomotor speed, synthesis of parts to whole. *Stimulus*: printed geometric design. *Response:* manipulation and arrangement of blocks.

WISC-III Object Assembly (Wechsler, 1991)

Involves the arrangement of pieces to form a picture as quickly as possible. Score: total score for speed and accuracy of assembly. *Task requirements:* visual-spatial organization, psychomotor speed, visual-motor coordination, synthesis of parts to whole. *Stimulus:* disarranged parts of picture. *Response:* manipulation and arrangement of parts.

WISC-III Symbol Search (Wechsler, 1991)

The test-taker is presented with a target and a second group of symbols and is asked to indicate whether either of the symbols from the target group is within the second group by checking a box. *Score*: number of incorrect responses subtracted from the number of correct responses. *Task requirements*: visual-perceptual scanning, visual discrimination. *Stimulus*: printed symbols. *Response*: manual checking of a box (yes or no).

APPENDIX III

DESCRIPTION OF SELECTED WIAT SUBTESTS[1]

WIAT Basic Reading (Wechsler, 1992)

The test-taker is presented with a series of printed words and is asked to read or sound-out each independently. Score: number of correctly pronounced words. *Task requirements*: phonetic decoding and word-reading ability, visual perception. *Stimulus*: printed word. *Response*: spoken word.

WIAT Mathematics Reasoning (Wechsler, 1992)

Requires the test taker to solve math problems of increasing difficulty, which are presented both visually and orally. The test-taker may use paper and pencil to calculate his or her responses. *Score*: number of correct responses. *Task requirements*: pattern analysis, calculation skills, knowledge of mathematical operations, problem solving skills. *Stimulus*: oral and visual description of math problem. *Response*: spoken answer.

WIAT Spelling (Wechsler, 1992)

The test taker is presented with a series of increasingly difficult dictated words, which he or she is required to spell. *Score*: number of accurately spelled words. *Task requirements*: phonetic skills, auditory processing skills, eye-hand coordination. *Stimulus*: spoken word. *Response*: written word.

WIAT Numerical Operations (Wechsler, 1992)

Requires the test taker to solve increasingly difficult paper and pencil calculation problems. *Score*: number of correct responses. *Task requirements*: calculation skills, knowledge of mathematical operations. *Stimulus*: printed math problem. *Response*: written response.

---

[1] Some descriptions of WIAT measures were adapted from Wechsler, D. (1992). Wechsler Individual Achievement Test: Manual. San Antonio, TX: Psychological Corporation.

APPENDIX IV

DESCRIPTION OF NEUROPSYCHOLOGICAL MEASURES[1]

(i) Measures of Auditory-Linguistic Domain

Auditory Closure Test (Kass, 1964)

The test-taker is required to blend 23 progressively longer chains of sound elements into words. *Score:* number correct. *Task requirements:* auditory perception, phonological awareness. *Stimulus:* tape-recorded sound elements. *Response:* spoken word.

Verbal Fluency Test

Requires the test-taker to name as many different words that begin with the sound "P" within 60 seconds. This is repeated, asking the test-taker to name as many words as he or she can that begin with the sound "C". *Score*: total number of words. *Task requirements*: initiation, linguistic fluency. *Stimulus*: verbal instructions. *Response*: spoken words.

Sentence Memory Test (Benton, 1965)

Requires the test-taker to repeat sentences of gradually increasing length and difficulty. *Score*: number of sentences correctly reproduced. *Task requirements*: auditory perception, attention, short-term memory. *Stimulus*: spoken sentences. *Response*: repeated sentences.

Speech-Sounds Perception Test (Reitan & Davison, 1974)

Requires the test-taker to identify from among 4 possible options, the written word that corresponds with the nonsense word presented via tape-recorder. *Score:* number correct. *Task requirements:* phonological processing, auditory perception. *Stimulus:* tape-recorded nonsense words. *Response:* underline word.

(ii) Measures of the Attention and Concentration Domain

Target Test (Reitan & Davison, 1974)

Following a brief delay, the test-taker is required to replicate visual-spatial configurations of increasing complexity, demonstrated (through tapping) by the examiner. *Score*: number correct. *Task Requirements*: visual perception, visual-spatial skills, and attention. *Stimulus:* tapped out pattern. *Response:* line drawings of reproduced visual-spatial patterns.

Trail Making Test (Reitan & Davison, 1974)

This test consists of two parts; part A and part B. In part A, using his or her pencil, the test-taker is instructed to rapidly connect the numbers one through 15 which are randomly placed on the page. In part B, the task is similar only the test-taker must alternate between numeric and alphabetic items as quickly as possible. *Score*: time to complete each task and number of errors. *Task Requirements*: eye-hand coordination, attention, visual scanning, and set shifting. *Stimulus*: encircled numbers and letters. *Response*: pencil drawn connecting lines.

