# THE RELATIONSHIP BETWEEN KUMON AND ACHIEVEMENT IN MATHEMATICS

**JENNA WOULD**

B.Sc., McMaster University, 2007

A Project
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

**MASTER OF EDUCATION**
(COUNSELLING PSYCHOLOGY)

**FACULTY OF EDUCATION**
LETHBRIDGE, ALBERTA

March 2010

# DEDICATION

I would like to dedicate this project to my Grandfather, Reginald Would, for all the encouragement he has given me to pursue higher education. He won’t be able to see me graduate, but the education that he has helped me obtain and his memory will be with me on graduation day and on all the other days of my life.

ABSTRACT

The purpose of the current study was to investigate the relationship between participation in the Kumon Mathematics programme and children's achievement in Mathematics and whether or not this relationship is different for children of varying academic abilities. Twenty-two students in grades four through six participated in the study. A nonexperimental, causal comparative research design was employed to answer the research question. Children were administered the Canadian Achievement Test, Third Edition (CAT-3) and the Canadian Cognitive Abilities Test (CCAT) shortly after beginning the Kumon programme. Six months later, the CAT-3 was re-administered. Results suggest that there may be a significant relationship between participation in the Kumon programme and development in computation skills ($p = 0.053$), but not with development in mathematical reasoning skills ($p = 0.867$). Results also suggest that there is a significant, negative relationship between pretest computation scores and gains made in computation skills ($p = 0.005$). The conclusions drawn from the results of this study are that Kumon may be more effective as a remediation programme than it is as an enrichment programme. More specifically, Kumon may be a more effective remediation programme for computation skills versus mathematical reasoning skills. However, there are major limitations to the current study, namely a small sample size. As such, statistical analyses were made for exploratory purposes only and no statements can be made regarding the effectiveness of Kumon for children.

ACKNOWLEDGEMENT

I would like to take this opportunity to thank the people that enabled me to complete this project. To my supervisor, Dr. Kerry Bernes: thank you for making yourself available for questions and discussion despite your busy schedule. Your suggestions and guidance were essential to the completion of this project. Dr. Thelma Gunn: thank you for your time and expertise. Dr. Nola Aitken: thank you for your valued suggestions and for going through the document with a fine-tooth comb, making it a much more cogent piece.

I must thank Lisa Kaul, a general manager at Kumon Canada. Without her belief in the value of investigating Kumon's effectiveness in a scientific and independent manor, this project would not have been possible. Lisa: thank you for your time and support. Thank you to the Kumon field staff in Toronto and Vancouver; my sample size would have been considerably smaller if they hadn't helped collect data. I would also like to thank all the Kumon Instructors who went beyond what is required of them to help orchestrate testing in their Centres. Last, I would like to thank all the children who participated in this study.

On a personal level I would like to thank my Mother. Thank you for always being the rock in my life. Without that solid foundation to stand on I could never have reached so high.

TABLE OF CONTENTS

LIST OF TABLES

# CHAPTER I: INTRODUCTION

## Summary of Literature Review

### *Typical Development in Mathematics*

Research has shown that some very basic Mathematics skills are innate (Butterworth, 2005; Shalev, 2004). The more advanced skills, such as enumeration, correspondence construction, counting, Arithmetic, and understanding of arithmetical concepts, develop sequentially through instruction (Ardila & Rosselli, 2002; Butterworth, 2005; Klein & Starkey, 1987). These skills, and the sub-skills that comprise them, progress from acquisition to mastery. Once the skills are mastered, they can be used as a foundation for the acquisition of the next, more advanced skill in the hierarchy. Most of the skills in the elementary school Mathematics curriculum are taught this way. For example, multiplication is often taught as a series of addition problems (Butterworth, 2005). The vast majority of children acquire mathematical abilities in this manner and their achievement matches the expectations outlined in their school board's curriculum. However, there are also many children who demonstrate either above or below average achievement.

### *Giftedness in Mathematics*

Many different criteria are used for classifying children as gifted in the literature. To formally identify an individual as mathematically gifted, the child must be assessed using a measure of cognitive ability [specifically a measure of nonverbal reasoning ability (Preckel, Goetz, Pekrun, & Kleine, 2008)]. However, several researchers have used achievement tests to group children who may be gifted (Ma, 2005; Mills, Ablard, & Gustin, 1994; Niederer, Irwin, Irwin, & Reilly, 2003; Swanson, 2006; Threlfall & Hargreaves, 2008). These researchers typically refer to these children as "talented" or "precocious", in recognition that achievement test scores are not adequate for classifying children as gifted. Similarly, the current study defines a child as "gifted" in Mathematics if he or she scores at or above the $97^{th}$ percentile on a measure of mathematical achievement.

Giftedness occurs in individuals of all ages and in both genders. However, of the individuals who meet criteria for giftedness in Mathematics, there are more males than females (Preckel et al., 2008). There are observable differences between gifted students and non-gifted students. There has been debate as to whether these differences are qualitative (life-long differences in thinking processes), quantitative (precocity), or both (Shore & Kanevsky, 1993; Winner, 2000a). There is also a nature/nurture debate as to the origins of giftedness. Again, it may be that both nature and nurture contribute to giftedness.

The need for Mathematics interventions has largely been ignored to the detriment of the mathematically gifted. Without enrichment, gifted students are at risk for loss of motivation, under-achievement, unhappiness, poor mental health, and social isolation (Ma, 2005; Winner, 2000b). Strategies that have proven to be effective in evading these risks and enhancing achievement in Mathematics for mathematically gifted students include acceleration, flexible pacing, ability grouping, and school-based enrichment classes. Further research into effective interventions that address the special needs of mathematically gifted children is warranted.

*Mathematics Disorder*

The expression of Mathematics Disorder (MD) is heterogeneous (Geary, 1993; Kronenberger & Dunn, 2003; Mazzocco & Myers, 2003). In 1983, Badian proposed five

MD subtypes based on the different ways MD is expressed: Alexia and Agraphia for numbers, Spatial Dyscalculia, Anarithmetria, Attentional-Sequential Dyscalculia, and Mixed Types of Dyscalculia. Later, Geary (1993) suggested that MD should be divided into three subtypes: Semantic Memory, Procedural, and Visuo-spatial. Both models have received empirical support (Mazzocco, 2001; Shalev, Manor, Auerbach, & Gross-Tsur, 1998); however, neither model is universally accepted, nor is it universally accepted that MD subtypes exist. Most researchers do agree, though, that there is heterogeneity in the expression of MD.

MD is fairly prevalent in the school-aged population. Although estimates vary throughout the literature (largely because of differences in researcher’s definitions and diagnostic criteria for the disorder), most estimates are between 5 and 8% (Geary, 2003, Gross-Tsur, Manor, & Shalev, 1996; Lewis, Hitch, & Walker, 1994**;** Shalev, Auerbach, Manor, & Gross-Tsur, 2000). Although comorbidity rates are also variable throughout the literature, children with a MD are often reported to have comorbid ADHD and/or RD (Badian, 1983; Geary, 2003; Gross-Tsur et al., 1996).

Several factors influence the prognosis for a child with a MD, including “the severity of the disorder at the time of initial diagnosis and the presence of Arithmetic problems in the siblings” (Shalev, 2004, p. 766). In approximately 50% of cases, teenagers diagnosed in childhood with a MD continue to experience significant difficulties with Mathematics (Shalev et al., 1998). In the remaining instances, the severe difficulties associated with MD abate, although many still display poor performance in Mathematics.

MD can be attributed to many factors, including genetic, neurological, environmental, and psychological. Twin (Alarcon, Defries, Gillis Light, & Pennington, 1997) and family studies (Shalev et al., 2001) have demonstrated that MD is heritable. Neuro-imaging has revealed differences in brain activation between typically developing individuals and individuals with a MD when solving mathematical problems. More specifically, the majority of studies find differences in the activation of the intraparietal sulcus areas (Butterworth, 2005; Dehaene, Piazza, Pinel, & Cohen, 2003; Issacs, Edmonds, Lucas, & Gadian, 2001; Molko et al., 2003). Environmental factors such as adequacy of instruction (Shalev, 2004), speed of teaching (Cumming & Elkins, 1999), classroom size, student diversity, mathematic curricula, and mainstreaming (Miller & Mercer, 1997) can all influence a child's achievement in Mathematics and compound the effects of a MD. However, they do not cause MD (Learning Disabilities Association of Ontario, [LDAO], 2009). Psychological factors, such as mathematical anxiety, can also contribute to, but do not cause, significant difficulties in Mathematics. In sum, the etiology of MD is multi-factorial and there are many variables that can further exacerbate difficulties in Mathematics.

Typically developing children and children with a MD experience difficulties in Mathematics. However, due to genetic and neurological differences, there are differences in the expression of Mathematics difficulties between the two groups. According to the behaviourist model, all skills are acquired by progressing through the stages of the learning hierarchy (acquisition, fluency, generalization, and adaptation) (Jolivette, Lingo, Houchins, Barton-Arwood, & Shippen, 2006; Scott, 1993; Shalev, 2004). Difficulties experienced by all children in achievement will occur within one of these stages. The difference between the difficulties experienced by children with a MD and typically developing children is quantitative according to the behaviourist model. Children with a

MD experience the same types of difficulties as typically developing children, they simply experience more of these difficulties. Cognitive theorists, on the other hand, believe that mathematical achievement is different for individuals with a MD and typically developing individuals because individuals with a MD have qualitatively different psychological processing skills (visuo-spatial processing, attention, memory, inhibition, and speed of processing).

Scientific knowledge of effective treatments for children with a MD is inadequate. Of the few strategies and programmes that are suggested in the literature, the majority are insufficiently researched and understood. In practice, many school aged children with significant difficulties in Mathematics receive withdrawal support at their school. This means that the student spends the majority of his or her time in a regular class, but that he or she also receives instruction outside the classroom (for less than 50 per cent of the school day) from a qualified special education teacher. There are also specific intervention programmes described in the literature for children with significant difficulties in Mathematics including Great Leaps Math and Cover, Copy, Compare. Unfortunately these programmes are inadequately researched.

