THE UNIVERSITY OF CALGARY

The Use of the Canadian Cognitive Abilities Test (CCAT) Administered Off-Level to Predict Future Achievement of Gifted Students

by

Victoria Lynn Plouffe

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DIVISION OF APPLIED PSYCHOLOGY

CALGARY, ALBERTA

APRIL, 2004

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "The Use of the Canadian Cognitive Abilities Test (CCAT) Administered Off-Level to Predict Future Achievement of Gifted Students" submitted by Victoria Lynn Plouffe in partial fulfilment of the requirements of the degree of Master of Science.

Michael C. Pyryt

Supervisor, Dr. Michael Pyryt, Division of Applied Psychology

Dr. Tanya Beran, Division of Applied Psychology

Dr. B. Lynn Bosetti, Graduate Division of Educational Research

04/04/23

Date

Abstract

The purpose of the current study was to examine whether the Canadian Cognitive Abilities Test (CCAT) administered off-level predicts future academic achievement of young gifted students. Archival data consisting of CCAT scores as well as Grade 3 and Grade 6 Provincial Achievement Test (PAT) scores were collected from 367 students attending a school for gifted children. Results indicated that the various batteries of the CCAT administered off-level do predict future academic achievement as evidenced by PAT scores. This study suggested that the concept of administering a test designed for older students to younger capable students has additional benefits aside from predicting future academic achievement.

Acknowledgements

I would like to begin by extending my sincere appreciation to my supervisor, Dr. Michael Pyryt, for guiding and supporting me throughout my thesis completion. His passion for gifted education has allowed me to develop a greater understanding of this unique population.

I also appreciate the efforts of Dr. Tanya Beran and Dr. Lynn Bosetti for reviewing my thesis as well as making my oral examination a rewarding learning experience.

I would like to acknowledge Westmount Charter School for allowing me to access the archival data used in this investigation.

I am also appreciative of my family and friends for being extremely supportive throughout this educational journey.

Finally, my love and gratitude is extended to my husband, Gordon Plouffe, for his unconditional love and support.

Dedication

To my wonderful husband,

Gordon Plouffe

Table of Contents

List of Tables

Chapter 1

Introduction

Purpose of the Study

Gifted students are an invaluable resource to our global community. Their intellectual ability, creative thinking, and domain specific talents will challenge and shape our future knowledge, thinking, and culture. In recognizing the potential contributions of our gifted community, it only stands to reason that these exceptional individuals need to be fostered and nurtured in their development. The first step in developing intellectual giftedness is to identify those individuals with potential talent.

The purpose of this study is to investigate whether employing the concept of off-level testing is effective in identifying intellectual talent among early elementary aged students. The effectiveness of this method will be judged on its predictive value of later academic achievement for these students.

Rationale for the Study

The Talent Search model developed by Julian Stanley and his colleagues in 1971 has been characterised as "one of the most successful ideas in the field of gifted education" (Van Tassel-Baska, 1998, p. 139). Although simplistic in design, the basic premise of the Talent Search model is to administer a more difficult test normed on older students to younger able students in order to identify academic precocity (Stanley, 1996).

The Talent Search concept has been widely implemented and practised throughout the United States of America (Van Tassel-Baska, 1999) with a tremendous

growth rate. Today four regional centers (The Center for Talented Youth (CTY) at The Johns Hopkins University, The Talent Identification Program (TIP) at Duke University, The Center for Talent Development (CTY) at Northwestern University, and The Rocky Mountain Talent Search (RMTS) at the University of Denver enable more than 300,000 students to participate in annual Talent Searches (Stanley & Brody, 2001).

Although the Talent Search concept is widely used in the United States of America, the same cannot be said for Canada. In fact, the only Canadian Talent Search recognized in the reviewed literature was sponsored by the University of Calgary (Lupkowski-Shoplik, Benbow, Assouline, & Brody, 2003). As such the current study will add to the paucity of Canadian based literature in this area.

The Talent Search concept of off-level testing has been well established as a positive identification tool for intellectual talent of students in later elementary and junior high school (Benbow, Lubinski, & Sanjani, 1999; Lubinski, Webb, Morelock, Benbow, 2001; Stanley, 1999). However, its use with students in early elementary school is an area that appears to be less studied. As such, this study will add empirical research in an area that is currently less well established.

Additionally, it has been suggested that any standardized test can be used for the purpose of off-level testing provided that it has a high enough ceiling to clearly differentiate among highly able students (George, 1979). As such, this study will examine whether the Canadian Cognitive Abilities Test (CCAT) is an appropriate off-level testing instrument.

Finally, researchers have primarily taken a retrospective approach (i.e. giving parents of gifted children questionnaires and /or checklists asking about aspects of early

development) when attempting to identify characteristics of promising young children (Perleth, Schatz, & Mönks, 2000). This study however, will prospectively examine whether the psychometric approach of using a standardized test off-level can predict the achievement of young talented students.

Research Questions

The effectiveness of off-level testing derived from the Talent Search model, to identify intellectual talent evidenced through academic achievement among early elementary aged students will be examined through the following research questions: The primary research question of this study asks:

1. Does the use of a specific standardized test, specifically the Canadian Cognitive Abilities Test (CCAT), administered off-level significantly predict future academic achievement of gifted young students?

Additional questions of this study ask:

1. How well does a young child's standard score on the Verbal battery of the Canadian Cognitive Abilities Test administered off-level predict their actual achievement as well as their range of achievement on the Language Arts portion of the Alberta Provincial Achievement Tests?
2. How well does a young child's standard score on the Non-Verbal battery of the Canadian Cognitive Abilities Test administered off-level predict their actual achievement as well as their range of achievement on the Language Arts portion of the Alberta Provincial Achievement Tests?

3. How well does a young child's standard score on the Quantitative battery of the Canadian Cognitive Abilities Test administered off-level predict their actual achievement as well as their range of achievement on the Mathematics portion of the Alberta Provincial Achievement Tests?
4. How well does a young child's standard score on the Non-Verbal battery of the Canadian Cognitive Abilities Test administered off-level predict their actual achievement as well as their range of achievement on the Mathematics portion of the Alberta Provincial Achievement Tests?
5. If the Canadian Cognitive Abilities Test administered off-level is predictive of later academic achievement, are the results more predictive of earlier (Grade 3 Provincial Achievement Test) results or later (Grade 6 Provincial Achievement Test) results?

These research questions will be examined through a statistical analysis of the data obtained by reviewing sample students' cumulative academic files and obtaining the following information: (a) standard scores on the three batteries of the Canadian Cognitive Abilities Test (Thorndike & Hagen, 1998) administered off-level; (b) Grade 3 Provincial Achievement Test scores; and (c) Grade 6 Provincial Achievement Test scores.

Definition of Terms

In the interest of clarity, the following section discusses conceptualisations and definitions of terms that may assist the reader in understanding the current study.

For the purpose of this study, giftedness is conceptualised as development that is advanced for the individual's chronological age (Stanley, 1991). It is recognized that advanced development may be evidenced in a variety of domains (Gardner, 1983; 1993). This study will specifically examine intellectual development in the area of Language Arts and Mathematics. In regards to the present study, giftedness will be used synonymously with talent.

It is noted that many definitions of intelligence exist (Sattler, 2001). For the purpose of this study, intelligence refers to an individual's general and specific cognitive abilities as measured by a standardized test, specifically the Canadian Cognitive Abilities Test (Thorndike & Hagen, 1998).

Academic achievement refers to what an individual has learned (Sattler, 2001). In this study, academic achievement will be measured by subjects' scores on the Grade 3 and Grade 6 Alberta Provincial Achievement Tests.

Identification refers to "the process of recognizing students' needs, strengths, talents, and interests, in order to design effective ways to nurture and enhance their potentials" (Alberta Learning, 2000, p. 34). This study will use off-level testing as an identification method.

Off-level testing refers to a process whereby a standardized test normed on older children is administered to younger capable students. A test is considered off-level if it is administered at least two years above a student's chronological age (Tsai, 2001).

In this study, early elementary aged students refers to those students in Grades 1 through 3.

Chapter 2

Literature Review

This chapter has four intentions. Firstly, this chapter will provide an overview of the general conceptual understandings of intelligence and their interrelationship with various models of giftedness. Secondly, various methods for identifying above-average intelligence, including standardized intelligence and achievement tests and multiple criteria procedures will be addressed. The identification of giftedness among young children will also be discussed. Thirdly, this chapter will review the Talent Search model, including relevant historical contributions and rationale for the use of off-level testing. It will also examine empirical studies stemming from the Talent Search model and off-level testing. Off-level testing will be suggested as a method for identifying intellectual giftedness among early elementary aged students. Finally, this chapter will conclude by detailing the purpose of this study.

Theories of Intelligence

There is an interrelationship between intelligence and giftedness, whereby various intelligence theories influence the conceptualizations, identification procedures, and programming methods pertaining to gifted education (Assouline, 1997). The concept of intelligence and the definition of giftedness or talent is an evolving debate among scholars. Throughout history, many definitions of intelligence have been conceptualized. A review of the literature suggests that contemporary theories of intelligence accept the

notion of general intelligence, while maintaining that multi-dimensional factors may contribute to intelligent behaviour (Carroll, 1993; Sattler, 2001).

General intellectual or mental ability is often referred to as "g", a term first proposed by Spearman (1927), which is usually measured by a variety of intelligence tests. In his book *Abilities of Man*, Spearman (1927) metaphorically referred to "g" as mental energy. He further clarified this concept by stating that "g" "is some force capable of being transferred from one mental operation to another different one" (p. 414). Although not described using today's terminology, many early psychologists, including Spearman, recognized that general intelligence is comprised of specialized abilities (Carroll, 1993).

Galton (1822-1911) proposed that individuals possess a general intellectual ability, which presents itself in their mental abilities and special aptitudes (Sattler, 2001). General ability can be explained as the power of the mind affecting everything that a person does, whereas the specialized aptitudes may be in areas such as linguistics, mathematics, or artistic pursuits (Jensen, 1998). Galton was interested in studying individual cognitive differences and suggested that a series of mental tests be given to people to determine their mental strengths and weaknesses (Carroll, 1993). Despite his interest, Galton was unsuccessful in measuring individual differences in intelligence (Carroll, 1993; Jensen, 1998).

Binet (1857-1911) regarded general intelligence as a high level of cognitive ability being made up of judgement, practical sense, initiative, and the ability to adapt to circumstances (Sattler, 2001). Binet has been credited with developing the first test of intelligence making general cognitive ability quantifiable. This was an important

movement in the field of psychology as it paved the way for the identification of intellectual giftedness to begin in earnest (Achter & Lubinski, 2003).

Terman (1916) defined general intelligence as the capacity to form concepts and grasp their significance. He used IQ scores as the criterion for identification of gifted individuals in his longitudinal studies on eminence. Terman's research subjects were "...within the top 1% in general intelligence..." (Terman & Oden, 1947, p. 22). Terman proclaimed that anyone scoring 140 on the Stanford-Binet was a genius, as the mean score was 100. Although Terman's initial definition of giftedness was based primarily on general intelligence, after 30 years of follow-up studies, he conceded that social and emotional confidence combined with a drive to achieve were also important determiners of achievement (Terman & Oden, 1947).

Similarly, Hollingworth (1926) defined general intelligence as the "power to achieve literacy and to deal with its abstract knowledge and symbols" (Pritchard, 1951, p. 49). She also used intelligence tests as a measure for identifying and defining highly intelligent children. Despite setting a strict cut-off score for high general intelligence, Hollingworth wrote:

> By a gifted child we mean one who is far more educable than the generality of children are. This greater educability may lie along the lines of one of the arts, as in music or drawing; it may lie in the sphere of mechanical aptitude; or it may consist in surpassing power to achieve literacy and abstract intelligence (as cited in Pritchard, 1951, p. 49).

Gottfredson (2003) defined "g" in a more global sense by stating, "g" is the ability to process information of any sort." (p. 26). It is manifested in daily life as the ability to learn, reason, and solve problems. She explained that the existence of a strong

"g" factor does not mean that intelligence is a unitary construct, but it forms the core for all mental abilities.

By recognizing that general intelligence has a relationship with many special abilities, aptitudes, achievements, and skills (Carroll, 1993), the current discussion must recognize explicit factor-analytic theories of intelligence. It is beyond the scope of this discussion to appropriately address the statistical underpinnings behind factor analysis; however it is understood that factor-analysis "makes it possible to determine the degree to which each of the variables is correlated (or loaded) with the factor that is common to all the variables in the analysis" (Jensen, 1998, p. 18).

Spearman (1863-1945) was an early advocate of this approach to understanding intelligence. He postulated a two-factor theory of intelligence whereby a general factor "g" and one or more specific factors "s" would account for performance on intelligence tests (Spearman, 1927). He arrived at this conclusion after realizing that the correlation between abilities, as evidenced through test scores, could consistently be divided into the two aforementioned parts. Spearman (1927) defined the two factors as follows:

> The one part has been called the "general factor" and denoted by the letter "g"; it is so named because, although varying freely from individual to individual, it remains the same for any one individual in respect to all of the correlated abilities.. The second part has been called the "specific factor" and denoted by the letter "s". It not only varies from individual to individual, but even for any one individual from each ability to another (p. 75).

Thurstone (1887-1955) moved away from Spearman's dual factor method of analysis by analyzing multiple factors in order to study cognitive abilities in various domains (Carroll, 1993). According to Thurstone's multidimensional theory of

intelligence, intelligence can be divided into multiple factors (verbal, perceptual speed, inductive reasoning, rote memory, deductive reasoning, word fluency, and space or visualization) all of which have equal weight (Sattler, 2001).