---

[1] Some descriptions of neuropsychological measures were adapted from Rourke, Bakker, Fisk, & Strang (1983). *Child Neuropsychology An introduction to theory, research and clinical practice*. New York: Guilford Press.

(iii) Measures of the Psychomotor Speed and Coordination Domain

Finger Tapping (Reitan & Davison, 1974)

Requires the test-taker to rapidly tap a lever with his or her index finger in 10-second intervals. Beginning with the dominant hand, three successive trials are completed before alternating to the non-dominant hand. Five consecutive trials within five points of one another or up to a total of 10 trials per hand are recorded. *Score*: average number of taps for each hand. *Task Requirements*: motor speed and coordination. *Stimulus*: Lever attached to a manual counting device, mounted on a wooden board. *Response*: tap index finger.

Grooved Pegboard (Klove, 1963)

Requires the test-taker to fit keyhole shaped pegs into similarly shaped holes oriented in random directions on a 4"X4" board. Beginning with the dominant hand, the child is instructed to place each peg as quickly as possible. When using the right hand pegs are placed from left to right and from right to left when using the left hand. *Score*: time to complete the task with each hand and number of times the pegs are dropped. *Task Requirements*: eye-hand and fine motor coordination and speed. *Stimulus*: Pegboard and 25 pegs. *Response*: peg placement.

(iv) Measures of Nonverbal Problem-Solving/Concept Formation Domain:

Category Test (Reitan & Davison, 1974)

Involves the presentation of a visual pattern on a screen attached to four response levers denoted by the colours red, blue, yellow, and green. The test-taker must depress one of the four levers to indicate his or her answer. A bell or buzzer is sounded to signal correct and incorrect responses, respectively. The test is divided into 5 subtests, each consisting of a uniform principle. The final subtest is a summary of previously viewed items. The principles tested are colour, quantity, oddity, and colour prominence. *Score*: total number of errors. *Task requirements:* concept formation, pattern analysis, appreciation and incorporation of verbal feedback, hypothesis generation and testing. *Stimulus*: visual patterns. *Response*: lever depression.

Tactual Performance Test (Reitan & Davison, 1974)

While blindfolded, the test taker is required to place 6 differently shaped wooden blocks into their proper spaces on an upright board. The test is administered in three trials: first using only the dominant hand, followed by the non-dominant hand and then both hands simultaneously. After the blindfold is removed, the test-taker is asked to draw as many shapes as he or she can recall in their relative position on the board. *Score:* time to complete each trial; number of blocks correctly placed within time limit; number of blocks correctly recalled; number of blocks placed in their correct position (for the purposes of the present work, only the total time score was analyzed). *Task Requirements:* tactile perception, kinesthetic feedback, nonverbal problem-solving. *Stimulus:* wooden blocks and board. *Response*: placement of blocks and written diagram of board.

REFERENCES

Aldenderfer, M.S., & Blashfield, R.K. (1984). *Cluster Analysis*. Beverly Hills, CA: Sage Publications.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders (4th ed.)*. Washington, DC: American Psychiatric Association.

Anastasi, A., & Urbina, S. (1997). *Psychological Testing, 7th edition*. New Jersey: Prentice-Hall.

Bannatyne, A. (1974). Diagnosis: A note on recategorization of the WISC scaled scores. *Journal of Learning Disabilities, 7*, 272-274.

Bannatyne. A. (1979). Spatial competence, learning disabilities, auditory-vocal deficits and a WISC-R subtest recategorization. *Journal of Clinical Child Psychology, 8*, 194-200.

Batchelor, E., Sowles, G., Dean, R.S., & Fischer, W. (1991). Construct validity of the Halstead-Reitan Neuropsychological Battery for children with learning disorders. *Journal of Psychoeducational Assessment, 9*, 16-31.

Benton, A.L. Sentence Memory Test. Iowa City: Author, 1965.

Boake, C. (2002). From the Binet-Simon to the Wechsler-Bellevue: Tracing the history of intelligence testing. *Journal of Clinical and Experimental Neuropsychology, 24*, 383-405.

Boder, E. (1973). Developmental Dyslexia: A diagnostic approach based on three atypical reading-spelling patterns. *Developmental Medicine and Child Neurology, 15*, 663-687.