The most promising intervention programme, in the current author's opinion, is Kumon. Millions of children are enrolled in Kumon, making it the most widely used after-school programme worldwide (Kumon North America, [KNA], 2008). Although the majority of children enrolled in Kumon are typically developing and gifted children, Kumon also serves thousands of children with disabilities (Kumon Toru Research Institute of Education, [KTRIE], 2002). There is currently very limited scientific evidence of Kumon's effectiveness despite its widespread use.

*Kumon*

The Kumon Method of Learning has seven components: individualized learning, independent learning, comfortable starting point, curriculum, repeated practice, mastery, and advanced level of study (Izumi, 2001). Both Kumon programmes (Reading and Mathematics) were created based on this Method.

The Mathematics programme has twenty-three levels (KNA, 2008). The skills taught in each level correspond with the skills taught from preschool to the university level (KNA, 2008). The difference between Kumon and public education is the method of instruction, or Method of Learning.

The only empirical study regarding Kumon located in a recent PsycInfo search was a study investigating the effect of the Kumon Mathematics programme on the mathematical achievement of economically disadvantaged children (McKenna, Hollingsworth, & Barnes, 2005). Although the study had several limitations, such as including grade two students in their sample [performance in Mathematics in grades one and two is highly variable and not indicative of true ability (American Psychiatric Association, [APA], Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision, [DSM-IV-TR], 2000)], their results suggest that Kumon enhances children's mathematical achievement (McKenna et al., 2005).

Kumon serves children of varying ability and achievement (KTRIE, 2002). There is no systematic research evaluating the effectiveness of Kumon for children outside of the average ability range (for either the below or above average ranges). However, there is one published case report that describes an adolescent with Down Syndrome and his experience with the Kumon Mathematics programme (Haslam, 2007). In the adolescent's mother's opinion, Kumon helped her son improve his Math skills (Haslam, 2007). It is reasonable to predict that the Kumon programme will help children with Learning Disabilities (LD) above average, and gifted children to develop their Mathematics skills because the special needs of these children are so well met by the tenets of the Kumon Method of Learning. Presently, systematic research is needed to determine whether this is true in practice.

## Potential Contribution of Current Study

The current research project is necessary and important for two reasons: 1) people need basic mathematical skills in order to live independently and thus mathematical ability is an important component of having a desirable quality of life, and 2) practitioners, teachers, parents, and children with special needs (remedial and enrichment) require a larger knowledge base of effective Math programmes and interventions.

### *Importance of Mathematical Ability*

Basic mathematical ability is essential to everyday life as it carries implications for practical, civic, recreational, and professional endeavors (Jolivette et al., 2006; Shalev, 2004). Math skills are used in the practical tasks necessary for independent living such as following a recipe (Patton, Cronin, Bassett, & Koppel, 1997; Shalev, 2004). On a civic level, individuals must be able to apply mathematical concepts to interpret information such as filing taxes (Jolivette et al., 2006). Math is also used in certain games and puzzles, affecting the individual on a recreational level (Jolivette et al., 2006). Failure to obtain mastery of basic mathematical skills can also negatively impact the individual's ability to obtain and maintain employment (Jolivette et al., 2006) as well as to achieve employment success (Erford & Klein, 2007).

*Importance of Effective Mathematical Enrichment and Remedial Programmes*

There is a clear, imminent need for evidence of effective Mathematics interventions (Burns, VanDerHeyden, & Jiban, 2006). Gifted children need enrichment programmes because they are often underserved by the general education Mathematics programme (Mills et al., 1994; Silverman, 1989). Unless these children are engaged in a more stimulating Mathematics programme, they are at risk for losing motivation, under-achieving, and being denied appropriate education (Ma, 2005).

Practitioners, teachers, parents, and children with MDs will benefit from a larger database of effective remedial Mathematics programmes. As noted above, mathematical ability is essential to independent living and to becoming a contributing member to society. Effective treatments will therefore directly improve children's quality of life and will therefore have an indirect positive influence on the quality of life of parents, teachers, and practitioners.

For children with a MD, early intervention is an even more critical issue. Achievement in elementary level Mathematics lays the foundation for later, more difficult concepts. Elementary school-aged children with LDs commonly do not catch up to their peers in high school (Cawley, Kahn, & Tedesco, 1989) and often drop out at the high school level (Phelps & Hanley-Maxwell, 1997). Programmes that are open to young children and thus allow for early intervention are especially valuable because they intervene before the child's difficulties are exacerbated.

It is also important to expand the knowledge base of effective Mathematics enrichment and remedial programmes because of the extensive resources invested in these intervention programmes. Children and parents invest time, money, and children's futures in mathematical interventions. Such precious resources are best invested in scientifically proven programmes.

The Kumon Mathematics programme can be used for both remediation and enrichment and children as young as two years old can enroll (personal communication, S. Vishnu, April 9, 2009). By investigating the effectiveness of Kumon, the current study has the potential to benefit children with MDs, average-achieving, above average, and gifted children, as well as parents, teachers, and practitioners.

## Research Problem and Hypotheses

The purpose of the study is to investigate the effectiveness of the Kumon Mathematics programme for children of varying abilities. Specifically, the research questions guiding this study are:

1. Is there a difference between male and female participants in terms of their pretest scores, posttest scores, or the gains they made on the CAT-3?
2. Is there a difference in the magnitude of gains made in Math skills among children of different achievement groups (below average, average, above average, and gifted)?
3. Is there a difference in the magnitude of gains made on a measure of computation skills versus mathematical reasoning skills after six months of participating in the Kumon Math programme?
4. What is the relationship between children's participation in the Kumon Mathematics programme and their achievement in Mathematics?

Based on a literature review, this study has several hypotheses as to how these research problems will be answered:

*Hypothesis One: Gender Differences in Mathematical Ability*

The null hypothesis is that there is no difference between males and females in terms of mathematical ability as measured by the CAT-3. I predict that the null hypothesis will be accepted; there will be no significant difference between genders in pretest scores, posttest scores, or gains made in Mathematics.

*Hypothesis Two: Children of Varying Mathematical Ability*

Participants will be categorized based on their pretest achievement scores into one of four achievement groups: below average, average, above average, or gifted. The null hypothesis is that there will be no difference among achievement groups in the gains made on the measure of mathematical ability. The directional hypothesis is that gifted students will demonstrate the greatest gains, followed by the above average students, followed by the average students, who will be followed by the below average students. I predict that the results from the current study will support the directional hypothesis.

Achievement groups will also be compared based on the number of steps the students advanced in the Kumon programme. The null hypothesis is that there will be no difference in the number of steps completed among the four achievement groups. The directional hypothesis is that gifted students will make the most advances, followed by above average students, followed by the average students, followed by the below average students. I predict that the results from the current study will support the directional hypothesis.

*Hypothesis Three: Kumon Mathematics Programme*

The Kumon Math programme is composed mostly of exercises that target computation skills. The null hypothesis is that there will be no difference in the gains made by students on a measure of computation skills versus gains made on a measure of mathematical reasoning skills. The directional hypothesis is that gains made on the CAT-3 computation and numerical subtest will be significantly greater than gains made on the CAT-3 Mathematics subtest (a measure of mathematical reasoning). Based on the structure of the Kumon worksheets, I predict that the results from the current study will support the directional hypothesis.

*Hypothesis Four: The Relationship between Kumon and Achievement in Mathematics*

The null hypothesis is that gains made in Mathematics cannot be predicted based on a measure of participation in the Kumon Math programme (number of worksheets completed). The directional hypothesis is that gains made on a measure of Mathematics can be predicted based on the number of Kumon Math worksheets a child completes. I predict that the results from the current study will support the directional hypothesis. Since Kumon worksheets focus on computation exercises, I predict that the results from the current study will support the directional hypothesis only for gains made in computation skills.

The second null hypothesis is that the relationship between gains made in computation skills cannot be predicted based on pretest achievement scores. The directional hypothesis is that there is a positive relationship between gains made on the computation and numerical estimation and pretest scores on the same subtest. Furthermore, a model that examines the predictive ability of the two variables (number of

Kumon worksheets completed and the pretest computation scores) together will explain more variance in gains made in computation skills than models that examine the relationships between the predictor and criterion variables separately. I predict that the results from the current study will support the directional hypothesis.

## Summary of Introduction

Mathematics skills develop at different rates among elementary school-aged children. While most children can be categorized into a “typically developing” or “average” group, there are children who develop more slowly and others that develop at an accelerated rate. The former are sometimes identified as having Mathematics Disorder, or below average achievement, and the latter are identified as above average. Children who demonstrate exceptional skills are sometimes categorized as gifted. Programmes that help build stronger Math skills are desirable for children along all trajectories of development in Mathematics, including gifted children. The purpose of the current study is to investigate whether the Kumon Math programme is effective as a remedial and/or enrichment programme.

Chapter two is an in-depth review of the current research regarding learning theories, typical development of mathematical abilities, mathematical giftedness, Mathematics Disorder, and Kumon. Chapter three describes the methods used to test the hypotheses. Chapters four and five present the results and the implications of the study, respectively. Chapter five will also include conclusions based on the results.

# CHAPTER II: LITERATURE REVIEW

Chapter two will discuss the findings and conclusions of the past studies most relevant to the current study. The purpose of this section is to interpret the current literature on this topic and provide support for how the current study will add to the current knowledge base. To achieve this goal, the literature review covers theories of learning, the typical acquisition of mathematical abilities, giftedness in Mathematics, Mathematics Disorder, Kumon, and a summary of the current state of knowledge.