Guilford's structure of intellect model is considered an extension of Thurstone's theory of intelligence as it incorporates all seven of Thurston's multiple factors, as well as adds new factors, bringing the total number of factors to 120 (Sternberg & Grigorenko, 2001). Guilford's model of intelligence does not acknowledge "g", but presents three facets of ability: Operations (cognition, memory, divergent production, convergent production, and evaluation), Contents (visual, auditory, symbolic, semantic, and behavioural), and Products (units, classes, relations, systems, transformations, and implications). Within the model, a distinct ability is defined by the intersection of a particular form of each facet (Guilford, 1985). According to Guilford (1985) each of these facets is required in every mental activity.

Vernon (1950) placed "g" in a hierarchical structure with "g" being the highest level. Vernon theorized that major (verbal-educational and spatial-mechanical) and minor (creative, verbal-fluency, numerical, spatial, psychomotor, and mechanical) group factors contribute to "g" and thus must be considered when attempting to understand or measure intelligence.

After reanalyzing more than 460 data sets found in factor-analytic research, Carroll (1993) postulated his three-stratum theory of intelligence. This factor analytic theory of cognitive abilities suggests that the relationship among individual differences in cognitive abilities can be stratified on three levels. The first level of the theory (Narrow-Stratum I) consists of abilities at a mastery level in various cognitive areas or

"specializations of ability" (Carroll, 1993, p. 634). The second level (Broad –Stratum II) is made up of eight factors that "are very general abilities that lie in broad domains of behaviour" (Carroll, 1993, p. 633). Finally, the third level (General – Stratum III) represents general intelligence.

Sternberg proposed the triarchic theory of intelligence. This theory is comprised of three subtheories, serving as the basis for understanding intelligence from his perspective. According to Sternberg (1986), intelligence must be viewed from the internal world of the individual, the external world of the individual, and the interaction between these two worlds of the individual's experience. The internal world can be exemplified by analytical thinking which addresses the mental mechanisms that lead to intelligent behaviour. The external world is characterized by intelligent behaviour in the everyday world through three classes of acts – environmental adaptation, selection, and shaping. Finally the interactive world can be thought of as a continuum of experiences with tasks or situations that involve the use of intelligence. Within this componential theory, the mechanisms or components by which intelligent behaviour is generated are specified. According to Sternberg (1986), components perform three functions.

> Metacomponents are higher-order processes used in planning, monitoring, and decision making in task performance. Performance components are processes used in the execution of a task. Knowledge components are processes used in learning new things (p. 225).

In generalizing his view of intelligence, Sternberg (2003) stated that intelligence is, "the ability to achieve success in life in term's of one's personal standards, within one's sociocultural context" (p 88).

Similarly to Sternberg's componential based triarchic theory of intelligence, Gardner developed his theory of Multiple Intelligences (MI) by combining studies of the brain with research on the contextual aspects of intelligence. In his book *Frames of Mind*, Gardner (1983) proposed that there is not one independent intelligence, but rather that the human intellect is made up of "...several relatively autonomous human intellectual competencies" (p. 8). The MI theory currently identifies the following intelligence areas: verbal-linguistic, logical-mathematical, bodily-kinesthetic, musical-rhythmic, visual-spatial, intrapersonal, interpersonal, naturalist, and existentialist. According to Gardner (1993), intelligences are manifested in different ways at different developmental levels. According to Gardner's (1983) theory of Multiple Intelligences an individual with intelligence or intellectual competency is one who has the ability to solve problems and to fashion products (p. 61). The above prerequisites to intelligence are enhanced by one's creativity. Gardner (1993) has asserted that creativity crosses all domains of intelligence, but he has questioned to what extent creativity should be fostered in students by structured classes.

The PASS theory of intelligence proposed by Das, Naglieri, and Kirby (1994) views intelligence in terms of four major cognitive functions: planning processes, attentional processes, simultaneous processing, and successive processing. According to the PASS theory, these cognitive processes are the essential elements involved in intellectual functioning, thus relating mental abilities to a theory of neurological functions.

> Planning is a mental process by which the individual determines, selects, applies, and evaluates solutions to problems...Attention is a mental process by which the

> individual selectively focuses on particular stimuli while inhibiting responses to competing stimuli presented over time...Simultaneous processing is the a mental process by which the individual integrates separate stimuli into a single whole or group...Successive processing is a mental process by which the individual integrates stimuli into a specific serial order that forms a chain-like progression (Naglieri & Das, 1997, p. 2-5).

Summary

The selected theories of intelligence presented suggest that a common conceptualization of intelligence is non-existent as theories are continually evolving. Although theorists may not agree whether intelligence is "a general unitary function or a composite of several independent abilities" (Sattler, 2001, p. 137), they do appear to all recognize that intelligence is a compilation of one's learning and experience. Moreover, "there is no contradiction in recognizing both general and specific mental abilities as long as one bears in mind the fact that mental abilities below the level of general intelligence are positively correlated" (Deary, Austin, & Caryl, 2000, p. 182).

Models of Giftedness

Prior to discussing the various theories of intelligence, it was stated that theories of intelligence influence the field of giftedness in many ways. This is a reasonable influence given that the conception of what constitutes giftedness has evolved and is explicitly linked to intelligence. It has changed from a uni-dimensional view linking giftedness to high general intelligence equated with a high IQ score to a multi-dimensional conception acknowledging outstanding abilities in different domains. Considering this paradigm shift, Alberta Learning (2000) has adopted the following

definition of giftedness. "Giftedness is exceptional potential and/or performance across a wide range of abilities in one or more of the following areas: general intellectual, specific academic, creative thinking, social, musical, artistic, and/or kinaesthetic" (p. 17). This definition has been influenced by various theoretical models.

Gardner's theory of Multiple Intelligences has been transformed into a model for gifted education (Alberta Learning, 2000). Ramos-Ford and Gardner (1997) stated that "MI theory is not so much a prescription for educational programs, with a set curriculum or methodology that must be carried out in all settings, but rather a provocation for educators and others to think about the assessment and education of all individuals" (p. 59). With this in mind, the MI based model of gifted education should address a broad range of human abilities and talents, in comparison to the linguistic and mathematical domains typically focussed on in schools. Such models should also emphasize learning in context, assess each intelligence directly as part of the learning experience, and understand the distinctive combination of intelligences that make up a learner (Blythe & Gardner, 1990).

In his model of human intelligence, Sternberg (2003) discriminates between three kinds of intellectual giftedness: analytic giftedness, synthetic giftedness, and practical giftedness. Analytic giftedness involves being able to dissect a problem and understand its parts. Synthetic giftedness involves being insightful, creative, and adept at coping with novel situations. Finally, practical giftedness involves applying whatever analytic and synthetic ability one may have, to everyday, pragmatic situations. According to Sternberg (1997) "an important part of giftedness lies in being able to coordinate these three aspects of abilities and in knowing when to use each one" (p. 44). Sternberg's model is

considered a cognitive component model approach, which is "task analytic and attempts to directly identify the information processing components of performance on tasks that have been generally used to assess mental abilities" (Pellegrino & Glaser, 1979 p. 188). Such models have been criticized for their primarily theoretical basis and relative lack of empirical support (Mönks & Mason, 1993). Countering such criticism, Sternberg, Grigorenko, Jarvin, Clinkenbeard, Ferrari, and Torff (2000) published an article validating the effectiveness of Sternberg's model. The article summarized some of the main findings of recent research. Overall it was concluded that instructional interventions can be used to improve analytical skills, creative, and practical thinking skills and more importantly, Sternberg's model can be applied to improve students' achievement in school.

Renzulli's (1978) three-ring model of giftedness suggests that giftedness is the result of an interaction between three clusters of human traits: above average ability, task commitment, and creativity. Renzulli (1978) stressed the fact that each cluster is an equal partner in determining giftedness, which may be represented by behaviours in general and specific performance areas. According to Gagné (1985), Renzulli's model is inapplicable to gifted underachievers as motivation is considered a critical component of giftedness. Gagné (1985) also faults the three-ring model for including the identification of creativity as an essential component. He proposes that "creativity may be regarded as a major determinant of exceptional performance in certain fields of endeavour, but not in all" (p. 106). Finally, Gagné (1985) criticizes Renzulli's model for not differentiating above average ability into separate ability domains.

The three-ring model of giftedness however, has formed the rationale for the Enrichment Triad Model which "was designed to encourage creative productivity on the part of young people by exposing them to various topics, areas of interest, and fields of study, and to further train them to apply advanced content, process training skills, and methodology training to self-selected areas of interest" (Renzulli & Reis, 2003, p. 186). A further manifestation places the Enrichment Triad model as the final component of the Schoolwide Enrichment Model (SEM) which provides a detailed plan for the development of talents and gifts and encourage creative productivity in students (Renzulli & Reis, 2003). Renzulli and Reis (2003) summarized numerous field tests applying the SEM, concluding that this model can successfully be implemented in a variety of settings and used with various populations of students.

Like Renzulli, Treffinger defines giftedness by addressing the same three attributes of giftedness: above average ability, task commitment, and creativity (Alberta Learning, 2000). This definition of giftedness is important to Treffinger's (1986) Individualized Programming Planning Model (IPPM) as "it establishes a context through which specific characteristics can be determined and examined" (p. 140). The focal point of this model is to provide effective independent learning based on individual student's strengths and talents. In addition to IPPM, Treffinger collaborated with Feldhusen on the Talent Identification and Development in Education (TIDE) model (Alberta Learning, 2000). Essentially this model "attempts to foster talent development in all youth rather than focusing on a small percentage of students" by recognizing that "talent development involves a combination of aptitudes, acquired knowledge and personality characteristics" (Alberta Learning, 2000, p. 32).

Van Tassel-Baska (1996) would agree with such models of giftedness arguing that the conception of high intelligence has changed from an emphasis on global ability or general intelligence to an emphasis on talent development. As such, in her view the current notion is that there is a paradigm shift from intellectual giftedness to talent development. Likewise, through his work with the Study of Mathematically Precocious Youth (SMPY), Julian Stanley extended the concept of general intelligence to major group factors for identifying and developing more specific intellectual areas of strength or talent (Achter & Lubinski, 2003). Giftedness is therefore synonymous with precocity. Stanley's model will be elaborated upon further in this chapter.

Gagné's Differentiated Model of Giftedness and Talent (DMGT) clearly distinguishes between the concepts of giftedness and talent (Gagné, 2000). According to Gagné (2000),

> The term 'giftedness' designates the possession and use of untrained and spontaneously expressed natural abilities (called aptitudes or gifts) in at least one ability domain, to a degree that places an individual within the top 10% of age peers. By contrast, the term 'talent' designates the superior mastery of systematically developed abilities (or skills) and knowledge in at least one field of human activity, to a degree that places the individual within the top 10% of age peers who are (or have been) active in that field (p. 67).

Within his model, Gagné (2000) recognizes four aptitude domains: intellectual, creative, socio-affective, and sensorimotor. He also acknowledges as many areas of talent as there are areas of human activity, recognizing that natural abilities and aptitudes form the constituent elements of talent development. This development is either facilitated or hindered by various catalysts, which can be either interpersonal and/or environmental (Gagné, 2000).

Summary

The above citations demonstrate that models of giftedness are related to theories of intelligence, recognizing that giftedness, like intelligence is not a unitary construct, but rather is a combination of various domains. Common to all models presented is the recognition of ability with the discrepancy being how ability is differentiated. Moreover, these models seek to identify the potential of individuals while acknowledging the processes required for achievement - the observable output of giftedness (Mönks & Mason, 1993).

Methods for Identifying Giftedness

Standardized Testing

Cognitive based measures of assessment such as achievement or intelligence tests are the most traditional and widely utilized screening measures for identification of intellectually gifted students (Callahan, Hunsaker, Adams, Moore, & Bland, 1995; Hadaway & Marek-Schroer, 1992; Pyryt, 1996) as these instruments "remain a most technologically advanced and sophisticated tool for providing essential and unique information to best serve the needs of children" (Kaufman, Lichtenberger, & Naglieri, 1999, p. 308). However, exclusive reliance on standardized test scores to determine talent has received criticism because it is believed that potentially gifted students may be overlooked (Reis & Renzulli, 1992; Sternberg, 1986; Von Karolyi, Ramos-Ford, & Gardner, 2003). According to such opponents, a high score on a standardized intelligence or achievement test warrants further consideration, but a low score should not rule out the possibility of intellectual giftedness.

Traditional IQ tests such as the Weschler Intelligence Scales and the Stanford-Binet measure the general ability construct of intelligence and were developed using contemporary psychological theories (Naglieri & Das, 1997) with high scores being significantly related to achievement (Kaufman, et al., 1999; Naglieri & Das, 1997). Assouline (2003) reports that individually administered intelligence tests remain the best instrument for identifying gifted students on the criterion of general ability. Similarly, McGrew, Keith, Flanagan, and Vanderwood (1997) purport that individually administered standardized tests provide reliable and valid estimates of an individual's global level of cognitive functioning.

In her discussion regarding the assessment of gifted children, Assouline (1997) suggested that scores obtained from well designed tests "can provide relevant and useful information about behaviours that typically are associated with intelligence" (p. 91). Sattler (1988) supported this view, adding that intelligence tests provide an overview of an individual's range of knowledge and cognitive skills at a given point in time. A review of twenty-two empirical studies addressing giftedness conducted by Tannenbaum (1992) found that every study listed IQ as the measure of choice for identifying giftedness (as cited in Tannenbaum, 2003).

According to Sternberg (1999), IQ scores reveal abilities that are "general only with respect to the academic or analytical aspect of intelligence. Once one includes creative and practical abilities in an assessment the general factor is greatly diminished or disappears" (p. 16). As such, intelligence measures only reveal a limited amount of information. Winner (1999) agreed that IQ tests measure only a narrow range of skills that are relevant for mastery in school curricula.