Bolen, L.M., Aichinger, K.S., Hall, C.W., & Webster, R.E. (1995). A comparison of the performance of cognitively disabled children on the WISC-R and WISC-III. Journal of *Clinical Psychology, 51*, 89-94.

Brown, S.J., Rourke, B.P., & Cicchetti, D.V. (1989). Reliability of tests and measures used in the neuropsychological assessment of children. *The Clinical Neuropsychologist, 3*, 353-368.

Butler, K, Rourke, B.P., Ruerst, D.R., & Fisk, J.L. (1997). A typology of psychosocial functioning in pediatric closed-head injury. *Child Neuropsychology, 3*, 98-133.

Burton, D.B., Sepehri, A., Hecht, F., VandenBroek, A., Ryan, J.J., & Drabman, R. (2001). A confirmatory factor analysis of the WISC-III in a clinical sample with cross-validation in the standardization sample. *Child Neuropsychology, 7*, 104-116.

Canivez, G.L. & Watkins, M.W. (1998). Long-term stability of the Wechsler Intelligence Scale for Children-Third Edition. *Psychological Assessment, 10*, 285-291.

Canivez, G.L. & Watkins, M.W. (2001). Long-term stability of the Wechsler Intelligence Scale for Children-Third Edition among students with disabilities. *School Psychology Review, 30*, 438-453.

Carlton, M., & Sapp, G.L. (1997). Comparison of WISC-R and WISC-III scores of urban exceptional students. *Psychological Reports, 80*, 755-760.

Casey, J., & Yawny, P. (2001, February). *The relationship between selected WIAT subtests and the WISC-III.* Poster presented at the 29th meeting of the International Neuropsychological Society, Chicago.

Chittooran, M.M., D'Amato, R.C., Lassiter, K.S., & Dean, R.S. (1993). Factor structure of psychoeducational and neuropsychological measures of learning-disabled children. *Psychology in the Schools, 30*, 109-118.

Crowe, S.F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. *Journal of Clinical Psychology, 54*, 585-591.

Daley, C.E., & Nagle, R.J. (1995). Relevance of WISC-III indicators for assessment of learning disabilities *Journal of Psychoeducational Assessment, 14*, 320-333.

D'Amato, R.C., Gray, J.W., & Dean, R.S. (1988). A comparison between intelligence and neuropsychological functioning. *Journal of School Psychology, 26*, 283-292.

Davis, R.D., Adams, R.E., Gates, D.O., Cheramie, G.M. (1989). Screening for learning disabilities: A neuropsychological approach. *Journal of Clinical Psychology, 45*, 423-428.

Dean, R.S. (1985). Review of the Halstead-Reitan Neuropsychological Test Battery. In J.V. Mitchell (ed.). *Ninth Mental Measurements Yearbook*. Lincoln, NE: Buros Institute of Mental Measurement.

Donders, J. (1996). Cluster Subtypes in the WISC-III standardization sample: Analysis of factor index scores. *Psychological Assessment, 8,* 312-318.

Donders, J., & Warchausky, S. (1997). WISC-III Factor Index score patterns after traumatic head injury in children. *Child Neuropsychology, 3*, 71-78.

Fisk, J.L., & Rourke, B.P., (1983). Neuropsychological subtyping of learning-disabled children: History, methods, implications. *Journal of Learning Disabilities, 16*, 529-531.

Flanagan, D.P. & Alfonso, V.C. (1993). Differences required for significance between Wechsler Verbal and Performance IQs and WIAT subtests and composites: The predicted-achievement method. *Psychology in the Schools, 30*, 125-132.

Fletcher, J.M., Foorman, B.R., Boudousquire, A., Barnes, M.A., Schatschneider, C., & Francis, D.J. (2002). Assessment of reading and learning disabilities. A research-based intervention-oriented approach. *Journal of School Psychology, 40*, 27-63.

Fletcher, J.M., Shaywitz, S.E., Shankweiler, D.P., Katz, L., Liberman, I.Y., Stuebing, K.K., Francis, D.J., Fowler, A.E., & Shaywitz, B.A. (1994). Cognitive profiles of reading disability. Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86*, 6-23.

Flynn, J.R. (1984). The mean IQ of Americans: massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29-51.

Francis, D.J., Epsy, K.A., Rourke, B.P., & Fletcher, J.M. (1991). Validity of intelligence test scores in the definition of learning disability: A critical analysis. In B.P. Rourke (Ed.) *Neuropsychological Validation of Learning Disability Subtypes* (pp. 15-44). New York: Guilford Press.