## Theories of Learning

### *Behavioural Psychology*

Behaviourism was the most popular learning theory in the 1950s and 1960s (Woodward & Montague, 2002). From this perspective, learning is best characterized in terms of observable stimuli and responses, as opposed to unobservable internal factors (Gray, 1999; Hetherington, Parke, & Schmuckler, 2005). There are two processes of learning: classical conditioning and operant conditioning (Gray, 1999). Classical conditioning is defined as, “a type of learning in which individuals learn to respond to unfamiliar stimuli in the same way they are accustomed to respond to familiar stimuli if the two stimuli are repeatedly presented together” and operant conditioning is defined as, “a type of learning in which learning depends on the consequences of behaviour; rewards increase the likelihood that a behaviour will recur, whereas punishment decreases that likelihood” (Hetherington, 2005, p. 11). Operant conditioning is the most relevant process to learning Mathematics, the focus of this literature review. There are many principles and constructs used in an operant conditioning framework to explain how people learn including “reinforcement”, “punishment”, “shaping”, and “chaining” (Gray, 1999). However, the most relevant to the current study is the learning hierarchy.

*Learning Hierarchy*

According to a behavioural-analytic, operant conditioning model, knowledge and skills are developed and mastered within a learning hierarchy: acquisition, fluency, generalization, and adaptation (Daly & Martens, 1994; Haring, Lovitt, Eaton, & Hansen, 1978). Proficiency at each stage is necessary for success at later stages. In the acquisition phase, the student is acquiring response accuracy (Haring et al., 1978). Acquisition “is the first step to skill mastery” (Poncy, Skinner, & Jaspers, 2007, p. 28). In other words, acquisition is the foundation upon which fluency, generalization, and adaptation lie, and therefore has a significant influence on skill development (Poncy et al., 2007). Suggested strategies for improving accuracy include increasing the number of learning trials (Albers & Geer, 1991), modeling, demonstrating, and errorless learning (Daly, Witt, Martens, & Dool, 1997; Haring et al., 1978).

Fluency refers to the ability to perform a skill both accurately and quickly (Haring et al., 1978). Fluency is a particularly important stage because it frees attentional resources. For example, children should become so proficient in solving basic Arithmetic equations that responding is so rapid and accurate that it requires little or no conscious monitoring (Goldman & Pellegrino, 1987). With automaticity, attentional resources can be allocated to other, more difficult tasks (Goldman & Pellegrino, 1987). Fluent responding not only requires less effort, but results in higher rates of reinforcement (Codding, Eckert, Fanning, Shiyko, & Solomon, 2007). Also, students who can respond automatically may have less anxiety and increased motivation to complete tasks (Billington, Skinner, &

Cruchon, 2004; Skinner, 2002). And finally, fluency allows more opportunities to practice, which further strengthens accuracy and fluency (Skinner, Pappas, & Davis, 2005). “The primary method for improving fluency is practice” (Skinner, McLaughlin, & Logan, 1997, p. 297). Time-based performance exercises and performance feedback are other suggested strategies for improving fluency (Daly et al., 1997; Haring et al., 1978).

In the generalization stage, the student can perform the skill both accurately and fluently, but is acquiring the ability to use the skill in a variety of settings (Haring et al., 1978). Generalization is essential because for a skill to be useful, the individual must be able to respond accurately and fluently across time, people, and settings (Stokes & Baer, 1977). Stokes and Baer summarized the generalization literature and categorized different strategies for enhancing generalization. These categories include training multiple stimulus and response exemplars, including common stimuli in programming, training in multiple contexts, incorporating salient and self-mediated stimuli, and using appropriate reinforcement techniques to reward generalization when it occurs (Stokes & Baer, 1977; Stokes & Osnes, 1989).

Last, adaptation refers to the ability to modify or adapt a skill to fit novel situations (Haring et al., 1978). A skill is mastered when the student is proficient in performing the skills in each stage of learning (Haring et al., 1978). Adaptation is particularly important when learning Mathematics because mathematical skills and concepts develop hierarchically (Fuchs & Fuchs, 2005). In order to acquire more advanced mathematical skills, the individual has to adapt previously mastered skills to novel problems. In fact, solving novel problems is the suggested strategy for enhancing adaptation of skills (Daly & Martens, 1994).

Concepts and methods such as the learning hierarchy are frequently drawn from behavioural psychology for the purpose of educating children. Equally influential in the field of education, however, is cognitive psychology.

*Cognitive Psychology*

The cognitive perspective on learning began to gain popularity in the late 1960s and early 1970s (Woodward & Montague, 2002). This orientation emphasizes the influence of cognitive factors in learning and development (Hetherington et al., 2005) and describes learning in terms of stored information (Gray, 1999). Emerging from behaviourism, cognitivists also attribute learning to the processes of classical and operant conditioning (Gray, 1999). The difference between the two schools of thought is that the cognitivists also consider the mental activity that occurs between stimulus and response (Gray, 1999). The three major subtheories within cognitive developmental psychology (Piagetian theory, Vygotsky's sociocultural theory, and information-processing theory) continue to influence educational research and practice (Woodward & Montague, 2002).

Piagetian theory is defined as, "a theory of cognitive development that sees the child as actively seeking new information and incorporating it into his knowledge base through the processes of assimilation and accommodation" (Hetherington et al, 2005, p. 15). Sociocultural theory is defined as, "a theory of development, proposed by Lev Vygotsky, that sees development as evolving out of children's interactions with more skilled others in their social environment" (Hetherington et al, 2005, p. 15). The information-processing approach, however, is the most useful for explaining cognitive changes such as the development of Mathematics skills because of its many relevant metaphors, models, and concepts (Hetherington et al., 2005).

*Information-Processing*

The information-processing approach uses the computer as a metaphor for how humans learn and think (Bjorklund, 2005; Hetherington et al., 2005). Few followers of the information-processing theory believe that the mind functions exactly like a computer. Rather, the theory is used for its concepts and language which are widely used to describe and understand learning processes (Bjorklund, 2005).

The information-processing approach comprises four tenets (Siegler, 1998; Siegler, 2001). First, thinking is information-processing and therefore involves perception, encoding, representation, storage, and retrieval of information is thinking. Second, encoding, strategy construction, automatization, and generalization are the mechanisms that allow a child to learn and develop cognitively. For example, information about a Mathematics problem is encoded and then, in combination with prior knowledge, the child constructs a strategy to solve the problem. New strategies are slow and effortful but become automatic and more effective with practice. Strategies must then be generalized to novel problems. Third, the information-processing approach holds that development is driven by self-modification. In other words, the child uses previously acquired knowledge and strategies to adapt his or her responses to more difficult problems. The fourth tenet of information-processing is that both the child's abilities and the nature of the task can influence the child's performance (Siegler, 1998; Siegler, 2001).

According to this theory, information moves through a system of cognitive structures (Bjorklund, 2005). Information from the environment, perceived through one or more of the five senses, enters the sensory register and is stored there temporarily (Hetherington et al., 2005). This information, if encoded, is then sent to short-term memory (STM)

(Hetherington et al., 2005). STM has a limited capacity so information must be rehearsed (to maintain the information) or transferred to long-term memory (LTM) before it dissipates (Hetherington et al., 2005). LTM holds knowledge and strategies permanently in most cases (Hetherington et al., 2005). When a particular response is required, the knowledge and strategies contained in the LTM can be transferred and temporarily held in STM, where responses can then be executed (Hetherington et al., 2005).

The transfer of information between the aforementioned cognitive structures is aided by attention, memory, and inhibition in typically developing children. Attention is a complex process that is influenced by several variables (Hetherington et al, 2005). One such variable is maturation; as the child ages, mental resources and the ability to allocate said resources improves (Hetherington et al., 2005). To learn a complex skill such as Mathematics, children must also be able to attend to the information at hand but also ignore irrelevant information (Hetherington et al., 2005). Miller and Weiss (1981) determined that selective attention improves over time. In sum, attention influences the processing of information and improves as children develop; therefore, the ability to learn improves over time.

Memory develops and becomes more efficient over time as a result of maturational changes in other cognitive processes (Hetherington et al., 2005). First, the STM's information capacity increases and overall information-processing improves when children's processing speed improves. Speed of processing influences all aspects of information-processing and, according to Kail (2000), has the most influence on developmental changes in cognitive ability. Processing speed typically increases with maturation, rather than as a result of practice (Miller & Vernon, 1999). Second, memory strategies, such as rehearsal, information organization, and elaboration, are increasingly observed in children as they develop (Hetherington et al., 2005). These strategies also make cognitive processing more efficient. And third, knowledge of the world influences how children understand information and also how they will recall that information at a later time (Hetherington et al., 2005). In other words, as children learn and make more mental connections between stored information, cognitive processes become more efficient. As these three areas develop, so does the child's memory.

Memory can also improve with practice. When a skill is practiced repeatedly, it can eventually be performed automatically (Hetherington et al., 2005). Automatic processes require none of the STM's limited capacity, require no mental effort, are unconscious, and do not interfere with other processes (Bjorklund, 2005). Storage and retrieval of new, more complex information is enhanced when lower-level information can be retrieved from LTM automatically.

Last, inhibition is required to ensure that information flows smoothly from input to output in the information-processing model. With age, children become better at ignoring irrelevant information in the environment and suppressing inappropriate responses (Bjorklund, 2005). When there is less irrelevant and/or inappropriate material in the child's STM there are more resources available for essential cognitive operations which makes cognitive processes, such as the transfer of information between STM and LTM and learning, more efficient (Bjorklund, 2005).

In conclusion, the information-processing theory is a useful model for studying and explaining children's learning because it provides many useful terms and constructs. The concepts of this theory will be used throughout this literature review to explain the development of mathematic skills as well as to explain the results of the current study in the Discussion section (Chapter 5).