Terman was also sensitive to the limitations of measuring intelligence and cautioned:

> We must guard against defining intelligence solely in terms of ability to pass the tests of a given intelligence scale. It should go without saying that no existing scale is capable of adequately measuring the ability to deal with all the possible kinds of material on all intelligence levels (as cited in Assouline, 2003, p. 125).

A study by Tyler-Wood and Carrie (1991) examined the scores students obtained on four different standardized cognitive abilities instruments (Stanford-Binet [LM], Stanford-Binet [4th Edition], the Otis-Lennon School Abilities Test, and the Cognitive Abilities Test) to determine if the scores on each measure were consistent. It was determined that students' scores were significantly different across measures. This prompted the researchers to caution that as student scores ranged significantly from one standardized measure to the next, using a single test of cognitive ability can have significant ramifications for the identification of gifted students (Tyler-Wood & Carrie, 1991), whereby a specific criterion or definition of giftedness must be established and test norms should be considered.

Sternberg (2003) believes that, "the idea of the testing is to expand our notion of giftedness and then be able to identify as gifted those individuals who may be adept in skills that are not measured by conventional tests" (p. 94). Sternberg (2003) supports the use of the Sternberg Triarchic Abilities Test as a method to identify gifted students. He reports that it measures analytical, practical, and creative abilities by providing seven separate scores: analytic, synthetic, automization, practical abilities, verbal, quantitative, and figural processing

(Sternberg, 2003, p. 94). According to Sternberg (2003) analytic abilities are generally not enough to make important contributions, there is a need for synthetic and practical thinking as well. Although supported by Sternberg, a review of the literature revealed that as of yet this newly developed test has received limited usage.

Like Sternberg, Gardner suggests that IQ tests should only be one factor in the identification of gifted individuals (Von Karolyi, et al., 2003). The theory of MI does not endorse any particular test, but advocates for employing multiple approaches in one's identification strategy (Von Karolyi, et al., 2003). Specific to the theory of MI, researchers concede that, "being gifted in one intelligence does not, and should not, foretell (or preclude) a similarly high level of ability on another area" (Von Karolyi, et al., 2003, p. 103). Similarly, it cannot be assumed that, "a child who performs poorly on an IQ test or a standardized achievement test will fail to excel in activities relying on one or more of the other intelligences" (Von Karolyi, et al., 2003, p. 103).

Project Spectrum is an educational program inspired by the theory of Multiple Intelligences. It "...provides a rich child-friendly environment in which a wide range of abilities in young children, can be comfortably and unobtrusively observed and assessed in a meaningful context" (Von Karolyi, et al., 2003, p. 106). For each domain or intelligence area an assessment measure has been developed to determine the child's level of intelligence.

The preliminary results from the Project Spectrum research suggest that domain specific strengths and weaknesses can be identified in young children. For example, when children were tested on specific intelligence area activities, a child could be labelled as

high functioning in visual-spatial and low functioning in the logical-mathematical domain, dependent on their scores. As well, a child's performance in one area may enhance performance in another. A child's ability to relate to music may be beneficial in the musical-rhythmic area, but may also be relied on to enhance performance in the bodily-kinesthetic area. In follow-up studies, the intelligence areas identified as high for children continued to be areas of strength (Gardner, 1993).

DISCOVER (Discovering Intellectual Strengths and Capabilities through Observation while allowing for Varied Ethnic Responses) is a performance based assessment grounded in Gardner's theory of MI (Sarouphim, 2000). Initial reliability and validity studies of DISCOVER revealed that interobserver reliability was high, although limited support was found for the concurrent validity of the instrument (Sarouphim, 1999). Similarly, a study conducted by Sarouphim in 2000 to investigate the internal structure of DISCOVER concluded that more research is required with this instrument before it can be positively recommended, as the results of the study were inconclusive.

Multiple Criteria Identification

In response to a more broadened view of giftedness, current identification procedures are being guided by the use of a variety of objective and subjective techniques for identification. Frasier (1997) referred to the use of a variety of methods for identification as multiple criteria. Specifically,

> Multiple criteria refers to the process of obtaining comprehensive information about a student's abilities by gathering and interpreting results from: standardized measures of aptitude, achievement, and creativity; observations by teachers, parents, the student, and others e.g. community members who are familiar with the student;

> and standardized evaluations of student products and performances e.g. juried performances, portfolios (p. 2).

The intent of the use of multiple criteria is to give professionals the most complete picture of the student and to allow many ways for the student to exhibit talent (Callahan, Hunsaker, Adams, Moore, & Bland, 1995).

The Frasier-Talent Assessment Profile (F-TAP) is designed to facilitate the collection, display, and interpretation of data from multiple test and non-test sources. Studies conducted evaluating this model have concluded that the F-TAP effectively facilitates the collection and use of information from multiple sources when making decisions about potentially gifted students (Frasier, 1997). Frasier (1997) reports that the performance of students identified through the application of this more comprehensive assessment model equals the performance of students identified by more traditional methods.

Jenkins-Friedman (1982) argues that ways must be found to tolerate a degree of subjectivity if we are to respect the multi-dimensional nature of giftedness that accompanies our acceptance of a broadened view of intelligence and giftedness. Assessment using only traditional intelligence tests appears questionable for assessing across all the constructs within the broadened conception of giftedness and inappropriate for use in identifying ability in specific academic areas, the arts, creativity, or leadership (Callahan, et al., 1995).

There is little research available supporting multiple criteria measures. The literature that is available is preliminary in nature and cites philosophical approaches. As there is little research in the area, there is concern over the reliability and validity of using

multiple criteria for identification. Less traditional aspects of giftedness do not lend themselves readily to standardization and quantification (Goldberg, 1986, as cited in Callahan et al., 1995).

In a similar vein, Richert (2003) notes that the trend to use data from a variety of sources can be counter-productive. The data, "may be unreliable, used at an inappropriate stage or sequence in the identification process, weighted in indefensible ways, or placed without validity in a matrix with other data (Richert, 2003, p. 148). The widely used practice of giving equal weighting to a variety of sources (standardized tests, grades, observations, checklists, etc.) for identification of gifted students is problematic (Richert, 2003).

Summary

Despite broadening definitions and understandings of intelligence and giftedness, standardized testing methods continue to hold prominence in terms of identification of gifted individuals (Feldhusen & Jarwan, 2000; Pyryt, 1996). Multiple criteria procedures have been suggested as a way to address ability theories and models that stress multiple factors (Gardner, 1991). According to Feldhusen and Jarwan (2000) "the task of identification shifts from a search for the gifted few to assessment of talent strengths and aptitudes of all students, and to identification of high level talent potential among those who are especially precocious or advanced in their talent development" (p. 279).

Identifying Giftedness in Young Children

Identifying giftedness or talent becomes an even more convoluted issue as the age of the child decreases. Early identification as it has been termed, is hampered by the fact

that physical, social, and cognitive development in young children is rapid and uneven (Smutney, 1999). Perleth, et al. (2000) question whether it is even necessary to identify talented children before school age, arguing that the primary goal during the early years should be optimal child development. Moreover, psychometric measures for identifying giftedness in young children have been viewed as controversial (Perleth, et al., 2000).

In regards to psychometric instruments, Clarke (2001) noted that although the psychometric view of giftedness is evident, this quality is not directly relevant to preschoolers. Sattler (1988) would agree, purporting that high reliability and stability of measured traits of preschoolers can not be assumed. Reliability, norms, and predictive validity are the test properties considered the most important for the identification of giftedness, yet there is a paucity of empirical research with preschool and early elementary school children addressing these test factors when identifying giftedness (Perleth, et al., 2000). Furthermore, Perleth, et al. (2000) stated that "nearly no longitudinal studies have been conducted which followed preschool and early elementary school children for a minimum of several years to analyze possible indicators and determinants of giftedness and achievement" (p. 299).

Perleth, Lehwald, and Browden (1993) commented on a few studies addressing the Stanford-Binet Intelligence Scale (SB) and the Weschler Intelligence Scale for Children (WISC) as psychometric instruments for identifying giftedness in young children. Robinson's (1987) study was cited which reported a retest correlation of $r = 0.75$ for preschoolers assessed at age 2 and 6 ($n=16$), $r = 0.59$ between children of age 3 and 6 ($n=117$), $r = 0.61$ between age 4 and 6 ($n=74$), and $r = 0.71$ between age 5 and 6 ($n=25$). Despite Robinson concluding that these results were low, Perleth, et al.

considered these correlations to indicate that intelligence was a stable trait as measured by the Stanford-Binet. A discriminant analysis study conducted by Shapiro, Palmer, Antell, Bilker, Ross, and Capute (1989) attempted to predict IQ scores as measured by the WISC-R at age 7.5 with developmental indicators assessed at 7-12 months and the SB (Form LM) scores at 3 years of age. As the correlations were medium, the authors concluded that it was not possible to predict giftedness at the age of 7.5 with the SB. They further stated that "the expectation for precocious infants to become precocious children, while appealing, is not supportable" (as cited in Perleth, et al., 2000, p. 300).

Other standardized tests such as the Kaufman Assessment Battery for Children, Tests of Nonverbal Reasoning Abilities, and/or the Peabody Picture Vocabulary Test are also instruments that are sometimes used to identify or screen for young gifted students (Perleth, et al., 2000). Although such tests are available and used with preschoolers and early elementary students, the research remains sceptical as to their validity, especially in predicting high ability and high achievement from an early age (Perleth, et al., 1993). Perleth, et al. (2000) acknowledge that psychometric instruments have their limitations for use with younger children, but advocate that further evaluation of well-known tests should take precedence over the construction of new instruments for the identification of young gifted children.

Similar to the multiple criteria methods suggested for identifying older gifted students, Clarke (2001) suggest that several different techniques should be combined to identify giftedness among the very young. She lists the following possibilities for the identification of preschool and early elementary students who may be gifted and talented: biographical data, checklists and rating scales, specific area achievement testing, group

intelligence testing, individual intelligence testing, formal and informal observations, parent nominations, peer nominations, work and behaviour samples, personal interviews, and professional judgements. Perleth, et al. (2000) cautioned that if a variety of indicators of giftedness are used for identification, "effective and practical statistical methods for the combination and weighting of the information have to be used" (p. 309).

Others such as Jackson (2003), believe that intellectual giftedness in young children may be better identified by their behaviour. She draws on the work of Sternberg and Zhang (1995 as cited in Jackson, 2003) suggesting that gifted behaviour in young children must be characterized by five qualities. Specifically,

> It must be excellent, relative to the performance of peers who are the same age or who have had the same degree of instruction; it must be rare among the same peers; it must be demonstrable on some reliable and valid assessment instrument; it must be productive or suggest potential for productivity; and it must have some special societal value (p. 470).

Such behaviours often include oral language production and comprehension or precocious reading and writing abilities (Jackson, 2003). Even if a young child does demonstrate remarkable behaviours at a young age, Jackson (2003) cautions that it must be questioned whether the behaviour is a result of genetic disposition or of an exceptionally supportive environment. Thus, it becomes difficult to tease out true gifted behaviour from an environmental advantage.

It is without question, that controversy exists regarding identifying giftedness in young children. Some question whether there is even a need at such a young age (Perleth, et al., 2000), while others question the methodology behind suggested identification procedures (Clarke, 2001: Perleth, et al., 1993).

The Talent Search Model

Overview

Julian C. Stanley has been recognized as the founder of the Talent Search concept, along with being distinguished as the most notable psychologist in the 20th century to study intellectually precocious youth (Achter & Lubinski, 2003). In 1972, Stanley pioneered the Talent Search concept by offering a challenging test designed for older students to capable younger students. The intended purpose was to identify exceptional intellectual talent in a specific domain. As mentioned previously, Stanley supported the idea that intellectual giftedness among youth is best construed in terms of precocity, or development that is advanced for its age (Benbow & Stanley, 1996), making giftedness synonymous with talent. Stanley further supported the notion that talent can be identified and nurtured in specific domains (Benbow & Stanley, 1996; Stanley, 1984). The identification component of the Talent Search concept (further explained below) allows for the testing of specific areas of aptitude rather than addressing only global intelligence (Van Tassel-Baska, 1996).

In 1972, Stanley held the first large scale Talent Search at The Johns Hopkins University, deemed the Study of Mathematically Precocious Youth (SMPY) with 450 examinees (Stanley, 1999). Using the Scholastic Aptitude Test (SAT) designed for Grade 12 students, he sought to identify advanced mathematical reasoning ability in seventh and eighth grade students (Lupkowski-Shoplik, et al., 2003).

The Talent Search concept is based on a two-step process. The first step is an initial screening designed to identify students who may benefit from participation in the off-level assessment of the Talent Search testing. This initial screening is typically

comprised of a grade-equivalent standardized test. A typical test used is the Iowa Test of Basic Skills (Lupkowski-Shoplik et al., 2003). Students who perform well on this initial screening test (usually at the 95th or 97th percentile) are then encouraged to move on to the second step in the process, an off-level test of their aptitude.

The second component of the Talent Search is to administer the specified off-level test to the eligible students. "A test is above [off]-level if it is at least two years above the student's present grade placement" (Tsai, 2001, p. 449). The goal of the Talent Search model is not just to identify talented individuals, but to provide them with educational experiences that will foster their outstanding ability (Cohn, 1991).

The educational options available to students who have been identified as talented through participation in a Talent Search have been referred to as a "smorgasbord" of choices (Lupkowski-Shoplik, et al., 2003; Stanley, 1999). These choices may include: grade skipping, taking selected classes with an older age group, independent studying/tutoring in advanced subject matter, testing out of courses, distance learning courses; fast-paced classes or compressed curricula, the International Baccalaureate Program, advanced placement courses, summer courses, dual enrolment in high school and college, early entrance to college, or concurrent undergraduate and graduate programs (Lupkowski-Shoplik, et al., 2003, p. 208).

VanTassel-Baska (1996) further highlighted the benefits to of the programs and opportunities resulting from the Talent Search model.