Francis, D.J., Fletcher, J.M., Rourke, B.P., & York, M.J. (1992). A five-factor model for motor, psychomotor, and visual-spatial tests used in the neuropsychological assessment of children. *Journal of Clinical and Experimental Neuropsychology, 14*, 625-637.

Francis, D.J., Fletcher, J.M., Shaywitz, B.A., Shaywitz, S.E., & Rourke, B.P. (1996). Defining learning and language disabilities: Conceptual and psychometric issues with the use of IQ tests. *Language, Speech, and Hearing Services in Schools, 27*, 132-143.

Frank, G. (1983). *The Wechsler Enterprise. An assessment of the development, structure and use of the Wechsler Tests of Intelligence*. New York: Pergamon Press.

Frankenberger, W., & Fronzaglio, K. (1991). A review of states' criteria and procedures for identifying children with disabilities. *Journal of Learning Disabilities, 24*, 495-500.

Fuerst, D.R., Fisk, J.L., & Rourke, B.P. (1989). Psychosocial functioning of learning-disabled children: Replicability of statistically derived subtypes. *Journal of Consulting and Clinical Psychology, 57*, 275-280.

Fuerst, D.R., & Rourke, B.P. (1993). Psychosocial functioning of children: Relations between personality subtypes and academic achievement. *Journal of Abnormal Child Psychology, 21*, 597-607.

Glutting, J.J., McDermott, P.A., Prifitera, A., & McGrath, E.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*, 619-639.

Glutting, J.J., Youngstrom, E.A., Ward, T., Ward, S., & Hale, R.L. (1997). Incremental efficacy of WISC-III factor scores in predicting achievement: What do they tell us? *Psychological Assessment, 9,* 295-301.

Goh, D.S., Teslow, C.J., & Fuller, G.B. (1981). The practice of psychological assessment among school psychologists. *Professional Psychology, 12*, 696-706.

Gravetter, F.J., & Wallnau, L.B. (1988). *Statistics for the behavioral sciences. Second Edition*. St. Paul, MN: West Publishing Co.

Gridley, B.E. & Roid, G.H. (1998). The use of the WISC-III with achievement tests. In A. Prifitera and D. Saklofske (Eds.). *WISC-III Clinical Use and Interpretation: Scientist-Practitioner Perspectives* (pp. 249-288). San Diego, California: Academic Press.

Grossman, F.M., & Galvin, G.A. (1987). Clinically and theoretically derived WISC-R subtest regroupings: Predicting academic achievement in a referral population. *Psychology in the Schools, 24*, 105-110.

Guerin, D.W., Griffin, J.R., Gottfried, A.W., & Christenson, G.N. (1993). Dyslexic subtypes and severity levels: Are there gender differences? *Optometry and Vision Science, 70*, 348-351.

Habib, M. (2000). The neurological basis of developmental dyslexia. An overview and working hypothesis. *Brain, 123*, 2373-2399.

Harnadek, M.C.S., & Rourke, B.P. (1994) Principal identifying features of the Syndrome of Nonverbal Learning Disabilities in children. *Journal of Learning Disabilities, 27*, 144-154.

Hooper, S.R., & Willis, W.G. (1989). Historical foundations and definitional issues. In S.R. Hooper & W.G. Willis (Eds.). *Learning Disability Subtyping. Neuropsychological Foundations, Conceptual Models, and Issues in Clinical Differentiation* (pp. 3-19). New York: Springer-Verlag.

Humphries, T. & Bone, J. (1993). Use of IQ criteria for evaluating the uniqueness of the learning disability profile. *Journal of Learning Disabilities, 26*, 348-351.

Hynd, G.W., Cohen, M.J., Riccio, C.A., & Arceneaux, J.M. (1998). Neuropsychological basis of intelligence and the WISC-III. In A. Prifitera and D. Saklofske (Eds.). *WISC-III Clinical Use and Interpretation: Scientist-Practitioner Perspectives* (pp. 203-226). San Diego California: Academic Press.

Individuals with Disabilities Education Act of 1990 (IDEA), Public Law Number 101-476, 20, U.S.C.

Joschko, M., & Rourke, B.P. (1985). Neuropsychological subtypes of learning-disabled children who exhibit the ACID pattern on the WISC. In B.P. Rourke (Ed.) *Neuropsychology of Learning Disabilities: Essentials of Subtype Analysis* (pp. 65-88). New York: Guilford Press.

Kamphaus, R.W. (1993). *Clinical assessment of children's intelligence: A handbook for professional practice*. Needham Heights, Massachusetts: Allyn and Bacon.