## Typical Development of Mathematical Abilities

Mathematical ability is partly innate (Butterworth, 2005; Shalev, 2004). Infants as young as four months old demonstrate the ability to compare, understand,, and respond to different quantities of items for groups containing up to four items (Starkey & Cooper, 1980). This ability has also been observed in monkeys (Nieder, Freedman, & Miller, 2002), supporting the notion that numeracy is a biologically based, innate ability. Using a habituation-dishabituation design, Wynn (1992) demonstrated that 4- to 5-month old infants are sensitive to simple addition and subtraction scenarios. Wynn (1992) placed a doll on a stage, covered the doll with a screen and then showed a second doll being placed behind the screen. When the screen was removed, infants looked longer if there were one or three dolls on the stage than if there were two (Wynn, 1992). Children, therefore, seem to have an innate ability to understand simple addition and subtraction problems.

The more advanced Mathematics skills are acquired through learning or nurture. These mathematical abilities develop hierarchically (Ardila & Rosselli, 2002); each stage (enumeration, counting, Arithmetic, and understanding arithmetical concepts) builds upon the skills acquired in the previous stage (Ardila & Rosselli, 2002; Butterworth, 2005; Klein & Starkey, 1987). Enumeration is the ability to sequence and distinguish between items in a group of objects (Ardila & Rosselli, 2002; Klein & Starkey, 1987). Infants show evidence of enumeration as young as six months old, although their number representations are often imprecise (Wood & Spelke, 2005). Enumeration continues to improve as the child develops (Wood & Spelke, 2005).

Counting is an even more complex type of enumeration, as it requires many subskills (Butterworth, 2005). Children must first learn the counting words; a skill that begins to emerge at about the age of two (Butterworth, 2005). Also in their second year, children acquire the one-to-one principle and the stable order principle (Ardila & Rosselli, 2002). The one-to-one principle is the understanding that each number word is linked with one and only one object (Gelman & Gallistel, 1978). The stable order principle embodies the concept that each number name is assigned to a permanent position in the list of numbers and that the sequence of numbers in that list never changes (Ardila & Rosselli, 2002; Gelman & Gallistel, 1978). At around age three, children will demonstrate the third principle required to be able to count: the cardinal principle (Butterworth, 2005; Gelman & Meck, 1983). The cardinal principle is the understanding that a collection of items contains the number of objects corresponding to the last counting word used (Ardila & Rosselli, 2002; Gelman & Gallistel, 1978). Children continue to refine these counting skills until about the age of six (Butterworth, 2005).

"Counting is the basis of Arithmetic for most children" (Butterworth, 2005, p. 8). Addition, therefore, is commonly taught by building on the child's counting skills. There are three addition strategies that make use of counting skills: counting all, counting on from first, and counting on from larger (Butterworth, 1999; Carpenter & Moser, 1982). Counting all develops first (Butterworth, 1999; Carpenter & Moser, 1982). Children in this stage represent both numbers in the addition problem using physical objects (their fingers, for example) and will then count all the objects (Butterworth, 1999; Carpenter &

Moser, 1982). In the next stage, counting on from first, children will no longer count the first addend and will instead start with that number and then use an external aid (such as their fingers) to add on the second number (Butterworth, 1999; Carpenter & Moser, 1982). The last stage (counting on from larger) is the most efficient (Butterworth, 1999; Carpenter & Moser, 1982). The child selects from the equation the largest number and then uses an external aid (such as fingers) to count on the smaller addend (Butterworth, 1999; Carpenter & Moser, 1982). Subtraction develops similarly. The acquisition of addition and subtraction continues into the second grade.

The next stage of development is fluency with Arithmetic facts. At this stage, children no longer need to mentally add or subtract the numbers in an equation. Instead, the solution is quickly and accurately retrieved from memory (Butterworth, 2005). Siegler and Shrager (1984) suggest that over time, children learn to associate specific answers with specific equations.

At around the age of eight or nine children begin to learn multiplication and division (Ardila & Rosselli, 2002). These new skills are also built upon previously mastered skills; multiplication and division are frequently explained in terms of repeated addition and subtraction (Butterworth, 2005).

Mathematical reasoning is a distinct domain of mathematical ability (Fuchs et al., 2008). It is defined as axiomatic reasoning, logical deduction, formal inference, and problem-solving (Arthur Steen, 1999). Relatively little is known about how mathematical reasoning develops (Arthur Steen, 1999). There is evidence, however, that suggests that logical and mathematical thinking can develop naturally through social interaction, games, commercial transactions, and discussions with others (Schliemann & Carraher, 2002). This is not to say that mathematical reasoning cannot be further developed via formal instruction. Rather, Arthur Steen (1999) suggests that active classroom tasks, such as discussion, projects, and teamwork, will be more effective than passive strategies, such as memorization, drill, and templates, in helping children develop mathematical reasoning.

In both research and practice, mathematical reasoning is often secondary to computation. Mathematics teachers often focus on teaching students how to do Mathematics with little focus on ensuring that their students understand how to reason mathematically (National Center for Educational Excellence, 1996). It is not surprising then, that Zhou, Peverly, and Lin (2005) found that first grade Chinese and American children's numerical operation skills were better developed than their mathematical reasoning skills. The imbalance in knowledge and emphasis between computation and mathematical reasoning is not due to differences in their importance. Mathematical reasoning is necessary for solving word problems, finding patterns, and verbalizing logical explanations (Zhou et al., 2005). Further research is needed on the typical and atypical development of mathematical reasoning ability.

## Gender Differences in Mathematical Ability

There is a widespread mainstream belief that males have a greater aptitude for Mathematics than females. This belief can be traced as far back as 1894 when Havelock Ellis wrote about male's cognitive superiority in a variety of domains, including Mathematics. Since then, gender differences in mathematical ability have been the focus of many research studies and the source of great debate.

Many early studies reported significant gender differences in mathematical ability, with the results favouring males. Research in this area continues to this day and the results are often conflicting. Some researchers report no gender difference; others find gender differences, sometimes in favour of males, other times the gender difference favours females. For example, Demie (2001), Gorard, Rees, and Salisbury (2001), and Penner (2003) have found gender differences in mathematical ability. The later researcher found that gender differences favoured males in all areas measured. The conclusion of the two other studies was that females outperformed males in basic arithmetic. In contrast, Ding, Song, and Richardson (2007), Georgiou, Stavrinides, and Kalavana (2007), and Leahey and Guo (2001) have found that there are no significant gender differences in mathematical ability among elementary school aged children. The reason for paradoxical results can likely be attributed to different research designs (Aunio, Aubrey, Godfrey, Pan, & Liu, 2008). Researchers may have come to different conclusions because of sample differences (for example, cultural differences and/or age range), and/or type of Math skill investigated.

Meta-analysis has revealed that the magnitude of gender differences varies with age. Gender differences seem to become pronounced near the end of high school (Georgiou et al., 2007; Hyde, Fennema, & Lamon, 1990; Leahey & Guo, 2001). Differences between the sexes increase slightly in college and adulthood (“effect sizes = .29, .41, and .59 for high school, college, and adult samples, respectively” (Bjorklund, 2005, p. 418)).

Gender differences may also vary depending on the specific Mathematics skill evaluated. Some research shows that females perform better on computation tasks, males do better on problem-solving tasks, and there is no gender difference in comprehension of mathematical concepts or basic mathematical knowledge (Anastasi, 1958; Demie, 2001; Gorard et al., 2001; Hyde et al., 1990; Stage, Kreinberg, Eccles, & Becker, 1985). Females demonstrate superiority in computation as early as in the elementary years while gender differences favouring males in problem solving do not emerge until the high school and college years (Hyde et al., 1990). Research in this area may have advanced and led to new conclusions in recent years. Unfortunately, recent research could not be located in a PsycInfo search.

Although gender differences are relatively small among the majority of elementary school aged children, it is still important to determine the origin of slight differences. Some researchers attribute gender differences to gender socialization. There is some evidence to support this hypothesis. Teachers, parents, and peers often have negative feelings about females acquiring advanced Math skills (Fox, 1976), which results in less encouragement for females to pursue advanced Mathematics. Some research has also shown that there are gender differences in interest in Mathematics. One possible cause for this disinterest is popular media; the message in popular media often discourages females from taking interest in Mathematics (Bjorklund, 2005). Last, some researchers have demonstrated that females tend to have lower self-confidence in Mathematics and higher Math anxiety than males (Felson & Trudeau, 1991). In sum, the gender socialization hypothesis is that females are conditioned by society to believe that they have a lower aptitude for Mathematics, that society expects males to demonstrate better performance in Mathematics than females, and that females will not need advanced Mathematics skills (Felson & Trudeau, 1991; Persson Benbow & Stanley, 1983).

Some researcher's results have challenged the socialization hypothesis. Felson and Trudeau (1991) found no gender differences in parental encouragement in Mathematics and they found that although females have greater anxiety about Mathematics, they have greater general test anxiety, rather than a specific fear of Mathematics (Felson & Trudeau, 1991).

An alternate explanation for gender differences in Mathematics is biologically-based. Some researchers have proposed that gender differences in brain lateralization are the cause of gender differences in mathematical ability (Halpern, 1986). In addition, several reports have shown that males demonstrate different brain activation patterns than females when solving Mathematics problems, suggesting hard wired differences between the sexes (Hyde, 2007). This hypothesis has also been challenged. It could be that "males and females have different experiences related to Mathematics as they grow up, and that these different experiences have, on average, enhanced synaptic connections in some regions for males and in other regions for females" (Hyde, 2007, p. 262). Further evidence against a biologically-based hypothesis comes from meta-analyses. Analysis shows that the magnitude of the gender differences in mathematical ability has declined in the past several decades (Bjorklund, 2005). It is highly unlikely that drastic evolutionary changes in neuroanatomy could have occurred over such a short time span.