> [T]he direct benefits to students as a result of these programs is enormous. For many it becomes a way: (1) to earn advanced high-school and even college credit in an economical manner, (2) to associate for the first time with an equally able peer group, (3) to develop the "habits of

> mind" associated with serious study on a college campus, and (4) to gain a sense of academic competence through receiving worthy challenge to learn more difficult material. This list does not mention the enormous personal gains in the area of social and emotional development that these students seem to experience as judged by anecdotal reports (p. 238).

In support of identifying intellectual talents of gifted children, Achter and Lubinski (2003) stated that "gifted children do not always find their own way; rather, they are more likely to achieve at exceptional levels when given appropriate educational and environmental opportunities" (p. 27).

Currently it is estimated that 300,000 students at the sixth to eighth grade levels participate in annual Talent Searches in the United States (Lupkowski-Shoplik, et al., 2003), and as a result are privy to the aforementioned educational options and program opportunities. Talent searches remain the most popular in the United States, although the model is expanding internationally (Pyryt, Colangelo, Assouline & Ihrig, 2003). The University of New South Wales in Sydney Australia and Dublin City University in Dublin Ireland host annual talent searches (Lupkowski-Shoplik et al., 2003). The Center for Gifted Education at the University of Calgary in Calgary Alberta conducts the only Talent Search in Canada (Pyryt, et al., 2003).

Contributions to the Talent Search Concept

It is without question, that there are several noteworthy pioneers who have significantly contributed to the field of giftedness. However, there are a few individuals in particular that can be credited with providing the foundational underpinnings of the Talent Search concept.

Lewis Terman may be one of the most recognized figures in the study of intellectual talent (Stanley, 1984). Terman used IQ scores as the criterion for identification of gifted individuals in his longitudinal studies on eminence (1925-1956). In his ambitious study, Terman followed 1528 children from the age of 11, who had IQ scores of 140 or higher (within top 1% in general intelligence), to document their success (Terman & Oden, 1947).

Components of Terman's study are still in operation today, and many findings have been documented as a result of the research stemming from this study. Terman's research, "provided a solid foundation for much of what we know today about the physical and psychological development of gifted persons over the lifespan" (Achter & Lubinski, 2003, p. 30). Terman's work substantiated the predictive power of general intelligence (Terman & Oden , 1947) and further advocated for the use of curriculum acceleration and enrichment to adequately meet the needs of gifted students (Achter & Lubinski, 2003).

A contemporary to Terman, Leta Hollingworth also used intelligence tests as a measure for identifying and defining gifted children. Hollingsworth was an advocate for identifying intellectual talent at an early age and many of her ideas were considered revolutionary (Stanley, 1990). She felt that students with profoundly high IQ's deserved an education that would enhance their talent. As such, she felt an appropriate response would be to allow gifted students to be instructed according to their competence rather than their chronological age (Stanley, 1990).

Influenced by the work of Terman and Hollingworth, Stanley and his colleagues were the first to systematically use college entrance exams to differentiate ability levels

of gifted students (Stanley, 1999). Not rejecting the construct of general intelligence, Stanley sought to "...exten[d] the psychometric approach to major group factors for identifying and developing more specific intellectual strengths"(Achter & Lubinski, 2003, p. 35). Stanley (1984) advocated for such an approach because it has the potential to provide valuable information and a foundation to focus future development for talented students, their parents, their teachers, and significant others in their lives by providing a clearer pictures of student's intellectual strengths and comparative weaknesses. As Achter, Benbow, and Lubinski (1997) noted, this same information could not be gleaned from performance on grade-level achievement tests or general intelligence tests.

Rationale for Off-Level Testing

The purpose of off-level testing is to allow those students who have reached the ceiling on an in-grade or grade-equivalent achievement test to demonstrate their advanced abilities. "Administering an above [off]-level test to students at the upper-end of the bell curve helps discriminate able students from exceptionally able students, and it provides a more precise assessment of aptitude and readiness for additional academic challenges" (Lupkowski-Shoplik, et al., 2003, p. 205). It is important to determine the extent of a student's abilities as different courses of action need to be taken for very able versus extremely able students. For example, education acceleration is appropriate for only extremely able students (Benbow, 1991).

"One of the characteristics of a well-designed standardized test is that its score distribution follows a bell curve" (Tsai, 2001, p. 449). Most commonly the population of gifted and talented students fall at the upper end of the bell curve on grade-level standardized achievement tests (Tsai, 2001). It has been rationalised that standardized

tests traditionally used to assess academic achievement of school children have a ceiling effect that is too low to identify those students whose talents are exceptional and thus deserve special educational opportunities (Goldstein, Stocking, & Godfrey, 1999). Similarly, Olszewski-Kubilius (1998) argued that "in-grade achievement tests are useful in determining how the overall achievement of a school compares to others from the nation or what percentage of the student body is achieving at certain levels. It is not helpful, however, in determining educational placement for the group of students who score very well on the test because there is a great deal of variability among this group of students" (p. 107).

Although it is exceptional for a child to score at the 99th percentile on a standardized test, there is still the question of the vast range of abilities within this top 1%. Benbow and Lubinski (1993) noted that the range of scores for individuals in the top 1% is as broad a range as that from the bottom 2% to the top 2%. An efficient way to acquire more information about students at the extreme ends of the curve is to use tests that are out-of-level (Assouline & Lupkowski-Shoplik, 1997). The above-level test creates a new distribution of scores for those students at the upper percentiles.

Administering a test off-level then is a logical solution to a mental measurement problem – How does one assess the intellectual capabilities of the extremely able? (Cohn, 1991). Standard achievement tests fail to provide the required analysis of students as they require prior knowledge that some children may not have. By using tests normed for older students it allows the truly exceptional students the opportunity to adequately demonstrate the full scope of their abilities. In turn, the Talent Search model was developed as a systematic assessment approach that effectively uses off-level testing to

discriminate able students from exceptionally able students (Lupkowski-Shoplik, et al., 2003).

Common Tests Used Off-Level

There are several assessment instruments that are commonly used in off-level testing. The uses of these tests are consistent with the theoretical considerations outlined by the Talent Search model. The SAT was the original instrument used by Stanley in the SMPY and is still the most widely used Talent Search instrument for middle school students (Lupkowski-Shoplik, et al., 2003). Other recognized and commonly used instruments include the ACT Assessment, School and College Abilities Test, Secondary School Admission Test, and EXPLORE. Each of these instruments will be discussed.

Scholastic Aptitude Test (SAT)

As mentioned the Scholastic Aptitude Test (SAT) was the original test chosen by Stanley to identify high-level mathematical reasoning abilities in exceptional students. For the purpose of his study, he used the Mathematics section of this test for identification of talented students because he deemed the test to be professionally prepared, standardized, secure, reliable, and difficult enough (Assouline & Lupkowski-Shoplik, 1997). The primary purpose of this test is to assess the mathematical and verbal reasoning abilities of university-bound eleventh and twelfth graders. In 1994 the SAT was revised and renamed the Scholastic Assessment Test (SAT-I). The original SAT was comprised of three areas: Mathematics, Verbal, and Test of Written English. The revised version contains a Mathematical as well as a Verbal section. Scores on the Mathematical and Verbal sections of the SAT are reported on a scale of 200 to 800, with Talent Search participants usually scoring across the full range. (Lupkowski-Shoplik, et al., 2003).

Since a full range of scores is represented by this gifted group, the test is neither too easy (ceiling effect) nor too hard (floor effect) and thus is appropriate from a psychometric point of view (Benbow, 1991).

ACT Assessment

The ACT Assessment (formally the American College Testing Program) was designed as a university entrance examination. Currently it includes four tests of educational development: mathematics, English, reading, and science reasoning. Assouline and Lupkowski-Shoplik (1997) noted that "the questions on these tests are designed to tap students' knowledge and skills in major curriculum areas and therefore measure educational progress in curriculum-related areas" (p.173). Scores from these subtests can be averaged to produce a composite score ranging from 1 to 36. The efficacy of this assessment instrument as a Talent Search tool is supported by the fact that the average college student score on each subtest is 21. In comparison, approximately 35 percent of Talent Search participants earned scaled scores that were equal to or surpassed those of the senior college student (Sawyer, 1995 as cited in Assouline & Lupkowski-Shoplik, 1997). This test is commonly used as a Talent Search assessment tool to identify academically talented seventh and eighth graders.

School College Abilities Test (SCAT)

The School and College Abilities Test was mentioned by J. Cohn (1991) as a possible identification tool of exceptionally bright elementary aged students (Grades 2-4). SCAT includes two subtests that measure quantitative and verbal reasoning ability and is available in three levels: elementary, intermediate, and advanced. The predictive validity of this test is evident, as it has been found that children who perform well on SCAT also

perform well in advanced-level, fast-paced courses of study (Cohn, 1991). This test is used by the Academic Talent Search at California State University Sacramento, as well as by the Iowa State University's Talent Search (Lupkowski-Shoplik, et al., 2003).

Secondary School Admission Test (SSAT)

Educational Testing Service developed the Secondary School Admission Test (SSAT). It is a multiple-choice, secure test that contains quantitative, verbal, and reading comprehension sections. There are two levels of this test: Lower Level (developed for fifth to seventh graders) and Upper Level (developed for eighth to eleventh graders). A study by Lupkowski-Shoplik & Assouline (1993) used the SSAT off-level to identify talented elementary students (third through fifth grade). They found that the SSAT was an appropriate off-level test as the sample student scores covered almost the entire possible scaled score range.

EXPLORE

EXPLORE was developed by American College Testing to measure eighth grade students' curriculum related knowledge as well as complex cognitive skills (Lupkowski-Shopliket al., 2003). EXPLORE is made up of four multiple-choice sections (English, mathematics, reading, and science reasoning), with scores reported on a scale of 1-25 for each test. This test was first used as an off-level Talent Search instrument for third to sixth grade students by Carnegie Mellon University and the University of Iowa and is now also used for similar purposes by Duke University and Northwestern University (Lupkowski-Shoplik, 2001). As such, EXPLORE is now the primary instrument used in the search for academically talented elementary students, as it is appropriately difficult

for fourth and fifth grade students who are in the top 5% of ability in their respective age groups (Lupkowski-Shoplik, 1999).

Empirical Research Supporting Off-Level Testing

The Talent Search concept is beginning to amass a considerable amount of empirical support, including longitudinal studies (Benbow, et al., 1999; Lubinski, et al., 2001; Stanley, 1999) and predictive research (Assouline, & Lupkowski-Shoplik, 1997; Benbow, et al., 1999; Cohn, 1991; Lupkowski-Shoplik, et al., 2003; Lupkowski-Shoplik, & Swiatek, 1999).

Stanley's Study of Mathematically Precocious Youth (SMPY) has significantly added to researchers, educators, and those involved with talented students' understanding of effective mechanisms that promote intellectual and social well-being among talented students so that their talent potential becomes a reality (Benbow, et al., 1999). Numerous studies have documented research findings stemming from SMPY (Benbow, et al., 1999; Lubinski, et al., 2001).

In a longitudinal study based on SMPY spanning 20 years, researchers sought to identify our top scientists and engineers and determine their educational experiences. For this study, Cohort 5 of the SMPY, composed of 715 graduate students from the top U.S. mathematics and physical science departments comprised the sample. The mean score on the quantitative section of the Graduate Record Exam approached 750, with the highest possible score being 800. This indicates that these students were reaching, if not surpassing, the ceiling of this test. Of these high scoring graduate students, 93 % felt they would have been identified as talented through a Talent Search. Characteristics of students from Cohort 5 were then compared with students identified by SMPY at age 12

who later went into engineering and physical sciences. The most important finding of this comparison was that students who have the potential to become our nation's great scientific achievers can be identified by the age of 13 (Benbow, et al., 1999). Specifically, students identified by the SAT at age 12 or 13 do disproportionately enter the math or science fields. It was also found that these individuals seize educational opportunities (i.e. acceleration) and work hard at developing their talents (Benbow, et al., 1999).

A complimentary 10-year follow-up study with 320 profoundly gifted individuals (IQ=180+) found that identifying precocity during early adolescence significantly isolates a population at promise for exceptional adult achievement and creative production. Within this study, 56% of the participants were pursuing doctorates at prestigious universities, many had published scientific and literary works, and a number had secured prestigious awards. Additionally, it was reported that 255 of the participants took advantage of tailored educational opportunities such as advanced credit, specialized course work, grade skipping, etc. (Lubinski, et al., 2001).

Research with the SAT has shown that talented individuals can be identified before the age of 13 and that the SAT has predictive validity for achievements in college, graduate schools, and careers (Lupkowski-Shoplik, et al., 2003). A study by Benbow, Lubinski, Shea, and Eftekhari-Sanjani (2000) reported that talented individuals identified by SMPY demonstrated exceptional academic achievements, with 90% earning a bachelor's degree, and over 25 % earning a doctorate degree. Cohn (1991) reported that students who do well on the SAT also do well in advanced-level course work, concurrent

enrolment in programs, such as high school and college, and special accelerated summer programs, all of which require accelerated passage through difficult material.

Mills and Barnett (1992) conducted a pilot study whereby the upper level of the SSAT (SSAT-U) was administered to fifth and sixth grade students. They reported that SSAT-U effectively differentiated highly able students. In a concurrent study by Assouline and Lupkowski (1992), the quantitative section of the lower level of the SSAT (SSAT-L) was given to students in Grades 2 through 6. Once again it was reported that this test effectively differentiated among the range of abilities of the previously identified academically able students. The SSAT-L was further supported as off-level testing instrument in a follow-up study by Lupkowski-Shoplik (1992). The lower level of the SSAT was administered to 520 students in Grades 3, 4, and 5. As a result of the study, the authors concluded that the SSAT-L as an off-level test was an efficient for identifying talented elementary students and had an adequate ceiling.