Kamphaus, R.W. (1998). Intelligence test interpretation: Acting in the absence of evidence. In A. Prifitera and D. Saklofske (Eds.). *WISC-III Clinical Use and Interpretation: Scientist-Practitioner Perspectives* (pp. 40-57). San Diego, California: Academic Press.

Kamphaus, R.W., Benson, J., Hutchinson, S., & Platt, L.O. (1994). Identification of factor models for the WISC-III. *Educational and Psychological Measurement, 54*, 174-186.

Kamphaus, R.W., Di Stefano, C., & Lease, A.M. (2003). A self-report typology of behavioral adjustment for young children. *Psychological Assessment, 15,* 17-28.

Kamphaus, R.W., Frick, P.J., & Lahey, B.B. (1991). Methodological issues and learning disabilities diagnosis in clinical populations. *Journal of Learning Disabilities, 24*, 613-618.

Kass, C.E. Auditory Closure Test. In J.J. Olson, & J.L. Olson (Eds.), Validity studies on the Illinois Test of Psycholinguistic Abilities. Madison, Wisc.: Photo, 1964.

Kaufman, A.S. (1976). A new approach to the interpretation of test scatter on the WISC-R. *Journal of Learning Disabilities, 9*, 160-168.

Kaufman, A.S. (1979). *Intelligent testing with the WISC-R*. John Wiley & Sons, Inc: New Yrok.

Kaufman, A.S. (1993). King WISC the third assumes the throne. *Journal of School Psychology, 31*, 345-354.

Kaufman, A.S. (1994). *Intelligent testing with the WISC-III*. John Wiley & Sons, Inc.: New York.

Kavale, K.A., & Forness, S.R. (1984). A meta-analysis of the validity of the Wechsler scale profiles and recategorizations: Patterns or parodies. *Learning Disability Quarterly, 7*, 136-156.

Keith, T.Z., & Witta, E.L. (1997). Hierarchical and cross-age confirmatory factor analysis of the WISC-III: What does it measure? *School Psychology Quarterly, 12*, 89-107.

Kessler, J.W. (1980). History of minimal brain dysfunctions. In H.E. Rie and E.D. Rie (Eds.). *Handbook of minimal brain dysfunctions* (pp. 18-51). New York: John Wiley & Sons.

Klonoff, H. (1971). Factor analysis of a neuropsychological battery for children aged 9 to 15. *Perceptual and Motor Skills, 32*, 603-616.

Klove, H. Clinical neuropsychology. In F.M. Forster (Ed.). *The medical clinics of North America*. New York: Saunders, 1963.

Knights, R.M. & Norwood, J.A. (1980). *Revised smoothed normative data on the Neuropsychological Test Battery for Children*. Unpublished manuscript, Carleton University, Department of Psychology, Ottawa, Canada.

Konold, T.R., Glutting, J.J., McDermott, P.A., Kush, J.C., & Watkins, M.M. (1999). Structure and diagnostic benefits of a normative subtest taxonomy developed from the WISC-III standardization sample. *Journal of School Psychology, 37*, 29-48.

Konold, T.R., Kush, J.C., & Canivez, G.L. (1997). Factor replication of the WISC-III in three independent samples of children receiving special education. *Journal of Psychoeducational Assessment, 15*, 123-137.

Krug, D., Dean, R.S., & Anderson, J.L. (1995). Factor analysis of the Halstead-Reitan Neuropsychological Test Battery for Older Children. *International Journal of Neuroscience, 83*, 131-134.

Kush, J.C. (1996). Factor structure of the WISC-III for students with learning disabilities. *Journal of Psychoeducational Assessment, 14*, 32-40.

Lange, R.T., Iverson, G.T., Senior, G.J., & Chelune, G.J. (2002). A primer on cluster analysis applications to cognitive rehabilitation research. *Journal of Cognitive Rehabilitation, 20*, 16-33.

Learning Disabilities Association of Canada (2001). *Official Definition of Learning Disabilities*. Retrieved from http://www.ldac-taac.ca/english/defined/definew.htm.

Leckliter, I.N., Forster, A.A., Klonoff, H., & Knights, R.M. (1992). A review of reference group data from normal children for the Halstead-Reitan Neuropsychological Test Battery for Older Children. *The Clinical Neuropsychologist, 6*, 201-229.

Little, S.G. (1992). The WISC-III: Everything old is new again. *School Psychology Quarterly, 7*, 148-154.

Livingston, R.B., Gray, R.M., Haak, R.A., & Jenning, E. (1997). Factor structure of the Halstead-Reitan Neuropsychological Test Battery for Older Children. *Child Neuropsychology, 3*, 176-191.