There is another alternative explanation for the gender differences in mathematical ability. It may be that differences in background knowledge and cognitive strategies lead to differences in mathematical ability (Bjorklund, 2005). Byrnes and Takahira (1993) found that males scored higher on tests of mathematical strategies, background knowledge, and aptitude [Scholastic Aptitude Test (SAT)]. Knowledge of strategies and background knowledge accounted for 50% of the variance in the SAT scores and when these two factors were controlled for statistically, gender differences in SAT scores were no longer significant (Byrnes & Takahira, 1993). Questions still remain, however, such as why there are differences in background knowledge and strategies in the first place, since grade-point average and number of Mathematics courses completed were comparable for males and females in Byrnes and Takahira's (1993) study. The researchers proposed that the source of gender differences in Mathematics is likely multifactorial, "including socialization, physiological, and cognitive factors" (Bjorklund, 2005, p. 420).

The magnitude of the gender difference in mathematical achievement is greatest among individuals who score in the top fifth percentile, with the gender difference favouring males (Bjorklund, 2005; Preckel et al., 2008). This differential representation of males at the upper extreme is likely due to the fact that males demonstrate significantly more variance in test scores than females (Feingold, 1992; Hedges & Nowell, 1995). "Sex differences in variance and mean lead to substantially fewer females than males who score in the upper tails of the Mathematics and Science ability distributions…." (Hedges & Nowell, 1995, p. 45). As a result, males may be more likely to enter and excel in mathematically-based occupations (Hedges & Nowell, 1995).

In conclusion, for most of the population, the difference between males and females in mathematical ability is small. The gap widens slightly with age, across different types of Mathematics skills, and at the upper extreme of ability levels. In other words, "a general statement about gender differences is misleading because it masks the complexity of the pattern" (Hyde et al., 1990, p. 151). It is important to understand the pattern, sources, and implications of gender differences in mathematical ability because misinformation and stereotypes can negatively influence people's self-image, self-esteem, achievement, occupational status, and earnings (Hedges & Nowell, 1995).

## Giftedness in Mathematics

Giftedness is often defined as extraordinary cognitive ability paired with extraordinary specific knowledge and achievement in a particular domain (Lohman, 2005; Threlfall & Hargreaves, 2008). Still, the criterion for identifying giftedness often only consists of a score above a certain cutoff on a measure of intellectual ability. Many school districts and researchers have selected an IQ of 130 as a cutoff score on a measure such as the Wechsler Intelligence Scale for Children–Fourth Edition (WISC-IV) in order to be formally identified as gifted (Mills et al., 1994; Montague & van Garderen, 2003; Tsui & Mazzocco, 2007; van Garderen & Montague, 2003). Other researchers use domain specific achievement scores to group children as "gifted" (Ma, 2005; Mills, Ablard, & Gustin, 1994; Niederer, Irwin, Irwin, & Reilly, 2003; Swanson, 2006; Threlfall & Hargreaves, 2008). This method allows researchers to group children who *may be* gifted. Without a measure of cognitive ability, children cannot be formally classified as gifted. As such, researchers who use achievement scores to group children into ability categories typically label the highest scoring group as "precocious" or "talented" to highlight the fact that these children are not formally identified as gifted.

### *Gender Differences*

Among the mathematically gifted, there are more males than females (Preckel et al., 2008). Benbow and Stanley (1983) compared the number of males and females who scored above 700 on the SAT-Mathematics and found a male-to-female ratio of 13:1. Gender differences in mathematical ability within the general population have been decreasing since the 1960s and the disparity continues to get smaller except among the gifted, where males continue to be disproportionately represented (Bjorklund, 2005).

*Qualitative versus Quantitative Differences*

Same-age comparative studies have shown that gifted learners differ from their peers in several ways (Shore & Kanevsky, 1993; Steiner & Carr, 2003). Threlfall and Hargreaves (2008) stated that gifted learners:

- have a broader and more inter-connected knowledge base
- are quicker at solving problems, while spending more time planning
- are more efficient at representing and categorizing problems
- have more elaborate procedural knowledge
- are more flexible in their use of strategies
- prefer complex, challenging problems [and]
- are more sophisticated in their meta-cognition, including self-regulation. (p. 84)

Reichel (1997) noted several signs of mathematical giftedness, specifically. When students create new words, symbols, and sketches, look for problems, have imagination, reduce complexity, formalize and generalize ideas, are amazed by facts and formulae, and want to discuss Mathematics, they are demonstrating signs of mathematical giftedness (Reichel, 1997).

There has been debate and conflicting evidence as to whether these differences and the superior ability/achievement of gifted children are due to qualitative (life-long, fundamental differences in thinking processes and strategies) or quantitative differences (precocity) between them and typically developing children (Shore & Kanevsky, 1993; Winner, 2000a). Many researchers purport that mathematically gifted children are qualitatively different from average-ability children because the gifted “take more time to orient to a problem, to utilize a wider range of problem-solving strategies, and to evaluate their progress both during and after completing a problem” (Niederer et al., 2003, p. 72).

However, it is unclear whether these differences remain over time and therefore whether they are truly qualitative differences. The majority of evidence supporting the theory that gifted children develop and process information in a qualitatively different way from typical children is anecdotal (Winner, 2000a). Systematic research is needed to determine whether extraordinary mathematical ability is always accompanied by qualitative cognitive differences throughout gifted individuals' lives (Winner, 2000b).

In contrast, Threlfall and Hargreaves (2008) propose that the difference between gifted and non-gifted children is quantitative. Threlfall and Hargreaves (2008) compared how nine-year old gifted students and 13-year old average-achieving students solve Mathematics problems. The results of this study suggest that in at least some respects, young, gifted children solve mathematical problems with a similar approach to that of older children of average ability (Threlfall & Hargreaves, 2008). This research supports the position that the difference between gifted and average ability children is a matter of precocity (a quantitative difference) (Threlfall & Hargreaves, 2008).

There are other researchers, still, who say that gifted children are both quantitatively and qualitatively different from average-achieving children (Winner, 2000a). There is clearly a strong need for further systematic research to solve this debate.

### *Nature versus Nurture*

It is a dominant cultural belief that giftedness is a product of the individual's genetic makeup and that it is innate (Winner, 2000b). Although there is some evidence to that support this position, research has not supported the nature perspective unequivocally (Threlfall & Hargreaves, 2008).

A contrasting view, the “expert performance” approach, proposes that there is a clear link between giftedness and deliberate practice (Ericsson, Roring, & Nandagopal, 2007). There is scientific evidence supporting this position. Researchers have shown that extraordinary achievement in a variety of domains (Science, Arts, Mathematics, athletics, piano, violin, chess, etc.) is predicted by the amount of time the individual spends practicing (Ericsson, Krampe, & Tesch-Romer, 1993). Despite this and other evidence of the importance of nurturing giftedness, no evidence “rules out the contribution of genetic factors as a source of individual differences in who will be able to develop a given amount of expertise” (Sternberg, 2001, p. 161). In other words, it is most likely that both nature and nurture contribute to giftedness.

### *Educating the Mathematically Gifted*

Giftedness is often overlooked as an area in need of research and interventions as giftedness is often viewed as something to be admired or envied, rather than something that needs to be addressed with an intervention (Winner, 2000b). Also, advocates of the nature perspective believe that giftedness is an innate attribute that cannot be developed through training and intervention (Winner, 2000b). In reality, much is to be gained by providing gifted children with interventions and/or enrichment programmes that are tailored to their needs.

If the special education needs of gifted children are not met, these children are likely to be underserved (Mills et al., 1994; Silverman, 1989). An “under-stimulating Mathematics curriculum results in the loss of motivation, under-achievement, and the denial of the value of education among mathematically talented students” (Ma, 2005, p. 105). Gifted children need an appropriate level of challenge to ensure their happiness, mental health, and social integration (Winner, 2000b). Also, gifted children will become our future leaders and innovators; it is in all of our best interest to provide them with an education that will maximize their potential (Winner, 2000b). The best way to ensure that gifted students' special needs are met and that they will reach their full potential is to provide them with individualized education (Lubinski & Persson Benbow, 2006; Mills et al., 1994).

Research suggests that "the most effective method to provide [gifted] students with the necessary challenge and appropriate pace of learning is acceleration" (Mills et al., 1994, p. 496). This is a process of advancing the student by providing him or her with material that is at an appropriately challenging level (Feldhusen, 1989). Hundreds of studies have shown that this is an effective method for enhancing the achievement of gifted students (Lubinski & Persson Benbow, 2006). What's more, the positive effects of acceleration are long term. Participants in a fast-paced, intellectually rigorous Mathematics programme were two times more likely to have Mathematics- or Science-based careers than non-participants 10 to 20 years after completion of the programme (Swiatek & Benbow, 1991). In sum, an effective gifted programme will include acceleration.

Research also supports the use of flexible pacing for the mathematically gifted (Stanley, 1991; Winner, 2000a). Mills and colleagues (1994) evaluated the effect of a Mathematics programme that incorporated flexible pacing. More specifically, the programme integrated four strategies: (a) the curriculum was linear (students did not advance until they demonstrated mastery), (b) the curriculum was flexibly paced (the pace was determined by each individual's unique rate of achievement), (c) no age or grade restrictions were placed on students, and (d) instructional placement was determined by the student's achievement level at the time of entry in the programme (Mills et al., 1994). Third, fourth, fifth, and sixth grade mathematically gifted students demonstrated greater achievement gains over a 7- month period than controls (Mills et al., 1994). These gains were maintained after a period of five months (Mills et al., 1994).

Another major benefit of Mills and colleagues' (1994) Mathematics programme was that it provided a venue for like-minded children to meet and interact. Although gifted children tend to be more introverted and enjoy solitude more than average-achieving children, they still desire peer contact and friendships (Winner, 2000b). Gifted children especially desire interaction with like-minded peers (Winner, 2000b). Enrichment programmes that bring gifted students together address this need.