A study by Assouline and Lupkowski-Shoplik (1997) determined that sixth grade students participating in the University of Iowa's Talent Search performed better than the eighth grade norm group in all sections of EXPLORE, with the exception of Mathematics. Results were similar for fifth grade students, however fourth grade students were close to the average eighth grade norms on all sections except Mathematics and third grade students were lower than the norms on all sections. These results suggest that EXPLORE facilitates the identification of high-ability elementary students. A similar study was conducted in 1999 by Lupkowski-Shoplik and Swiatek to determine if the 95th percentile qualifying cut-off score was appropriate for third to sixth graders who take the EXPLORE test. The data obtained suggested that EXPLORE was an appropriate off-level

test for gifted elementary students because, "it raises the ceiling of in-grade tests in a way that allows gifted students to obtain a more accurate measure of their abilities" (p. 269). Looking specifically at the Mathematics portion of this test, Lupkowski-Shoplik (1999) found that, "EXPLORE is an appropriately difficult test fourth and fifth grade students in the top five percent of ability in their respective age groups" (p. 463).

In a similar vein, results from the Belin Exceptional Talent Search (BESTS) conducted in Alberta Canada has found that over a five year period 91.3 % of the participating third, fourth, fifth, and sixth grade participants "have scores that exceed the performance of the normative sample of eighth graders in the United States on at least one of the EXPLORE tests or composites" (Pyryt, et al., 2003, p. 6).

Summary

It is clear that much has been learned from the Talent Search model and the subsequent research findings. Julian Stanley developed this concept and was the first to systematically apply it in his Study of Mathematically Precocious Youth.

A key element of the Talent Search model is the identification process. In this unique process, students are screened using a grade-equivalent standardized achievement test. Those individuals scoring in the upper percentiles (usually 95^{th} to 97^{th}) are encouraged to take an off-level test. The off-level tests helps to further discriminate between able and highly-able students.

Numerous tests have been cited in the literature as appropriate off-level instruments, including Scholastic Aptitude Test, ACT Assessment, School and College Abilities Test, Secondary School Admission Test, and EXPLORE. Although the above instruments were specifically reviewed, there is flexibility in the choice of an off-level

test. George (1979) noted that the off-level instrument may vary as long as the ceiling is high enough to differentiate among the upper 2-5% of the initially screened students.

As a result of the research in this area, it is acknowledged that over the past three decades the Talent Search concept has had positive implications for many outstanding students. Most notably, the identification of profound talent in students has lead to educational enrichment and accelerative options that may not have otherwise been possible. It is also recognised that with early identification, this talent can be appropriately nurtured at a young age.

It is only equitable then that as a society who values excellence, we recognize and reward it in our youth. Thus, a consideration for the further development of the Talent Search concept would be to apply it to the identification of even younger precocious youth. Hollingworth (as cited in Assouline & Lupkowski, 1992) would agree with this suggestion. She specifically noted that, "it is within the primary and elementary school that the very intelligent child most especially needs a supplement to the standard curriculum" (p. 223). Currently, the Talent Search concept is most commonly applied to late elementary, early junior high age students, with a few experimental searches being conducted with youth as young as second grade (Lupkowski-Shoplik, & Assouline, 1993; Lupkowski-Shoplik & Swiatek, 1999).

Required Research

There is a paucity of empirical research addressing the identification of intellectual talent among early elementary aged students using the Talent Search model.

Similarly, there is a need for longitudinal research addressing the long-term predictive significance of young children's demonstration of giftedness.

The purpose of this study then, is to investigate whether employing a derivation of Stanley's Talent Search concept of off-level testing is efficacious in identifying intellectual talent among early elementary aged students. Specifically, this study asks whether a test designed for older students can be given to younger students to predict achievement at a later date. It has been suggested that if extreme talent in individuals is recognized at the same time that formal cognitive operations of the mind develop (Cohn, 1991) (Grades 1 through 3), then educational enrichment opportunities could also be initiated at this point.

Chapter 3

Method

Sample

The convenience sample, also referred to as the complete database sample, consisted of 367 students who were currently or previously enrolled at Westmount Charter School in Calgary, Alberta. Westmount Charter School is a K-10 school for gifted students. Giftedness is defined through the school's charter as a dynamic, multifaceted, multidimensional phenomenon. It involves a record of creative accomplishments over an extended period of time, and consists of an interaction among three basic clusters of human traits: well above average ability (in one or more of the intelligences), high levels of task commitment, and high levels of creativity. The school began as ABC Charter Public School in 1996 with 116 students in Grades 1-3. In 2002, the name officially changed to Westmount Charter School growing to its current student population of 814. As part of the admission requirements of this school, students are required to take the Canadian Cognitive Abilities Test (CCAT) administered off-level prior to acceptance. As such, each of the students included in the complete database had available CCAT scores. Additionally, available Grade 3 and Grade 6 Provincial Achievement Test (PAT) scores were collected for students in the complete database sample ($n=367$). Those students with complete CCAT and Grade 3 and 6 PAT scores made up the subsequent analysis sample of this study ($n=123$). As the data were collected anonymously, the precise gender ratio of either sample group is not known, nor is the

chronological ages of the participants. An analysis of the school's enrolment history revealed that during the years examined in this study the distribution of males and females was approximately equal. The distribution of all students who wrote the various levels of the CCAT is shown in the following table:

Table 1

Distribution of CCAT Levels Administered

CCAT Level Administered	Number of Students
Level 2	198
Level A	57
Level B	48
Level D	28
Level E	25
Level F	10

(*n=367*)

Measures

Two sources of achievement results were utilized. The Canadian Cognitive Abilities Test – Form K was used as the off-level testing instrument. Student scores on this instrument were then correlated to student scores on their Grade 3 and Grade 6 Alberta Provincial Achievement Tests.

The Canadian Cognitive Abilities Test (CCAT)

The Canadian Cognitive Abilities Test (CCAT) (1994) is described by its authors (Robert L. Thorndike & Elizabeth P. Hagen) as, "an integrated series of tests that provides information on the level of development of general and specific cognitive skills

of students from kindergarten through grade twelve" (p. 1). The Canadian version of the Cognitive Abilities Test, the CCAT's primary purpose is to appraise individual differences of students and provide a description of the student's cognitive resources for learning. This test is a group-administered, norm-referenced testing instrument appropriate for students in grades K-12. Levels 1 and 2 of the CCAT are most commonly used for students in Kindergarten and Grade 1 and Levels A through H are most commonly used for students in Grades 2 through 12. Each level of the test has a mean of 100 and a standard deviation of 16.

The CCAT assesses the development of three major cognitive abilities: Verbal, Quantitative, and Non-Verbal. The Verbal battery assesses verbal inductive and deductive reasoning, problem solving, and verbal comprehension. The Verbal battery subtests are Verbal Classification, Sentence Completion, and Verbal Analogies. The Quantitative battery assesses general abstract reasoning skills and specific mathematical skills. The Quantitative battery subtests are Quantitative Relations, Number Series, and Equation Building. The Non-Verbal battery assesses inductive reasoning skills. The Non-Verbal battery subtests are Figure Classification, Figure Analogies, and Figure Analysis.

The development of this testing instrument was based on Vernon's (1961) and Cattell's (1987) theoretical models of human abilities (Thorndike & Hagen, 1998). Both of these models stress "g" or general overall reasoning skills as the cognitive construct central to all learning and problem solving. The CCAT defines "g" as abstract reasoning skill with additional emphasis on inductive reasoning (Thorndike & Hagen, 1998). Aside from recognizing overall general cognitive ability, the CCAT is constructed to appraise specific ability factors differentiated by the three distinct batteries.

Form K of the CCAT was re-normed in the Spring of 1997. The standardization sample consisted of approximately 30,000 students drawn from a stratified random sample of Canadian schools in which English was the major language of instruction. A review of the literature revealed that the CCAT has not been the focal point of much empirical research. According to the CCAT Technical Manual, reliability estimates are about .91 for the Verbal battery, .90 for the Quantitative battery, and .91 for the Non-Verbal battery. There were no current reviews of the CCAT in the Mental Measurements Yearbook, although two reviews were available for the previous version, Form 7. These reviews were mixed with Anderson (1995) finding the test to be a useful instrument for providing information on a student's level of general functioning and Hattie (1995) concluding that the CCAT does not reflect good measurement practice as time required to administer the test does not justify the lack of information provided.

Thorndike and Hagen (1998) have suggested that this testing instrument can be appropriately used as an off-level assessment measure. In fact they state that one of the appropriate uses of the CCAT is, "to determine eligibility to gifted programs" (Thorndike & Hagen, 1995 p. 10). As mentioned previously, students in this charter school are required to take the CCAT administered off-level as a precursor to acceptance. Levels 2, A and B are administered two years above grade equivalency and Levels D, E and F are administered three years above grade equivalency.

Alberta Provincial Achievement Tests (PAT)

Each year in the province of Alberta, students in Grades 3, 6, and 9, write provincial achievement tests. As outlined by Alberta Learning (2002) the purpose of the Alberta Achievement testing program is to:

• determine if students are learning what they are expected to learn

• report to Albertans how well students have achieved provincial standards at given points in their schooling

• assist schools, authorities, and the province in monitoring and improving student learning (p. 2).

Students in Grade 3 annually write tests in English Language Arts and English and French Mathematics, whereas students in Grade 6 and 9 write tests in English and French Language Arts, Mathematics, Social and Science. The tests vary in format with the Grade 3 Language Arts possible raw score being 100 and Mathematics possible raw scores being 43. The possible raw scores for the Grade 6 tests are 100 in Language Arts, 54 in Mathematics, 50 in Science and 50 in Social Studies.

The Provincial Achievement Tests are criterion referenced and are based on specific learner outcomes of each grade level.

> School authorities and schools report results on all performance measures in their education plans, including the following required provincial measures: the percentage of students writing achievement tests in grades 3, 6, and 9 who achieved the *acceptable standard* and the percentage who achieved the *standard of excellence* in relation to school targets, provincial results, and provincial standards, for the past five years (Alberta Learning, 2002, p. 3).

The standard set for the province is that 85% of students are expected to meet the *acceptable standard* and that 15% are also expected to meet the *standard of excellence.*

Data Collection Procedure

The researcher established a general database at Westmount Charter School by gathering archival data on students' CCAT scores at the time of admission acceptance

into the school. Additionally Grade 3 and Grade 6 PAT scores were documented. All of the aforementioned information was collected by reviewing students' cumulative record files at the school. Cumulative records for students attending the school from 1996-2002 were accessed for review.

Data Analysis

Data collected from the original database were entered into a spreadsheet and then analyzed using the Statistical Package for the Social Sciences (SPSS). Data analysis in this investigation included the following:

(1) Descriptive statistics were calculated to summarize the means and standard deviations of the standard scores of the three batteries of the CCAT (Verbal, Quantitative, and Non-Verbal) and the standard scores of the PAT (Grade 3 L.A., Grade 3 Mathematics, Grade 6 L.A., Grade 6 Mathematics). Descriptive statistics were generated for both the complete database group as well as the analysis sample.

(2) Pearson product-moment correlations were performed to examine the degree of relationship between a student's score on the various batteries of the CCAT and their subsequent achievement as measured by their PAT scores in Grade 3 and 6. Correlations were determined for both their actual achievement as well as their range of achievement.

The higher the value of the resulting correlation coefficient the more accurately one can predict scores on one measure from scores on another measure. Values can range from +1.00 to -1.00 indicating both the strength and direction of the relationship.

Sattler (2001) suggested that a correlation coefficient of .50, either positive or negative, indicates a moderate to strong relationship between the two variable. Alternatively, Cohen (1983) suggested that a correlation coefficient of .10 is considered small, a correlation coefficient of .30 is considered medium, and a correlation coefficient of .50 is considered large.

Chapter 4

Results

Introduction

This chapter reports the results of the data analysis. The first section presents the descriptive statistics for the complete database sample on the CCAT standard scores and the PAT actual achievement scores. Descriptive statistics are also presented for the same measures of the analysis sample. An overview of the percentage of students in the complete database achieving at the various achievement standard levels on the PAT is also presented. The second section of this chapter reports the relationship between the CCAT standard scores and the PAT actual scores as indicated by the results of Pearson product-moment correlations. Similar correlations are also presented for the CCAT standard scores and PAT achievement ranges.

Descriptive Statistics

The complete database sample (*n=367*) consisted of students who attended Westmount Charter School between 1996-2002, whereas the analysis sample (*n=123*) is made up of students from the entire database who had available data for all variables under investigation. The mean scores and standard deviations obtained by the complete database sample on the CCAT and PAT are presented in Table 2.

Table 2

Means and Standard Deviations for the CCAT and the PAT for the Complete Sample

	Number	Mean	Standard Deviation
CCAT			
Verbal Battery	366	104.92	15.17
Non-Verbal Battery	366	105.04	14.55
Quantitative Battery	366	108.53	16.01
PAT			
Grade 3 Language Arts	325	79.07	10.21
Grade 3 Mathematics	320	36.72	6.32
Grade 6 Language Arts	150	75.52	11.19
Grade 6 Mathematics	160	44.76	6.43

Based on the complete database sample, the CCAT standard scores ranged from 53.65 to 146.35 on all three batteries. These results indicate that on each battery of the CCAT, scores ranged from $\leq 1^{st}$ percentile to $\geq 99^{th}$ percentile when administered off-level. A review of the frequency tables showed that on the Verbal battery one student scored $\leq 1^{st}$ percentile and one student scored $\geq 99^{th}$ percentile. On the Non-Verbal battery, two students scored $\leq 1^{st}$ percentile and one student scored $\geq 99^{th}$ percentile. On the Quantitative battery, five students scored $\leq 1^{st}$ percentile and six students scored $\geq 99^{th}$ percentile.