Macmann, G.M., & Barnett, D.W. (1997a). A critical appraisal of intelligent testing with the WISC-III: Introduction to the series. *School Psychology Quarterly, 12*, 193-197.

Macmann, G.M., & Barnett, D.W. (1997b). Myth of the master detective: Reliability of interpretations for Kaufman's "Intelligent Testing" approach to the WISC-III. *School Psychology Quarterly, 12*, 197-234.

McDermott, P.A., Fantuzzo, J.W., & Glutting, J.J. (1990). Just say no to subtest analysis: A critique on Wechsler theory and practice. *Journal of Psychoeducational Assessment, 8*, 290-302.

McKay, M.F., Neale, M.D., & Thompson, G.B. (1985). The predictive validity of Bannatyne's WISC categories for later reading achievement. *Journal of Educational Psychology, 55*, 280-287.

Mercer, C.D., King-Sears, P, & Mercer, A. (1990). Learning disability definitions and criteria used by state education departments. *Learning Disability Quarterly, 13*, 141-152.

Meyer, M.S. (2000). The ability-achievement discrepancy: Does it contribute to an understanding of learning disabilities? *Educational Psychology Review, 12,* 315-337.

Milligan, G.W. (1981). A Monte Carlo study of thirty internal criterion measures for cluster analysis. *Psychometrika, 46*, 187-199.

Milligan, G.W., & Cooper, M.C. (1987). Methodology review: Clustering methods. *Applied Psychological Measurement, 11*, 329-354.

Morey, L.C., & Agresti, A. (1984). The measurement of classification agreement: An adjustment to the Rand statistic for chance agreement. *Educational and Psychological Measurement, 44*, 33-37.

Morris, R., Blashfield, R., & Satz, P. (1981). Neuropsychology and cluster analysis: Potential and problems. *Journal of Clinical Neuropsychology, 3*, 79-99.

Morris, R.D., & Fletcher, J.M. (1988). Classification in Neuropsychology: A theoretical framework and research paradigm. *Journal of Clinical and Experimental Neuropsychology, 10*, 640-658.

Morris, R.D., Stuebing, K.K., Fletcher, J.M., Shaywitz, S.E., Lyon, G.R., Shankweiler, D.P., Katz, L., Francis, D.J., & Shaywitz, B.A. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology, 90*, 347-373.

O'Donnell, J.P., MacGregor, L.A., Dabrowski, J.J., Oestreicher, J.M., & Romero, J.J. (1994). Construct validity of neuropsychological tests of conceptual and attentional abilities. *Journal of Clinical Psychology, 50*, 596-600.

Overall, J.E., Gibson, J.M., & Novy, D.M. (1993). Population recovery capabilities of 35 cluster analysis methods. *Journal of Clinical Psychology, 49*, 459-470.

Paget, K.D. (1982). Intellectual patterns of conduct problem children on the WISC-R. *Psychology in the Schools, 19*, 439-445.

Prifitera, A., & Dersh, J. (1993). Base rates of WISC-III diagnostic subtest patterns among normal, learning disabled and ADHD samples. *Journal of Psychoeducational Assessment Monograph Series – Advances in Psychoeducational Assessment: Wechsler Intelligence Scale for Children – Third Edition*, (pp. 43-55).

Prifitera, A. Weiss, L.G., & Saklofske, D.H. (1998). The WISC-III in context. In A. Prifitera & D.H. Saklofske (Eds.). *WISC-III clinical use and interpretation: Scientist-practitioner perspectives (pp. 1-39).* San Diego, California: Academic Press.

Reinecke, M.A., Beebe, D.W., Stein, M.A. (1999). The third factor of the WISC-III: It's (probably) not freedom from distractibility. *Journal of the American Academy of Child and Adolescent Psychiatry, 38*, 322-328.

Reitan, R.M. (1971). Trail Making Test results for normal and brain-damaged children. *Perceptual and Motor Skills, 33*, 575-581.

Reitan, R.M. (1974). Cerebral lesions in young children. In R.M. Reitan & L.A. Davison (Eds.). *Clinical Neuropsychology: Current status and applications*. Washington, D.C.: Winston & Sons.

Reitan, R.M., & Davison, L.A. (1974). *Clinical neuropsychology: Current status and applications*. Washington, D.C.: Winston & Sons.

Reitan, R.M. & Wolfson, D. (1992). *Neuropsychological evaluation of older children*. Tuscon, AR: Neuropsychology Press.