In practice, the most common intervention for gifted children is placement in an enrichment or advanced class within the child's school. These classes combine effective strategies for educating the gifted such as advanced placement and ability grouping (Winner, 2000b). Advanced classes are common at the secondary level but rare at the elementary level (Winner, 2000b). This is unfortunate because enrichment classes are needed at all school levels as giftedness occurs at all ages.

Enrichment classes are most often headed by a special education teacher. This is an important component of the programme as gifted students still need support despite their advanced knowledge. "Enrichment that consists of 'busy work' or irrelevant topics has limited academic value for gifted students" (Diezmann & English, 2001, p. 13). The teacher is required to design an enrichment programme that is appropriately challenging. The quality of this programme, therefore, will depend on the teacher's training, theoretical orientation, skill, personality, etc. There is much variability in these teacher variables and thus there is much variability in the quality of enrichment programmes. Despite this disadvantage, informal enrichment programmes are commonly used interventions for gifted students.

Formal programmes are employed much less frequently as interventions for gifted students although are many advantages to using formal programmes. Formal programmes provide the consistency that is lacking in informal programmes as instruction and evaluation are standardized. Also, programmes are periodically reviewed and revised by multiple professionals which often leads to a superior product. Third, formal programmes are more likely to be designed based on theory and research. The disadvantages of formal programmes are that they often cost money and that they can be too rigid and not individualized to the unique client's needs. A formal programme that also allows for individualization may be the optimal intervention for gifted children. Unfortunately, there is currently no research on programmes of this nature to the best of the author's knowledge.

Giftedness is defined as extraordinary performance or ability in one or more domains. Further research is needed to resolve the debate regarding whether the differences between gifted and non-gifted children are quantitative, qualitative, or both. Similarly, there is a nature-nurture debate as to the origins of giftedness; likely, both nature and nurture contribute. Last, research on interventions for gifted children has largely been ignored. The effectiveness of formal programmes as interventions for gifted students is in particular need of research.

# Mathematics Disorder

In comparison to other disorders, including Reading Disorder (RD), Mathematics Disorder (MD) is far less researched and understood (Geary, 1993; Mazzocco & Myers, 2003). The knowledge gap is so vast that a universally accepted definition for MD does not exist (Mazzocco & Myers, 2003). One proposed definition is an "unexpected difficulty in the Arithmetic module, conceptual knowledge base, or problem-solving space of the domain-specific functional Math system, given the student's verbal, quantitative, and/or visual-spatial reasoning ability" (Busse, Berninger, Rury Smith, & Hildebrand, 2001, p. 151). Adding to the confusion, a variety of labels has been used to describe the same disorder throughout the scientific literature (Geary & Hoard, 2001). Terms such as *Developmental Dyscalculia* (Mazzocco & Myers, 2003; Shalev, 2004; Shalev & Gross-Tsur, 1993), *Mathematical Disability* (Geary, 1993), *Arithmetic Learning Disability* (Geary & Hoard, 2001; Koontz & Berch, 1996), *Arithmetical Disability* (Geary, 2003), and *Number Fact Disorder* (Temple & Sherwood, 2002) have all been used to refer to the same disorder. The author of the current article will use the term *Mathematics Disorder* (MD), as MD is the terminology used by the DSM-IV-TR.

## *Diagnosis of Mathematics Disorder*

The diagnostic criteria for MD according to the DSM-IV-TR (2000) are:

> (a) mathematical ability, as measured by individually administered standardized tests, is substantially below that expected given the person's chronological age, measured intelligence, and age-appropriate education; (b) the disturbance in Criterion A significantly interferes with academic achievement or activities of daily living that require mathematical ability; and (c) if a sensory deficit is present, the difficulties in mathematical ability are in excess of those usually associated with it. (p. 54)

MD occurs in spite of emotional, environmental, cultural, social, or motivational factors (APA, DSM-IV-TR, 2000; Learning Disability Association of Ontario, [LDAO], 2009). The DSM-IV-TR's diagnostic criteria are based on a discrepancy model. The DSM-IV-TR specifies that the discrepancy between cognitive ability and academic achievement must be significant, but it does not specify the level of significance.

Often, a difference of two or more standard deviations (*SD*) between cognitive ability and academic achievement scores is defined as a significant discrepancy (Shalev, 2004). Other researchers and practitioners use a variation of the discrepancy-based model. For example, some researchers' criteria for inclusion in an LD group are a low-average to high IQ score combined with a ranking below the 25th percentile on a measure of mathematical achievement (Geary, 2003). However, the cutoff rank on the academic measure varies widely (from the 8th to 48th percentile) among researchers (Swanson & Jerman, 2006).

There is a third variation of the discrepancy-based model. Within this model, if there is a discrepancy of at least 2 years between the child's actual grade and his or her grade level of achievement on a measure of Mathematics, that child meets criteria for a diagnosis of MD (Hammill, 1990).

Although the discrepancy-based model is used frequently in the diagnosis of MD, there is evidence that this model is inappropriate and ineffective (Mazzocco & Myers, 2003). First, the discrepancy-based model sometimes leads to false negatives (Mazzocco & Myers, 2003) and/or false positives (Geary, 1990; Geary, Brown, & Samaranayake, 1991). For example, children who score in the gifted range on a measure of cognitive ability and in the average range on a Mathematical achievement measure would be diagnosed with an MD. Although these children are underachieving, a diagnosis of MD may not be appropriate. Second, the commonly used cutoff percentile ranks (between the

$10^{th}$ and $25^{th)}$ on Mathematics achievement tests does not reflect the reported prevalence rates of MD (between 5 and 8%) (Geary, 2003).

An alternate diagnostic model is the criterion-based model. According to this model, cognitive ability does not need to be considered in the identification of a Learning Disability (LD). For example, some researchers have classified children as having a MD if they rank below the $45^{th}$ percentile on a measure of mathematical achievement regardless of their intellectual ability (Mazzocco & Myers, 2003). Although this model may be the most efficient for use in research, it will lead to many false positives as it fails to differentiate between children with LDs and children with Mental Retardation (MR).

Change over time is also used as a method of diagnosing a LD (Geary, Hoard, & Hamson, 1999). MD, for example, is indicated when students perform poorly in Mathematics in two or more consecutive grade levels (Geary et al., 1999). This criterion would likely lead to many false positives and may, therefore, be better used as an adjunct to either a discrepancy- or criterion-based model. The DSM-IV-TR uses these criteria to diagnose a MD. Not only does the DSM-IV-TR require that the individual demonstrates a significant discrepancy between cognitive ability and mathematical achievement, but also that the mathematical difficulties incurred by the individual causes significant impairments in academic achievement (at school).

Variations in diagnostic models and criteria limit practitioners' and researchers' abilities not only to accurately diagnose the disorder, but also to determine the core deficits, existence of subtypes, prevalence, prognosis, etiology, characteristics, and effective treatments for MD. Universally accepted diagnostic criteria are an essential future endeavor.

*Subtypes*

Children with MD are heterogeneous in their mathematical achievement (Geary, 1993; Kronenberger & Dunn, 2003; Mazzocco & Myers, 2003). Based on cognitive theory, research on acquired dyscalculia, and behaviour genetic studies of mathematical ability, Geary (1993) divided MD into three subtypes: Semantic Memory, Procedural, and Visuo-spatial. The Semantic Memory subtype is characterized by difficulty retrieving mathematical facts, frequent errors in fact retrieval, and variable response times for correct retrieval (Geary, 2003; Mazzocco & Myers, 2003). Studies have shown that inefficient inhibition of irrelevant associations [for example, answering 7 or 3 to the problem 6 + 2 (the numbers that follow 6 and 2)] disrupts individual's ability to retrieve Mathematics facts from long-term memory (Geary, Hamson, & Hoard, 2000). Retrieval errors occur because irrelevant information cannot be inhibited from entering working memory which reduces STM capacity which in turn reduces the amount of mental resources that can be allocated to making correct associations (Geary, 2003). The existence of the Semantic Memory subtype "has received the most consistent support across studies of MD" (Mazzocco & Myers, 2003, p. 223).

The Procedural subtype is characterized by developmentally immature procedures, frequent errors in procedures, poor comprehension of the concepts underlying procedures, and difficulty sequencing steps in complex procedures (Geary, 2003). STM deficits and poor conceptual knowledge of procedures contribute to procedural deficits in children with MD (Geary, 2003; McLean & Hitch, 1999; Swanson, 1993).

Third, the Visuo-spatial subtype is characterized by difficulties in spatially representing mathematical information and relationships as well as poor comprehension of spatially represented information (Geary, 2003). Geary, Hoard, Byrd-Craven, and DeSoto (2004) suggest that these deficits are due to poor monitoring of the sequence of steps in a mathematical problem and poor detecting and correcting of errors, rather than deficits in spatial abilities.

There is preliminary scientific evidence that MD subtypes exist. For example, there is some evidence that heterogeneity in mathematical deficits reflect abnormalities in related neural structures (Swanson & Jerman, 2006). Other researchers have challenged the existence of MD subtypes. Further research is needed to resolve this debate because there may be multiple underlying deficits among individuals with MDs which would have significant implications for treatment (Mazzocco & Myers, 2003).

*Prevalence*

Prevalence estimates of MD vary widely in the literature due to variations in the definition and diagnostic criteria for MD (Butterworth, 2005; Mazzocco & Myers, 2003). Estimates range from as low as 1% of school-age children (APA, DSM-IV-TR, 2000) to as high as 8% (Ostad, 1998). Variations in prevalence rates may also be due to the fact that there has been no large-scale epidemiological study of the prevalence of MD (Geary, 2003). The majority of currently existing, well designed studies report prevalence rates between 5 and 8% (Geary, 2003, Gross-Tsur et al., 1996; Lewis et al., 1994; Shalev et al., 2000). Prevalence ratios are equal between girls and boys (Lewis et al., 1994; Gross-Tsur et al., 1996; Shalev, 2004).