On the PAT, the Grade 3 Language Arts scores ranged from 38 to 100, with the mean score being 79%. Grade 3 Mathematics scores ranged from 14 to 89, with the mean score being 85%. Grade 6 Language Arts scores ranged from 32 to 98, with the mean score being 76% and scores in Grade 6 Mathematics ranged from 22 to 54, with the mean score being 83%. In all cases of the PAT scores for the complete database sample, mean scores were within the Acceptable Standard range, with the Grade 3 Mathematics mean score being in the Standard of Excellence.

The mean scores and standard deviations obtained by the analysis sample on the CCAT and PAT are presented in Table 3.

Table 3

Means and Standard Deviations for the CCAT and the PAT for the Analysis Sample

	Mean	Standard Deviation
CCAT		
Verbal Battery	106.21	14.57
Non-Verbal Battery	106.54	16.08
Quantitative Battery	108.80	17.12
PAT		
Grade 3 Language Arts	80.07	10.08
Grade 3 Mathematics	37.49	4.54
Grade 6 Language Arts	76.06	10.90
Grade 6 Mathematics	45.08	5.92

(n=123)

Based on the analysis sample, standard scores the CCAT Verbal battery and the Quantitative battery ranged from 50.56 to 149.44, indicating a full range of scores from $\leq$ 1st percentile to $\geq$ 99th percentile. On the Non-Verbal battery standard scores ranged from 62.78 to 149.44. As the mean scores in all cases were above the standardization mean of 100, it can be generalized that students' abilities as measured by the CCAT lie 2-3 years above their chronological age.

On the PAT, the Grade 3 Language Arts scores ranged from 38 to 100, with the mean score being 80%. Grade 3 Mathematics scores ranged from 23 to 44, with the mean score being 87%. Grade 6 Language Arts scores ranged from 32 to 98, with the mean score being 76% and scores in Grade 6 Mathematics ranged from 24 to 54, with the mean score being 83%.

An examination of the means and standard deviations of the complete database sample and the analysis sample reveal that the scores are similar across all variables between the two samples. As such, it is reasonable to present only the analysis sample correlations.

Table 4 provides an overview of the percentage of students in the complete database sample achieving at the various achievement standard levels on the PAT. Scores for the analysis sample were similar, and thus will not be presented. In all areas of the PAT, a disproportionate percentage of students scored in either the Standard of Excellence or Acceptable Standard range of achievement. A higher percentage of students scored in the Acceptable Standard range on the Grade 3 and Grade 6 Language Arts PAT, whereas the opposite occurred for the corresponding Mathematics component

of the PAT, with a higher percentage of students scoring in the Standard of Excellence, albeit the percentage difference between the two ranges was minimal.

Table 4

Percent of Students in the Various PAT Range of Achievement Levels

	Standard of Excellence	Acceptable Standard	Unacceptable Standard
PAT			
Grade 3 Language Arts	29.5	69.5	0.9
Grade 3 Mathematics	60.6	37.5	1.9
Grade 6 Language Arts	18.7	80.0	1.3
Grade 6 Mathematics	54.4	43.8	1.9

(n=367)

Pearson Product-Moment Correlation Coefficients

Pearson product-moment correlation coefficients were computed to determine whether the correlations between the CCAT standard scores and the PAT actual achievement scores were positive and significant. This analysis produced a correlation matrix (*n=123*), presented in Table 5. Significant correlations ($p < .01$) observed in the correlation matrix between the CCAT standard scores and the PAT actual achievement scores are denoted by an asterisk.

Table 5

Pearson Product-Moment Correlations Between CCAT Scores and PAT Actual Achievement Scores for the Analysis Sample

	Grade 3 Language Arts	Grade 3 Mathematics	Grade 6 Language Arts	Grade 6 Mathematics
CCAT				
Verbal Battery.	.37*	.25*	.26*	.14
Non-Verbal Battery	.20*	.42*	.16	.43*
Quantitative Battery	.20*	.47*	.19	.46*

(*n=123*)
* Correlation is significant at the .01 level (two-tailed)

In terms of the three CCAT batteries, statistically significant correlations were found between the CCAT Verbal battery and the Grade 3 Language Arts PAT, Grade 3 Mathematics PAT, and the Grade 6 Language Arts PAT, but not with the Grade 6 Mathematics PAT. Statistically significant correlation coefficients ranged from .25 to .37. Statistically significant correlations were observed between the CCAT Non-Verbal battery and the Grade 3 Language Arts PAT, Grade 3 Mathematics PAT, and the Grade 6 Mathematics PAT, but not with the Grade 6 Language Arts PAT. Statistically significant correlation coefficients ranged from .20 to .43. Finally, statistically significant correlations were found between the CCAT Quantitative battery and all forms of the PAT, with statistically significant correlation coefficients ranged from .19 to .47.

Similarly, Pearson product-moment correlation coefficients were computed to determine whether the correlations between the CCAT standard scores and the PAT

range of achievement were positive and significant. A correlation matrix (*n=123*), was also produced for this analysis and is presented in Table 6. Significant correlations ($p<$ *.01*) observed in the correlation matrix between the CCAT standard scores and the PAT range of achievement are denoted by an asterisk.

Table 6

Pearson Product-Moment Correlations Between CCAT Scores and PAT Achievement Range for the Analysis Sample

	Grade 3 Language Arts	Grade 3 Mathematics	Grade 6 Language Arts	Grade 6 Mathematics
CCAT				
Verbal Battery.	.27*	.23*	.16	.16
Non-Verbal Battery	.18*	.44*	.00	.34*
Quantitative Battery	.11	-.06	.40*	.35*

(*n=123*)
* Correlation is significant at the .01 level (two-tailed)

Significant correlations were found between the CCAT Verbal battery and the range of achievement for the Grade 3 Language Arts PAT, and the Grade 3 Mathematics PAT, but not with the Grade 6 Language Arts PAT or the Grade 6 Mathematics PAT. Statistically significant correlations were observed between the CCAT Non-Verbal battery and the range of achievement for the Grade 3 Language Arts PAT, Grade 3 Mathematics PAT, and the Grade 6 Mathematics PAT, but not with the Grade 6

Language Arts PAT. Statistically significant correlations were observed between the CCAT Quantitative battery and the range of achievement for the Grade 6 Language Arts PAT, Grade 6 Mathematics PAT, but not with the Grade 3 Language Arts PAT, or the Grade 3 Mathematics PAT.

Summary

The summary of this chapter will be presented in terms of the research questions posed by this study. Standard scores on the CCAT Verbal battery were significantly correlated with actual achievement scores on both the Grade 3 and Grade 6 Language Arts PAT. The CCAT Verbal battery standard scores were also significantly correlated with the range of achievement on the Grade 3 Language Arts PAT. Standard scores on the CCAT Non-Verbal battery were significantly correlated with actual achievement scores on the Grade 3 and Grade 6 Mathematics PAT, as well as the Grade 3 Language Arts PAT. Standard scores on this battery were also significantly correlated with the range of achievement scores on the same three PAT batteries. Finally, the standard scores on the CCAT Quantitative battery were significantly correlated with the actual achievement scores on the Grade 3 and Grade 6 Mathematics PAT. The CCAT Quantitative battery standard scores were significantly correlated with the range of achievement on the Grade 6 Mathematics PAT.

Chapter 5

Discussion

This research study investigated whether employing the concept of off-level testing is effective in identifying intellectual talent among early elementary aged students. This task was accomplished by correlating standard scores obtained on the three batteries (Verbal, Non-Verbal, and Quantitative) of the Canadian Cognitive Abilities Test (CCAT) when administered off-level with subsequent achievement scores on the Grade 3 and Grade 6 Provincial Achievement Tests (PAT) in Language Arts and Mathematics.

This chapter provides a discussion of the results of the descriptive statistics of the means of the CCAT and PAT scores. This is followed by a discussion of the obtained Pearson product-moment correlations, addressing the research and ancillary questions, between the three CCAT battery standard scores and the Grade 3 and 6 Language Arts and Mathematics PAT scores. The CCAT's utility as an off-level test to predict future academic achievement is addressed. Finally, limitations of the study are presented, directions for future research are suggested, and implications of the study are discussed.

Descriptive Statistics

The mean scores obtained by the complete database sample as well as the analysis sample on all batteries of the CCAT when administered off-level were higher than those reported in the manual for the standardization sample (mean=100). Similarly, mean scores obtained by both sample groups on the four PAT's evaluated, were considered high (ranging from 76%-87%) and thus, did not follow a normal distribution. In fact, in all cases the mean scores on the PAT were in the High Acceptable range of achievement

and one mean score (Grade 3 Mathematics) met the Standard of Excellence for achievement. This may be explained by the present study sample being a convenience sample comprised of students who had been identified, most commonly through informal methods (parent checklist), as being talented. Perleth, et al. (2000) criticized gifted research with young children for the lack of studies with representative samples suggesting that obtained results are not generalizable to the larger population of young gifted children.

Research Questions

The primary research question of this study asked whether the CCAT administered off-level predicts future academic achievement of gifted young students. As each of the ancillary questions adds further clarity to this primary question, they will be addressed in turn.

The first ancillary question asked how well does a young child's standard score on the Verbal battery of the CCAT predict their actual as well as their range of achievement on the Language Arts portion of the PAT. Addressing actual achievement, the relationship between scores on the Verbal battery of the CCAT and the Language Arts PAT was significant and positive, though moderate. The correlation between the CCAT Verbal battery and the Grade 3 Language Arts PAT was .37 and the correlation between the CCAT Verbal battery and the Grade 6 Language Arts PAT was .26. These relationships are expected given that the verbal reasoning abilities tapped by the CCAT play an important role "in developing skills in reading comprehension, critical thinking,

writing and other types of verbal learning tasks" (Thorndike & Hagen, 1998, p. 8) addressed by the two levels of the Language Arts PAT.

When reviewing the range of achievement on the same PAT's as related to the CCAT Verbal battery, there was a positive relationship between the CCAT Verbal battery and the Grade 3 Language Arts PAT range of achievement. The correlation was again moderate at .27. There was no relationship observed between the CCAT Verbal battery and the Grade 6 Language Arts PAT range of achievement. The moderate and non-significant relationship between the CCAT Verbal battery and the Language Arts PAT range of achievement are explained given that the mean score on the Grade 3 Language Arts Pat was 80% and the mean score on the Grade 6 Language Arts PAT was 76%. Due to the restricted range of scores falling in the higher end of a normal distribution correlation potential is limited. Pyryt (1996) pointed out that correlations will diminish when range is restricted on either the predictor or the criterion.

The second ancillary question asked how well does a young child's standard scores on the Non-Verbal battery of the CCAT administered off-level predict their actual achievement as well as their range of achievement on the Language Arts portion of the PAT. In response to this question, the only significant and positive, although weak, relationships were observed between the CCAT Non-Verbal battery and the Grade 3 Language Arts actual achievement as well as between the CCAT Non-Verbal battery and the Grade 3 Language Arts PAT range of achievement. The correlation between scores on the CCAT Non-Verbal battery and the Grade 3 Language Arts PAT actual achievement scores was .200. The correlation between the CCAT Non-Verbal battery and the Grade 3 Language Arts PAT range of achievement was .18. There was not a

significant relationship between the CCAT Non-Verbal battery and either the Grade 6 Language Arts PAT actual achievement scores or range of achievement. As the correlations related to the second ancillary question were low, it can be concluded that the Non-Verbal battery of the CCAT when administered off-level is a weak predictor of both actual achievement and range of achievement on the Language Arts portion of the PAT.

The third ancillary question of this study asked how well does a child's standard score on the Quantitative battery of the CCAT administered off-level predict their actual achievement as well as their range of achievement on the Mathematics portion of the PAT. Correlations between the CCAT Quantitative battery and actual achievement scores for both levels of the Mathematics PAT were significant and positive and indicated a strong relationship. The correlation between the CCAT Quantitative battery and the Grade 3 Mathematics PAT was .47. In a similar fashion, the correlation between the CCAT Quantitative battery and the Grade 6 Mathematics PAT was .46. The strong relationship between these scores is expected given that the three subtests of the CCAT Quantitative battery (Quantitative Relations, Number Series, and Equation Building) require that a student have both deductive and inductive reasoning skills in working with quantitative symbols and concepts (Thorndike & Hagen, 1998). Likewise, the Mathematics PAT requires the formal development through appropriate schooling of these same skills to achieve a high score on the basic facts and problem solving components of this test (Alberta Learning, 2002).

There was no significant relationship found between the CCAT Quantitative battery and the Grade 3 Mathematics PAT range of achievement. Like the Language Arts

PAT, this may be attributed to the negatively skewed distribution, as the mean score for this test was 87%. There was however, a significant and positive relationship between the CCAT Quantitative battery and the range of achievement on the Grade 6 Mathematics PAT, with the correlation being .35.

The fourth ancillary question asked how well does a young child's standard score on the non-verbal battery of the CCAT administered off-level predict their actual achievement as well as their range of achievement on the Mathematics portion of the PAT. Although not as strong as the Quantitative battery, the scores on the Non-Verbal battery were significantly and positively related to both grade levels of the Mathematics PAT. The correlation between the Non-Verbal battery and actual achievement scores on the Grade 3 Mathematics PAT was .42 and the correlation between the Non-Verbal battery scores and the Grade 6 Mathematics PAT actual achievement was .43. It makes sense that the correlations were stronger between the Non-Verbal battery and the Mathematics PAT as compared to the Language Arts PAT, as this test uses geometric shapes and figures to measure general inductive reasoning skills. The CCAT test authors state that no reading or outside knowledge is required for the Non-Verbal battery and that the geometric shapes and figures have little direct relationship to formal school instruction (Thorndike & Hagen, 1998), however the CCAT Non-Verbal test questions are more similar in nature to the Mathematics PAT as compared to the Language Arts PAT.