Reynolds, C.R. & Kaufman, A.R. (1990). Assessment of children's intelligence with the Wechsler Intelligence Scale for Children-Revised (WISC-R). In C.R. Reynolds and R.W. Kamphaus (Eds.). *Handbook of psychological and educational assessment of children* (pp.127-165). New York: Guilford Press.

Riccio, C.A., Cohen, M.J., Hall, J., & Ross, C.M. (1997). The third and fourth factors of the WISC-III: What they don't measure. *Journal of Psychoeducational Assessment, 15*, 27-39.

Rispens, J., van Yperen, T.A., & van Duijn, G. (1991). The irrelevance of IQ to the definition of learning disabilities: Some empirical evidence. *Journal of Learning Disabilities, 24*, 434-438.

Roid, G.H., & Worrall, W. (1997). Replication of the Wechsler Intelligence Scale for Children-Third Edition four-factor model in the Canadian normative sample. *Psychological Assessment, 9,* 512-515.

Roid, G.H., Prifitera, A. & Weiss, L.G. (1993). Replication of the WISC-III factor structure in an independent sample. In B.A.Bracken and S.R. McCallum (Eds). *Journal of Psychoeducational Assessment Monograph Series – Advances in Psychoeducational Assessment: Wechsler Intelligence Scale for Children – Third Edition*, (pp. 6-21). Brandon, VT: Clinical Psychology Publishing Co., Inc.

Rourke, B.P., (1989). *Nonverbal Learning Disabilities. The syndrome and the model.* New York: Guilford Press.

Rourke, B.P. (1998). Verbal-Performance discrepancies. In A. Prifitera and D. Saklofske (Eds.). *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 139-156). San Diego, California: Academic Press.

Rourke, B.P. (1999). Neuropsychological and psychosocial subtyping: A review of investigations within the University of Windsor Laboratory. *Canadian Psychology, 41*, 34-51.

Rourke, B.P., Bakker, D.J., Fisk, J.L., & Strang, J.D. (1983*). Child neuropsychology. An introduction to theory, research and clinical practice.* New York: Guilford Press.

Rourke, B.P., Dietrich, D.M., & Young, G.C. (1973). Significance of WISC Verbal-Performance discrepancies for younger children with learning disabilities. *Perceptual and Motor Skills, 36*, 275-282.

Rourke, B.P. & Telegdy, G.A. (1971). Lateralizing significance of WISC Verbal-Performance discrepancies for older children with learning disabilities. *Perceptual and Motor Skills, 33*, 875-883.

Rourke, B.P., Young, G.C., & Flewelling, R.W. (1971). The relationships between WISC Verbal Performance discrepancies and selected verbal, auditory-perceptual, and problem-solving abilities in children with learning disabilities. *Journal of Clinical Psychology, 27*, 475-479.

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.

Sattler, J.M. (1992). *Assessment of children – Revised and updated Third Edition.* Jerome M. Sattler, Publisher, Inc.: San Diego.

Saunders, C.D., Casey, J.E., & Jones, D.A. (2001, August). *Patterns of WISC-III performance in a heterogeneous neuropsychological sample.* Poster presented at the 109th annual meeting of the American Psychological Association, San Francisco.

Saunders, C.D., Strang, J.D., & Jones, D.A. (2001, August). *Subtypes of WIAT performance in a heterogeneous neuropsychological sample.* Poster presented at the 109th annual meeting of the American Psychological Association, San Francisco.

Schuerholz, L.J., Harris, E.J., Baumgardner, T.L., Reiss, A.L., Freund, L.S., Church, R.P., Mohr, J., & Denckla, M.B. (1995). An analysis of two discrepancy-based models and a processing-deficit approach in identifying learning disabilities. *Journal of Learning Disabilities, 28*, 18-29.

Shaywitz, S.E., Shaywitz, B.A., Fletcher, J.M., & Escobar, M.D. (1990). Prevalence of reading disability in boys and girls: Results of the Connecticut Longitudinal Study. *Journal of the American Medical Association, 264*, 998-1002.

Siegel, L.S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22*, 469-478.

Slate, J.R. (1994). WISC-III correlations with the WIAT. *Psychology in the Schools, 31*, 278-285.

Slate, J.R. (1995a). Two investigations of the validity of the WISC-III. *Psychological Reports, 76*, 299-306.

Slate, J.R., Jones, C.H., & Saarnio, D.A. (1997). WISC-III IQ scores and special education diagnosis. *The Journal of Psychology, 131*, 119-120.

Smith, T.D., & Smith, B.L. (1998). Relationship between the Wide Range Achievement Test 3 and the Wechsler Individual Achievement Test. *Psychological Reports, 83*, 963-967.