MD typically occurs as a specific learning disability (Shalev, 2004), however, attention-deficit/hyperactivity disorder (ADHD) and RD sometimes occur comorbidly (Geary, 2003). Gross-Tsur and colleagues (1996) found that 26% of children with MD had comorbid ADHD and 17% had comorbid RD. Just as prevalence rates vary throughout the literature, so do comorbidity rates. Ostad (1996), for example, found that half of the children identified as having a MD had a comorbid Written Expression Disorder and Badian (1983) found that almost half of children with a MD had a comorbid RD. Children with comorbid disorders are more profoundly impaired in Arithmetic than children with MD alone (Jordan, Hanich, & Kaplan, 2003; Shalev, Manor, & Gross-Tsur, 1997). Children with MD alone also outperform co-morbid children (RD and MD) on measures of literacy, visual-spatial problem solving, LTM, STM for words, and verbal working memory (Swanson & Jerman, 2006).

*Prognosis*

The symptoms of MD may be expressed as early as Kindergarten and is usually diagnosed before children enter the fifth grade (APA, DSM-IV-TR, 2000). Typically, referral occurs during the second or third grades when formal instruction is introduced as difficulties in Mathematics become more apparent at this level.

Children’s academic strengths and weaknesses change over time and therefore not every delay in Mathematics is an indication of a MD (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1994; Shalev, 2004). In other words, not all children display stable, lifelong difficulty with Mathematics (Mazzocco & Myers, 2003). Mazzocco and Myers (2003) found that only 65% of children identified as having a MD in Kindergarten were classified as having persistent MD in grades one, two, and/or three and 63% of children who had been classified as MD in any grade (Kindergarten to grade three) were identified as having a MD for 2 or more years. Shalev and colleagues (1998) found that of 123 grade 4 students classified as having a MD, 47% were classified as having persistent MD three years later, yet all participants performed poorly in Mathematics (ranked in the lowest quartile). Also, Mazzocco and Myers (2003) found that RD occurred more frequently in children with persistent MD than children with non-persistent MD (25% and 7% incidence rates, respectively). Similarly, severity of the MD at the time of initial assessment and the presence of significant problems in Mathematics in siblings were associated with persistent MD in Shalev and colleagues' (1998) study (socioeconomic status, gender, presence of another LD, and educational interventions were not).

In sum, ability in Mathematics seems to fluctuate significantly between Kindergarten and grade 3 and to be more stable in grades four and up. Of all children diagnosed with a MD, approximately half of these individuals continue to meet criteria for a diagnosis throughout their lives although the expression of the disorder likely changes with age.

*Etiology*

The etiology of MD is multifactorial, including genetic, environmental, neurological, and psychological factors (Kronenberger & Dunn, 2003; Shalev, 2004). Heredity was proposed as a cause of MD as early as 1974 (Kosc, 1974). Alarcon and colleagues' (1997) study indicated that MD is in fact significantly heritable. This twin study demonstrated that 58% of monozygotic co-twins and 39% of dizygotic co-twins had a MD (Alarcon et al., 1997). In other words, MD was 8-12 times more prevalent in twin pairs than in the general population (Alarcon et al., 1997; Shalev, 2004). In a family study, Shalev and colleagues (2001) found that about half of all siblings with a MD have a MD themselves; siblings were 5 to 10 times more likely be to diagnosed with MD than members of the general population (Shalev et al., 2001). Genetics studies strongly suggest that MD is partly hereditary.

"Functional neuroimaging reveals that the parietal lobes, especially the intraparietal sulci, are active in numerical processing and Arithmetic" in typically developing individuals (Butterworth, 2005, p. 14; Dehaene et al., 2003). Functional neuroimaging suggests that different brain areas are activated when individuals with a MD are presented with mathematical tasks (Shalev, 2004). For example, when shown an Arithmetic exercise, the brain areas activated in young adults with a MD were the right intraparietal sulcus and left middle frontal gyrus, compared to controls who showed activation in the right and left intraparietal sulcus (Morocz et al., 2003). Research has also shown anatomical differences in brain structures between individuals with a MD and controls. Molko and colleagues (2003) found anatomical disorganization within the right intraparietal sulcus among individuals with Turner Syndrome who demonstrated difficulty with numeracy. Issacs and colleagues (2001) compared two groups of adolescents; one group had average abilities in Mathematics and the other displayed deficits in numerical operations. The group with arithmetical deficits had comparatively less grey matter in the left intraparietal sulcus (Isaacs et al., 2001). In sum, research suggests that individuals with a MD activate different brain regions and have anatomical differences in brain structure compared to typically developing individuals (Shalev, 2004).

Environmental and psychological factors can exacerbate difficulties in Mathematics but they cannot cause MD. Environmental influences include speed of teaching (Cuming & Elkins, 1999), class size, and Mathematics curricula (Miller & Mercer, 1997) and psychological influences include Mathematics anxiety (Shalev, 2004). These factors should be considered in the diagnosis and treatment of MD as they influence the expression and persistence of MD.

In conclusion, genetic and neurological factors can lead to the development of a MD while environmental and psychological factors can further exacerbate the disorder. It is most likely that MD develops as a result of a combination of the aforementioned variables, as opposed to any one factor in isolation. It is not surprising, then, that the expression of MD is so diverse.

## *Qualitative versus Quantitative Differences*

The majority of children (typically developing and otherwise) experience difficulty with Mathematics at some point in their life. The difference between the difficulties expressed by typically developing children and children with a MD can be described as quantitative or qualitative. Behaviourists theorize that any difficulties that arise in acquiring a skill, including Mathematics, occur at specific stages of the learning hierarchy. The difference between individuals with MDs and typically developing individuals is quantitative; individuals with MDs experience more difficulties at all stages in comparison with typically developing individuals. Cognitivists, contrastingly, theorize that the difficulties expressed by individuals with a MD are qualitatively different from those expressed by typically developing individuals because people with a MD have qualitatively different psychological processing abilities.

### *Learning Hierarchy*

Children with a MD express difficulty at all four stages of the learning hierarchy: acquisition, fluency, generalization, and adaptation (Daly & Martens, 1994; Haring et al., 1978). Shalev (2004) defines MD as “a specific learning disability affecting the normal acquisition of Arithmetic skills" (p. 765); children with LDs often struggle with the acquisition of new skills. For children with developmental disabilities, fluency in math skills is particularly difficult (Jolivette et al., 2006). Even if math skills do become accurate and fluent, children with a MD often have difficulty generalizing those skills (Scott, 1993). Although evidence of difficulty in the adaptation stage did not surface in a recent PsycInfo search, it is reasonable to assume that this stage, too, would be challenging for a child with a MD.

*Information- Processing*

The difficulties expressed by children with a MD can be attributed to four types of deficits in psychological processes: visual-spatial processing, attention, memory, and strategy use (Augustyniak, Murphy, & Kester Phillips, 2005; Shalev, 2004; Strang & Rourke, 1985). According to information-processing theory, deficits in these processes will lead to overall deficits in the ability to process information and thus the ability to learn.

Visual-spatial processing deficits refer to the misinterpretation and misunderstanding of visual information and difficulties in the spatial organization of mathematical information (Augustyniak et al., 2005; Geary, 2003; Strang & Rourke, 1985). Common errors experienced by individuals with visual-spatial processing deficits when working on Mathematics problems include difficulties placing numbers in columns and errors in reading Arithmetic signs (Ardila & Rosselli, 2002). Geary (1993) found that this cognitive profile was so prevalent and consistent among a subgroup of children diagnosed with a MD that it warranted classification as a specific subtype of MD.

Attentional difficulties are also common among children with a MD (Ardila & Rosselli, 2002). Rosenberger (1989) found that children with significant math difficulties scored higher on a measure of inattention and lower on a measure of freedom from distractibility than controls. Although Rosenberger (1989) found no significant differences between children with and without significant math difficulties on a measure of hyperactivity, impulsivity and poor peer relations, Gross-Tsur and colleagues (1996) found that 26% of children with a MD had comorbid ADHD. Another study found symptoms of Attention Deficit Disorder (ADD) in 32% of children with a MD (Shalev, Auerbach, & Gross-Tsur, 1995). Overall, poor regulation of attention is a distinct characteristic of many individuals with significant problems in Mathematics.

Memory deficits, such as difficulties in memory retrieval and working memory, are frequently identified skill deficits in individuals with a MD. Koontz and Berch (1996) compared children with a MD and a control group on digit and letter span tasks. Results showed that children with a MD had significantly lower scores than control children on both tasks (Koontz & Berch, 1996). Poor numerical digit span has been associated with MD in several other studies (Geary & Brown, 1991; Swanson & Beebe-Frankenberger, 2004). However, there have been other studies that compared the working memory of children with and without a MD and the results have been less conclusive (McLean & Hitch, 1999; Temple & Sherwood, 2002), suggesting that working memory deficits are not sufficient to cause a MD (Butterworth, 2005). Overall, though, it is widely accepted that children with a MD have memory difficulties and that these memory deficits contribute to difficulties in performing Mathematics (Geary, 1993; Geary & Hoard, 2001; Ginsburg, 1997; Jordan & Montani, 1997; Shalev & Gross-Tsur, 2001).

Children with LDs in general tend to overextend and inconsistently apply cognitive strategies (Augustyniak et al., 2005). Children with a MD in particular tend to use immature strategies (Geary, 1994) such as using fingers as an aid in single digit addition problems beyond the developmentally appropriate age (approximately age seven). Immature strategies and poor strategy execution are likely to lead to inefficient and unreliable cognitive processing and thus, poor academic performance.

In sum, individuals with LDs tend to have specific psychological processing difficulties that impede the individual’s ability to acquire, retain, and use knowledge and skills, such as Mathematics. Visual-spatial processing, attention, memory, and strategy use are noted as particular weaknesses among children with a MD. Information-processing theorists would argue that these qualitative differences in psychological processing between typically and atypically developing children are the reason children with a MD experience significantly more difficulty acquiring Mathematics skills than typically developing children.