Once again, scores on the CCAT Non-Verbal battery were significantly and positively related to both levels of the Mathematics PAT range of achievement. The correlation between the CCAT Non-Verbal battery and the Grade 3 Mathematics PAT

range of achievement was .44 and the correlation between the CCAT Non-Verbal battery and the Grade 6 Mathematics PAT range of achievement was .34. As these correlations are moderate in strength, it may be suggested that the CCAT Non-Verbal battery is a reasonable predictor of the range of achievement in both levels of the Mathematics PAT.

Before addressing the final ancillary question of this study, ancillary questions one through four will be summarized. Each of the first four ancillary questions asked how well a specific domain or battery of the CCAT administered off-level could predict later actual and range of achievement as measured by the PAT. In terms of actual achievement, the Verbal battery of the CCAT was the best predictor of Grade 3 Language Arts, although it did also predict Grade 6 Language Arts. The CCAT Non-Verbal battery was the best predictor of Grade 6 Mathematics, although it also predicted Grade 3 Mathematics and to a lesser extent Grade 3 Language Arts. Finally, the Quantitative battery was the best predictor of Grade 3 Mathematics, although it also did an acceptable job of predicting Grade 6 Mathematics. In respect to the range of achievement observed in the analysis, the Verbal battery predicted Grade 3 Language Arts performance. The Non-Verbal battery predicted Grade 3 Language Arts, Grade 3 Mathematics and Grade 6 Mathematics performance and the Quantitative battery predicted Grade 6 Mathematics performance ranges.

Despite the significant and positive correlations found in this study, it should be understood that the strength of the correlations varied with none being considered strong by Sattler's (2001) requirement of a correlation coefficient of .50, and some being considered moderate in strength by Cohen's (1983) recommended requirement of .30.

Although the degree of the relationship would be lessened in terms of the numerical value of the correlation coefficient, a further degree of clarity would be obtained by reporting the coefficient of determination (r^2) for each relationship examined (Gravetter & Wallnau, 1999). For each variable pair this would indicate the variance in one variable accounted for by the other variable.

Given that significant and positive correlations were found between the CCAT administered off-level and later academic achievement, the final ancillary question asked what the longitudinal validity of the predictions were. With the exception of two cases, CCAT results were more predictive of earlier (Grade 3 PAT's) actual achievement as well as range of achievement results. These results are expected given that short-term predictions usually yield higher correlations than long term predictions (Pyryt, 1996). The one exception was the CCAT Non-Verbal battery was a better predictor of Grade 6 Mathematics PAT actual achievement. The second exception was that the CCAT Quantitative battery was a better predictor of Grade 6 Mathematics PAT range of achievement.

Recognizing that a criterion problem, specifically "What is success?" exists in gifted education (Callahan, et al., 1995; Frasier, 1997), the relevance of the low correlations must be questioned. As students are already receiving acceptable scores on the PAT's, one must ask if it really matters if the correlations are weak. On the other hand, if the school's objective is to have more students scoring in the Excellence range, then higher correlations may be of more importance. As such, not just in the case of off-level testing, but with any testing, a clearly defined objective should be established.

Additionally, low correlations observed in this study may be explained by the context of the school. The school used for this study, addresses education from a Multiple Intelligences perspective, attempting to provide rich learning experiences in many different areas. If a specific talent search approach were followed where core academic areas, such as Mathematics and Language Arts, received more focus, the results may be different. Moreover, it is clear from research that more than just school experiences impact achievement (Lytton & Pyryt, 1998).

Using the CCAT Off-Level to Predict Later Achievement

Although rarely the focus of empirical research, the CCAT is used as a screening instrument for giftedness in the school utilized for this study, as well as other local school districts (Mendaglio, 2003). The CCAT is considered an acceptable off-level testing instrument given that it can be administered at least two years above a student's chronological age as recommended by Tsai (2001) and a full range of scores can be represented because the test is neither too hard nor too easy, an important off-level test quality according to Benbow (1991).

Of the tests reviewed in Chapter Two, the CCAT most closely resembles the School College Abilities Test (SCAT) in that it can be used with a variety of age levels (elementary to high school) and measures quantitative and verbal reasoning ability. The SCAT administered off-level was found to be a good predictor of students who will perform well in advanced-level, fast-paced courses of instruction (Cohn, 1991). By comparison, this study found that the CCAT varied in its ability to predict later academic achievement when used with young children.

As such, despite being an appropriate off-level testing instrument, the CCAT administered off-level may not be the most effective method for predicting later academic achievement of young students. This study would add further credibility to Perleth, et al. (2000) claim that the validity of using standardized tests with preschool and elementary aged students to predict high ability and high achievement from an early age is sceptical. All three batteries of the CCAT did however, serve the function of differentiating highly able students who would likely score at the upper percentiles of a normal curve on an in-grade test. The new distribution of off-level scores provides more information about students at the extremes of a normal distribution (Assouline & Lupkowski-Shoplik, 1997).

Limitations of the Study

There are a number of limitations that must be acknowledged in this exploratory study. First, the data gathered was archival in nature which leads to a few inherent risks. Especially in the case of the CCAT, the researcher had to assume that standardized administration procedures were followed during the assessments. This may not have always been the case, since classroom teachers not specifically trained in the administration and scoring of this test were responsible for the assessments.

Second, although still related to scoring, is that all CCAT scores obtained from student's cumulative school files were reported in percentiles. As such, there is a degree of specificity lost when standard scores are converted to percentile ranks and can be misleading (Assouline, 1997). In order for the student's CCAT percentile ranks to be compared with their PAT results, the percentile ranks had to be converted back to

standard scores, thus increasing the likelihood that errors may have occurred when manipulating or transferring the data.

Third, participants in this study formed a sample of convenience. As such, these students were not representative of a normal population. The students in the study were achieving at a higher level compared to those of a normal population, leading to a negatively skewed sample group.

Recommendations for Future Research

This was an exploratory study investigating whether the practice of off-level testing of young students to predict future academic achievement is effective. As such, similar studies with this age group are required, as additional research in this area would either support or refute this practice for use with young children. Similarly, standardized tests beyond the CCAT could be explored as possible instruments for off-level use with young children. Finally, more longitudinal research is required to determine whether the identification of young gifted students is a reasonable endeavour and the methods used are reliable and valid. Such studies could focus on possible indicators and determinants of giftedness and talent.

Implications of the Study

Results from this study indicated that the practice of administering an off-level test to young students to predict their future academic achievements may be limited. This practice should not be discounted altogether though. As in the case of the participating school in this study, off-level testing is used a component of the admission requirements. Lacking formal research, but evidenced through experience, school staff have found that

students who score at or above the mean on any one of the CCAT batteries administered off-level are usually successful candidates in the program (Joanne Koch, personal communication, 2003). Furthermore, by acknowledging talent in any one of the three areas assessed a larger percentage of students may be fostered in their development, an original goal of Stanley's Talent Search Model. This method of identification is also aligned with Feldhusen's Talent Identification and Development in Education (TIDE) model in that talent can be recognized in specific areas and verified through standardized testing or achievement based measures (Feldhusen, 1998). The aforementioned suggests that off-level testing with young students may be better used as an identification and selection device.

Although results from this study suggest that off-level testing may have more initial impact for admission purposes, perhaps more emphasis could be placed on the off-level results for alternative uses. For example, the results could be used to determine ability groupings in specified areas. As Stanley (1984) suggested groupings should be based on special ability scores relevant to the subject being studied. Additionally, as suggested by gifted education models (Pyryt, 1999; Stanley, 1984; 1991), results could be incorporated into a prescriptive teaching methodology whereby students are taught what they don't already know. According to Pyryt (1999), this form of teaching involves "pretesting to determine a student's current knowledge level, analyzing errors to determine instructional needs, [and] implementing instruction based on these needs" (p. 18).

If children are thought by other means to be talented (i.e. parent referrals, checklists, etc.), the process of administering an off-level test such as the CCAT is a

time-efficient, cost-effective method for confirming hypotheses and possibly guiding educational decisions, as is the case at Westmount Charter School. By discerning areas and levels of abilities, educators are given a rough estimate of whether the compacted, differentiated program offered by this school for elementary aged students is suitable.

Finally, although the correlations found in this study between the CCAT and PAT achievement were not considered strong enough to suggest that off-level testing provided by the CCAT can predict future academic achievement on the PAT's, the researcher feels confident to suggest that administering a test designed for older students to younger students has other benefits and thus should continue to receive further support both in research and in practice.

References

Achter, J. A., Benbow, C.P., & Lubinski, D. (1997). Rethinking multipotentiality among the intellectually gifted: A critical review and recommendations. *Gifted Child Quarterly, 41*, 5-15.

Achter, J. A., & Lubinski, D. (2003). Fostering exceptional development in intellectually talented populations. In W. B. Walsh (Ed.), *Counselling psychology and optimal human functioning* (pp. 25-54). Hillsdale, NJ: Lawrence Erlbaum Associates.

Alberta Learning. (2000). *Teaching students who are gifted and talented.* Edmonton: Alberta Learning Cataloguing in Publication.

Alberta Learning. (2002). Achievement testing program grade 3, 6, and 9 – General information bulletin. Edmonton, Alberta: Learner Assessment Branch.

Anderson, J. O. (1995). Review of the Canadian Cognitive Abilities Test, Form 7. In J. C. Colony & J.C. Imprava (Eds.), *The twelfth mental measurements yearbook* (pp. 153-154). Lincoln, NE: Buros Institute.

Assouline, S. G. (1997). Assessment of Gifted Children. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 89-108). Boston: Allyn & Bacon.

Assouline, S. G. (2003). Psychological and educational assessment of gifted children. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 124-145). Needham Heights, MA: Allyn & Bacon.

Assouline, S. G., & Lupkowski, A. E. (1992). Extending the talent search model: The potential of the SSAT-Q for identifying mathematically talented elementary students. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent Development: Proceedings from the 1991 Henry B. and Jocelyn Wallace National*

Research Symposium on Talent Development (pp. 223-232). Scottsdale, AZ: Gifted Psychology Press.

Assouline, S. G., & Lupkowski-Shoplik, A. (1997). Talent searches: A model for the discovery and development of academic talent. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 170-179). Boston: Allyn & Bacon.

Benbow, C. P. (1991). Mathematically talented children: Can acceleration meet their educational needs? In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 154-165). Boston: Allyn & Bacon.

Benbow, C. P., & Lubinski, D. (1993). Psychological profiles of the mathematically talented: Some sex differences and evidence supporting the biological basis. In G.R. Bock & K. Ackrill (Eds.). *The origins and development of high ability* (pp. 44-66). Chicester, UK: Wiley.

Benbow, C. P., Lubinski, D., & Sanjani, H. E. (1999). Our future leaders in Science: Who are they? Can we identify them early? In N. Colangelo & S. G. Assouline (Eds.), *Talent Development III: Proceedings from the 1995 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 59-70). Scottsdale, AZ: Gifted Psychology Press.

Benbow, C. P., Lubinski, D., Shea, D. L., & Eftekhari-Sanjani, H. (2000). Sex differences in mathematical reasoning: Their status 20 years later. *Psychological Science, 11*, 474-480.

Benbow, C. P., & Stanley, J. (1996). Inequity in equity: How "equity" can lead to inequity for high-potential students. *Psychology, Public Policy, and Law, 2*, 249-292.

Blythe, T. & Gardner, H. (1990). A school for all intelligences. *Educational Leadership, 47* (7), 33-38.

Callahan, C. M., Hunsaker, S. L., Adams, C. M., Moore, S. D., & Bland, L. C. (1995). *Instruments used in the identification of gifted and talented students*. Storrs, CT: The National Research Center on the Gifted and Talented.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Clarke, A. (2001). Early identification of gifted students and catering for them in the early stages of kindergarten. *Primary Educator, 7* (1), 21-29.

Cohen, J. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohn, S. (1991). Talent searches. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 166-177). Boston: Allyn & Bacon.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes*. Needham Heights, MA: Allyn & Bacon.

Deary, I. J., Austin, E. J., & Caryl, P. G. (2000). Testing versus understanding human intelligence. *Psychology, Public Policy, and Law, 6*, 180-190.

Feldhusen, J. F. (1998). Programs for the gifted few or talent development for the many? *Phi Delta Kappan, 79*, 735-738.

Feldhusen, J. F., & Jarwan, F. A. (2000). Identification of gifted and talented youth for educational programs. In K. A. Heller, F. J. Mönks, R. J. Sternberg, & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2nd ed., pp. 271-282). UK: Elsevier.

Frasier, M. M. (1997). Multiple criteria: The mandate and the challenge. *Roeper Review, 20*, Gifted Education Supplement (pullout in centre of magazine).

Gagné, F. (1985). Giftedness and talent: Reexamining a reexamination of the definitions. *Gifted Child Quarterly, 29*, 103-112.

Gagné, F. (2000). Understanding the complex choreography of talent development through DMGT-based analysis. In K. A. Heller, F. J. Mönks, R. J. Sternberg, & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2nd ed., pp. 67-79). UK: Elsevier.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gardner, H. (1991). *The unschooled mind*. New York: Basic Books.

Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.

George, W. C. (1979). The talent search concept: An identification strategy for the intellectually gifted. *The Journal of Special Education, 13*, 221-237.

Gottfredson, L. S. (2003). The science and politics of intelligence in gifted education. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 24-40). Boston: Allyn and Bacon.

Goldstein, D., Stocking, V., & Godfrey, J. (1999). What we've learned from talent search research. In N. Colangelo & S. G. Assouline (Eds.), *Talent Development: Proceedings from the 1995 Henry B. and Jocelyn Wallace National Research*

Symposium on Ttalent Development (3rd ed., pp.143-152). Scottsdale, AZ: Gifted Psychology Press.