Smith, T., Smith, B., Matthews, N., & Kennedy, S. (1993). Subtest scatter and Kaufman regroupings on the WISC-R in non-learning-disabled and learning-disabled children. *Psychology in the Schools, 30*, 24-28.

Snow, J.B. & Sapp, G.L. (2000). WISC-III subtest patterns of ADHD and normal samples. *Psychological Reports, 87*, 759-765.

Speece, D.L. & Shekitka, L. (2002). How should reading disabilities be operationalized? A survey of experts. *Learning Disabilities Research and Practice, 17*, 118-123.

Stanford, G., & Oakland, T. (2000). Cognitive deficits underlying learning disabilities. Research perspectives from the United States. *School Psychology International, 21*, 306-321.

Statistics Canada's Health and Activity Limitation Survey (1991). Retrieved January 2, 2004, from the Learning Disabilities Association of Canada Web site: http://www.ldac-taac.ca/english/indepth/bkground/stats01.htm.

Sternberg, R.J., & Grigorenko, E.L. (2002). Difference scores in the identification of children with learning disabilities. It's time to use a different method. *Journal of School Psychology, 40*, 65-83.

Stinnett, T.A., Havey, J.M., & Oehler-Stinnett, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment, 12*, 331-350.

Tabachnick, B.G., & Fidell, L.S. (1996). *Using Multivariate Statistics.* (3rd ed.). New York, NY: Harper Collins College Publishers.

Torgesen, J.K. (1991). Learning disabilities: Historical and conceptual issues. In B.Y.L. Wong (Ed.). *Learning about Learning Disabilities* (pp. 3-37). Academic Press: San Diego, California.

U.S. Department of Health, Education and Welfare. (1977). *Education of handicapped children: Implementation of Part B of the Education for All Handicapped Children Act.* Federal Register, 42, 42474-42518.

Ward, T.J., Ward, S.B., Glutting, J.J., & Hatt, C.B. (1999). Exceptional LD profile types for the WISC-III and WIAT. *School Psychology Review, 28*, 629-643.

Ward, S.B., Ward, T.J, Jr., Hatt, C.V., Young, D.L., & Mollner, N.R. (1995). The incidence and utility of the ACID, ACIDS, and SCAD profiles in a referred population. *Psychology in the Schools, 32*, 267-276.

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

Waxman, R.S., Casey, J.E., & Fuerst, D.R. (2003, August). *Replication of an empirically derived WISC-III taxonomy*. Poster presented at the 111th Annual Meeting of the American Psychological Association, Toronto.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children – Revised*. New York: Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children – Third Edition: Manual.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1992). *Wechsler Individual Achievement Test: Manual.* San Antonio, TX: Psychological Corporation.

Weiss, L.G., & Prifitera, A. (1995). An evaluation of differential prediction of WIAT achievement scores from WISC-III FSIQ across ethnic and gender groups. *Journal of School Psychology, 33*, 297-304.

Weiss, L.G., Prifitera, A., & Roid, G. (1993). The WISC-III and the fairness of predicting achievement across ethnic and gender groups. In B.A.Bracken and S.R. McCallum (Eds). *Journal of Psychoeducational Assessment Monograph Series – Advances in Psychoeducational Assessment: Wechsler Intelligence Scale for Children – Third Edition,* (pp. 36-42). Brandon, VT: Clinical Psychology Publishing Co., Inc.

Yawny, P., Casey, J.E., & King, B. (2001). *WISC-III Factor structure in a neuropsychological sample.* Paper presented at the meeting of the International Neuropsychological Society, Chicago, IL.

VITA AUCTORIS

Rosemary Susan Waxman (nee Schorr) was born on March 22, 1971 in Toronto, Ontario. She attended Concordia University in Montreal, Quebec, where she was conferred the degree of Bachelor of Arts (Honours) in Psychology in June 1994. Following graduation, Rosemary moved to Edmonton, Alberta and was employed as a Psychometrist at the University of Alberta Hospital, Keegan Psychological Services, and the Canadian Study of Health and Aging II. Rosemary returned to Toronto in the summer of 1997 where she worked as a Research Assistant at the Rotman Research Institute of Baycrest Centre for Geriatric Care. In 1998, Rosemary commenced graduate studies in Clinical Neuropsychology at the University of Windsor, completing her Master of Arts degree in 2001. In September 2002, she proposed the present study. At the same time, Rosemary began a predoctoral internship at the Bloorview MacMillan Children's Centre (2002-2003) in Toronto, Ontario, where she continues to be employed. Rosemary defended this dissertation on March 18, 2004.