Gravetter, F. J. & Walleau, L. B. (1999). *Essentials of statistics for the behavioural sciences* (3rd ed.). New York: Brooks/Cole Publishing Company.

Guilford, J. P. (1985). The structure of intellect model. In B. B. Wolman (Ed.) *Handbook of intelligence: Theories, measurements, and applications* (pp. 225-266). New York: Wiley.

Hadaway, N., & Marek-Schroer, M. F. (1992). Multidimensional assessment of the gifted minority student. *Roeper Review, 15*, 73-78.

Hattie, J. (1995). Review of the Canadian Cognitive Abilities Test, Form 7. In J.C. Colony & J. C. Imprava (Eds.), *The twelfth mental measurements yearbook* (pp. 154-155). Lincoln, NE: Buros Institute.

Hollingsworth, L. S. (1926). *Gifted children: Their nature and their nurture.* New York: MacMillan Company.

Jackson, E. N. (2003). Young gifted children. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp. 470-482). Boston: Allyn & Bacon.

Jenkins-Friedman, R. (1982). Myth: Cosmetic use of multiple selection criteria! *Gifted Child Quarterly, 26*, 24-26.

Jensen, A. R. (1998). *The "g" factor – The science of mental ability.* Westport, CT: Praeger.

Kaufman, A. S., Lichtenberger, E. O., & Naglieri, J. A. (1999). Intelligence testing in the schools. In C. R. Reynolds & T. B Gutkin (Eds.), *The handbook of school psychology* (3rd ed., pp. 307-349). New York: Wiley.

Lubinski, D., Webb, R. M., Morelock, M. J., & Benbow, C. P. (2001). Top 1 in 10,000: A 10-year follow-up of the profoundly gifted. *Journal of Applied Psychology, 86*, 718-729.

Lupkowski-Shoplik, A. (1999). A closer look at EXPLORE-Mathematics. In N. Colangelo & S. G. Assouline (Eds.), *Talent Development III: Proceedings for the 1995 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 461-464). Scottsdale, AZ: Gifted Psychology Press.

Lupkowski-Shoplik, A. (2001). Carnegie Mellon student talent search: Establishing appropriate guidelines for qualifying test scores. In N. Colangelo & S. G Assouline (Eds.), *Talent Development: Proceedings from the 1998 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (4th ed., pp. 379-385). Scottsdale, AZ: Gifted Psychology Press.

Lupkowski-Shoplik, A. & Assouline, S. G. (1993). Identifying mathematically talented elementary students: Using the lower level of the SSAT. *Gifted Child Quarterly, 37*, 118-123.

Lupkowski-Shoplik, A., Benbow, C. P., Assouline, S. G., & Brody, L. (2003). Talent searches: Meeting the needs of academically talented youth. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp. 204-218). Boston: Allyn & Bacon.

Lupkowski-Shoplik, A., & Swiatek, M. A. (1999). Elementary student talent searches: Establishing appropriate guidelines for qualifying test scores. *Gifted Child Quarterly, 43*, 265-272.

Lytton, H. & Pyryt, M. (1998). Predictors of achievement of basic skills: A Canadian effective schools study. *Canadian Journal of Education, 23*, 281-301.

Mendaglio, B. A. (2003). *Concurrent validity of the Canadian Cognitive Abilities Test (CCAT) with the Weschler Intelligence Scale for Children-Third Edition (WISC-III)*. Unpublished mater's thesis, University of Calgary, Calgary, Alberta, Canada.

Mills, C. J., & Barnett, L. B. (1992). The use of the secondary school admission test (SSAT) to identify academically talented elementary students. *Gifted Child Quarterly, 36*, 155-159.

Mönks, F. J. & Mason, E. J. (1993). Developmental theories and giftedness. In K. A. Heller, F. J. Mönks, & A. H. Passow (Eds.), *International handbook of giftedness and talent* (pp. 89-101). UK: Elsevier.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive assessment system interpretive handbook.* Chicago: Riverside.

Olszewski-Kubilius, P. (1998a). Research evidence regarding the validity and effects of talent search educational programs. *The Journal of Secondary Gifted Education, 9*, 134-138.

Olszewski-Kubilius, P. (1998b). Talent search. *The Journal of Secondary Gifted Education, 9*, 106-113.

Pellegrino, J. W. & Glasser, R. (1979). Cognitive correlates and components in the analysis of individual differences. *Intelligence, 3*, 187-214.

Perleth, C., Lehwald, G., & Browder, C. S. (1993). Indicators of high ability in young children. In K. A. Heller, F. J. Mönks, & A. H. Passow (Eds.), *International handbook of giftedness and talent* (pp. 283-310). UK: Elsevier.

Perleth, C., Schatz, T, & Mönks, F. J. (2000). Early identification of high ability. In K. A. Heller, F. J. Mönks, R. S. Sternberg, & R. F. Subotnik (Eds), *International handbook of giftedness and talent* (2nd ed., pp. 297-316). UK: Elsevier.

Pritchard, M. C. (1951). The contributions of Leta S. Hollingworth to the study of gifted children. In P. Witty (Ed.) *The gifted child* (pp. 47-85). Boston: D.C. Health.

Pyryt, M. C. (1996). IQ: Easy to bash, hard to replace. *Roeper Review, 18*, 255-259.

Pyryt, M. C. (1999). Putting the "g" back in gifted education. *Understanding our Gifted, 12*, 16-18.

Pyryt, M. C., Colangelo, N., Assouline, S. G., & Ihrig, D. (2003, May). *Importing the talent search model into Canada: The Alberta experience*. Paper presented at the meeting of the Canadian Society for Studies in Education.

Ramos-Ford, V., & Gardner, H. (1997). Giftedness from a multiple intelligence perspective. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education*. (pp. 54-66). Needham Heights, MA: Allyn and Bacon.

Reis, S. M., & Renzulli, J. S. (1992). Using curriculum compacting to challenge the above-average. *Educational Leadership, 50*, 51-57.

Renzulli, J. S. (1978). What makes giftedness: Reexaming a definition. *Phi Delta Kappan, 60*, 180-184.

Renzulli, J. S., & Reis, S. M. (2003). The schoolwide enrichment model: Developing creative and productive giftedness. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp. 184-203). Needham Heights, MA: Allyn and Bacon.

Richert, E. S. (2003). Excellence with justice in identification and programming. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 146-158). Needham Heights, MA: Allyn and Bacon.

Sarouphim, K. M. (1999). DISCOVER: A promising alternative assessment for the identification of gifted minorities. *Gifted Child Quarterly, 43*, 244-251.

Sarouphim, K. M. (2000). Internal structure of DISCOVER: A performance-based assessment. *Journal for the Education of the Gifted, 23*, 314-327.

Sattler, J. M. (1988). *Assessment of children* (3rd ed.). San Diego: Author.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego: Author.

Smutney, J. F. (1999). A special focus on young gifted children. *Roeper Review, 21*, 172-174.

Spearman, C. (1927). *Abilities of man*. New York: Macmillan.

Stanley, J. C. (1984). Use of general and specific aptitudes in identification: Some principles and certain cautions. *Gifted Child Quarterly, 28*, 177-180.

Stanley, J. C. (1990). Leta Hollingworth's contributions to above-level testing of the gifted. *Roeper Review, 12*, 166-171.

Stanley, J. C. (1991). An academic model for educating the mathematically talented. *Gifted Child Quarterly, 35*, 36-42.

Stanley, J. C. (1996). In the beginning - The study of mathematically precocious youth. In. C.P. Benbow & D. Lubinski (Eds.), *Intellectual Talent – Psychometric and Social Issue.* (pp. 225-235). Baltimore: The Johns Hopkins University Press.

Stanley, J. C. (1999). SMPY's first decade. In N. Colangelo & S. G. Assouline (Eds.), *Talent development: Proceedings from the 1995 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (3rd ed., pp. 139-141). Scottsdale, AZ: Gifted Psychology Press.

Stanley, J. C. & Brody, L. (2001). History and philosophy of the talent search model. *Gifted and Talented International, 16*, 94-96.

Sternberg, R. J. (1986). A triarchic theory of intellectual giftedness. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 223-243). New York: Cambridge University Press.

Sternberg, R. J. (1997). A triarchic view of giftedness. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (2nd ed., pp. 88-99). Boston: Allyn and Bacon.

Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology, 3*, 1-25.

Sternberg, R. J. (2003) Giftedness according to the theory of successful intelligence. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 88-99). Boston: Allyn and Bacon.

Sternberg, R. J. & Grigorenko, E. L. (2001). Guilford's structure of intellect model of creativity: Contributions and limitations. *Creativity Research Journal, 13*, 309-316.

Sternberg, R. J., Grigorenko,E. L., Jarvin, L., Clinkenbeard, P., Ferrari, M., & Torff ,B. (2000). The effectiveness of triarchic teaching and assessment. *National Research Center on the Gifted and Talented Newsletter*, (pp. 3-9).

Tannenbaum, A. J. (2003). Nature and nurture of giftedness. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd ed., pp. 45-59). Boston: Allyn and Bacon.

Terman, L., & Oden, M. H. (1947). *Genetic studies of genius: The gifted child grows up.* Stanford, CA: Stanford University Press.

Thorndike, R. L. & Hagen, E. P. (1998). *Canadian Cognitive Abilities Test, Form K.* Scarborough, Ontario: Nelson.

Treffinger, D. J. (1986). Fostering effective, independent learning through individualized programming. In J. S. Renzulli (Ed.), *Systems and models for developing programs for the gifted and talented* (pp. 429-460). Mansfield Center, CT: Creative Learning Press.

Tsai, T. (2001). Investigating grade-by-gender interaction effect of academically talented elementary students on the EXPLORE English test. In N. Colangelo & S. G. Assouline (Eds.), *Talent Development: Proceedings from the 1995 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development.* (4th ed., pp. 449-451). Scottsdale, AZ: Gifted Psychology Press.

Tyler-Wood, T., & Carri, L. (1991). Identification of gifted children: The effectiveness of various measures of cognitive ability. *Roeper Review, 14,* 63-64.

VanTassel-Baska, J. (1996). Contributions of the talent-search concept to gifted education. In C. P. Benbow & D. Lubinski (Eds.), *Intellectual talent: Psychometric and social issues* (pp. 236-245). Baltimore: The Johns Hopkins University Press.

Van Tassel-Baska, J. (1999). Introduction to talent search panel. In N. Colangelo & S. G. Assouline (Eds.), *Talent Development III: Proceedings from the 1995 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 135-138). Scottsdale, AZ: Gifted Psychology Press.

Vernon, P. E. (1950). *The structure of human abilities*. New York: Wiley.

Von Karolyi, C., Romos-Ford, V., & Gardner, H. (2003). International perspective on gifted education. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 100-112). Needham Heights, MA: Allyn & Bacon.

Winner, E. (1996). *Gifted children: Myths and realities*. New York: Basic Books.

Appendix A

UNIVERSITY OF CALGARY

CERTIFICATION OF INSTITUTIONAL ETHICS REVIEW

This is to certify that the Conjoint Faculties Research Ethics Board at the University of Calgary has examined the following research proposal and found the proposed research involving human subjects to be in accordance with University of Calgary Guidelines and the Tri-Council Policy Statement on *"Ethical Conduct in Research Using Human Subjects"*. This form and accompanying letter constitute the Certification of Institutional Ethics Review.

File no:	**CE101-3779**
Applicant(s):	**Victoria Plouffe**
Department:	**Applied Psychology, Division of**
Project Title:	**The Use of Standardized Tests Administered Off-Level to Predict Future Academic Achievement of Gifted Students**
Sponsor (if applicable):	

Restrictions:

This Certification is subject to the following conditions:

1. Approval is granted only for the project and purposes described in the application.
2. Any modifications to the authorized protocol must be submitted to the Chair, Conjoint Faculties Research Ethics Board for approval.
3. A progress report must be submitted 12 months from the date of this Certification, and should provide the expected completion date for the project.
4. Written notification must be sent to the Board when the project is complete or terminated.

Janice Dickin

Janice Dickin, Ph.D, LLB,
Chair,
Conjoint Faculties Research Ethics Board

29 November 2003

Date:

Distribution: (1) Applicant, (2) Supervisor (if applicable), (3) Chair, Department/Faculty Research Ethics Committee, (4) Sponsor, (5) Conjoint Faculties Research Ethics Board (6) Research Services.

Appendix B

WESTMOUNT CHARTER SCHOOL
The Wind Beneath Our Wings

WESTMOUNT CHARTER SCHOOL
Calgary's Public School for Gifted Students Since 1958
2519 Richmond Road SW
Calgary, AB T3E 4M2
(403) 217-0428 Fax (403) 217-0252
www.abccharter.com
admin@abccharter.com

Elaine McDougall, Ph.D.
Principal

Marylyn Waters, M.Ed.
Rhonda Watrin, Ph.D.
Vice-Principals

September 10, 2003

To Whom It May Concern:

Please be advised that as part of Victoria Plouffe's responsibility as a teacher at Westmount Charter School she has been authorized to collect and prepare a database of student information. Included in this information will be students' scores on their entrance examination to the school (the Canadian Cognitive Abilities Test) as well as students' scores on any Provincial Achievement Test they have written. She has also been granted permission to access any other data of pertinence to the database.

I understand that this database will become part of Victoria Plouffe's thesis research through the University of Calgary. I am aware that all information for the purpose of her research will be presented in a confidential manner. I am also aware that I may contact Victoria Plouffe directly or her supervisor, Dr. Michael Pyryt, should I have any questions or concerns.

I wish Victoria Plouffe the best of luck in her academic studies and am looking forward to the results of her research. Please feel free to contact me should you require any further clarification.

Sincerely,

Marylyn Waters

Marylyn Waters, B.Ed., M.Ed.