### *Treatment*

Since the overall research on MD is relatively rudimentary, it is not surprising that the research on interventions for children with a MD is extremely limited. Codding, Eckert, and colleagues (2007) noted this dearth of research investigating the effectiveness of Mathematics interventions. The treatments that are described are diverse, reflecting the various definitions and etiologies within the existing literature (Kronenberger & Dunn, 2003). The reported interventions can be divided into two categories: strategies and formal programmes (where a *programme* is defined as a collection of strategies that were selected based on theory and research in order to achieve a specific goal and a *formal programme* is a programme that has been designed and evaluated objectively).

Augustyniak and colleagues (2005) suggest several strategies for remediation of specific deficits in Mathematics. For example, for students whose difficulties in Mathematics are primarily caused by visual-spatial deficits, Augustyniak and colleagues (2005) suggest providing students with copies of the Math problems (instead of having them copy from the chalkboard), providing students with grid paper, and using visual aids, such as an overhead projector. Kronenberger and Dunn (2003) suggest flash cards and repeated practice as interventions to build fluency with the multiplication tables. Last, Shalev (2004) suggests assistive technology as a potentially successful strategy. For example, interactive computer games have been suggested for building problem-solving skills (Lewis, 1998) and calculators can be used to compensate for poor fact retrieval (Ginsburg, 1997). The majority of the available literature on Mathematics interventions consists of common-sense strategies to build specific Mathematics skills.

Schools most often employ informal remediation programmes, such as withdrawal support. When a student receives withdrawal support, he or she attends regular class for the majority of the school day. However, for less than 50 per cent of the school day, he or she receives instruction outside of the classroom from a qualified special education teacher. The content, pace, instruction, etc. of the withdrawal support are determined by the special education teacher. As such, there will be considerable variability in the quality of withdrawal support across different schools; some children with a MD will receive a high quality intervention while others will be underserved. In addition, the quality of programming will likely go undetected as the special education teacher who designs the intervention is likely the same person responsible for evaluating the programme's effectiveness.

Formal programmes offer more consistency, objective evaluation, and are more likely to be theory-based. There are a handful of formal remediation programmes for children experiencing significant difficulties in Mathematics. For example, a school website listed several formal Mathematics intervention programmes for elementary school children such as KeyMath Teach and Practice and the Lightspan Math Programme (Mississippi Department of Education, 2006). KeyMath Teach and Practice serves as a supplemental or remedial programme to help develop children's mathematical concepts, operations, and applications (Pearson, 2006). Lightspan has a series of curriculum-based educational video games that target reading, Language Arts, and Mathematics for students in Kindergarten to grade eight (Business Editors, 2000).Unfortunately, a PsycInfo search of these programmes returned no results; the effectiveness of these programmes is unknown.

There are several formal intervention programmes targeting specific mathematical deficits that have been scientifically evaluated. *Great Leaps Math* is a supplemental Math programme designed to target fluency of Math facts (Jolivette, Lingo, Houchins, Barton-Arwood, & Shippen, 2006). This programme is built upon a 5-step strategy: (a) the child's teacher greets the student and tells him or her that the lesson is about to begin; (b) reviews previously acquired Math facts; (c) teaches new concepts; (d) administers a one-minute fluency probe; and (e) graphs correct and incorrect responses (Jolivette et al., 2006). Jolivette and colleagues (2006) found that this intervention programme helped two children with developmental delays and one child with ADHD to obtain greater mathematical fluency in addition. This study does have its limitations. First, the sample size is extremely small. The results, therefore, may not be generalizable to the intended population. Second, the study did not include a control group so the gains measured at post-testing cannot unequivocally be attributed to the *Great Leaps* programme. Last, the *Great Leaps* programme only targets single-digit Arithmetic problems and only requires oral responses to fluency probes. As such, there is limited room for growth in numerical operations and mathematical reasoning. This programme has not been systematically evaluated since Jolivette and colleagues' (2006) study according to a recent PsycInfo search.

The Cover, Copy, and Compare (CCC) method appears to hold the most empirical support for a Mathematic intervention programme targeting fluency of basic math facts. CCC is a self-managed intervention that involves 5 simple steps: "a) look at the Mathematics problem with the answer, b) cover the Mathematics problem with the answer, c) record the answer, d) uncover the Mathematics problem […] and e) compare the answer" (Codding, Shiyko, et al., 2007). Although the principles of this intervention may seem overly simplistic, CCC has been proven through empirical research to be an effective intervention (Codding, Eckert, et al., 2007). Significant increases in computational fluency were found when CCC was introduced as an intervention for second and third grade general education students (Codding, Shiyko, et al., 2007). CCC was also successful in improving Math fact accuracy and fluency for a 10-year-old student with moderate MR (Poncy et al., 2007). A CCC intervention also helped a third grade student with a MD increase her accuracy for multiplication problems (Stading, Williams, & McLaughlin, 1996); however, it is unclear whether her fluency improved as a result of the CCC intervention because this variable was not measured. In sum, there are various intervention programmes designed to enhance mathematical skills although the majority of these programmes lack empirical support.

In conclusion, MD is an area in need of further research and development from its' diagnostic criteria to its intervention. This research is important because MD is fairly prevalent and the disorder can have a serious, lifelong impact if left untreated.

## Kumon

### *Kumon Method of Learning*

Kumon is the most popular after school academic programme in the World with over 2.6 million students (Izumi, 2001). Kumon has Centres in 44 countries and has over 40,000 students in Canada alone (Izumi, 2001; L. Kaul, personal communication, June 26, 2008). Kumon's Mathematics and Reading programmes are used by children with varying abilities, from individuals with disabilities (Autism, MR, LDs, etc.), typically developing individuals to gifted students. The programmes are also used by children in varying grades (from preschool to postsecondary education) (KTRIE, 2002; McKenna et al., 2005).

The Kumon programmes were designed to supplement public education (McKenna et al., 2005). Most school boards take a spiral approach to Mathematics and Reading instruction; teachers introduce many skills and then spiral back at a later time to further develop each skill (KNA, 2008). Kumon's approach is linear (KNA, 2008). Kumon instructors introduce one topic at a time and will not advance a student until he or she demonstrates mastery of the lower level skill (KNA, 2008). Still, Kumon is meant to complement, not replace, general education received at school (KNA, 2008).

Toru Kumon developed the Kumon Method of Learning 50 years ago in Japan (KNA, 2008). Toru was a high-school Mathematics teacher as well as a Father (Izumi, 2001). Toru's son was struggling in Mathematics so Toru created a series of worksheets for his son to complete every day after school (Izumi, 2001). Toru's son started working with these worksheets in grade two and by the time he was in the sixth grade he was able to do Calculus (KNA, 2008).

The Kumon Method of Learning comprises seven components (Izumi, 2001):

- Individualized learning–Kumon students work and advance at a pace dictated by the individual's abilities and initiative (KNA, 2008). The instructor's primary goal is to pursue and help children attain the highest potential of each unique student (KNA, 2008).
- Independent learning–the level of difficulty progresses so gradually that students advance smoothly and independently. As a result, students develop self-motivation and self-reliance (KNA, 2008).
- Comfortable starting point–each student's starting point is determined by a Placement Test (KNA, 2008). The starting point is intentionally set low so that initial work is completed and mastered easily (KNA, 2008). This initial success fosters concentration, study habits, confidence, and proficiency with later steps (KNA, 2008).
- Curriculum–the study materials are organized so that skills develop incrementally and in a logical progression. The levels of the Math programme are divided into smaller steps. For each step, students complete a 10 page worksheet booklet (or set).

- Repeated practice–the Kumon Method ensures comprehension and retention by requiring daily practice (KNA, 2008). The amount of practice for each worksheet and each level is determined by the student's needs and performance.
- Mastery–Kumon assesses speed and accuracy to determine mastery**.** Students are said to have mastered a level when they achieve 100% accuracy on an assignment within a prescribed time period (KNA, 2008).
- Advanced level of study–"Kumon's goal is for all students to attain advanced student status as early as possible" (Izumi, 2001, p. 65). Students are expected to advance three levels or more from their entry level within one year of study (Niikura, 2006b).

When a student first enrolls in Kumon, he or she must write a placement test. Based on the student's performance on this test, the student is assigned a starting level; there are 26 levels in the Math programme and 24 levels in the Reading programme. Following placement, students are to complete daily worksheets that are completed in about 20 minutes. Twice a week, Kumon students must also go to a Kumon Centre where they are to complete their daily worksheets and pick up more worksheets for the rest of the week. Immediately after students complete their worksheets they should be graded by either a parent (if at home) or the instructor (if at the Centre). Students then correct any mistakes. After completing all the required worksheets for a given level with adequate accuracy and speed, students are given an Achievement Test to ensure that that level is in fact mastered. If the student is able to complete the Achievement Test with accuracy and speed, the student advances to the next level (KNA, 2008).

The Kumon Method of Learning, including individualized learning, independent learning, comfortable starting point, curriculum, repeated practice, mastery, and advanced level of study, has helped millions of students worldwide, according to anecdotal reports. Despite Kumon's popularity and personal testimonials, only one empirical study of its effectiveness surfaced in a recent PsycInfo search (McKenna et al., 2005). McKenna and colleagues (2005) measured the effectiveness of the Kumon Mathematics programme for economically disadvantaged children in grades two through five. Results showed that children receiving Kumon Mathematics instruction had greater improvements in Mathematics skills than a control group after 7 months of instruction and that those gains were maintained 2 years later (McKenna et al., 2005). However, this study has several limitations. First, children's achievement in Math is highly variable in grades one and two (APA, DSM-IV-TR, 2000) and thus including grade two students in the sample may have skewed the results. Second, 98% of the participants in this study were African American so the study's results may not be generalizable to the general population. Far more troublesome than the limitations of McKenna and colleagues' (2005) study is the fact that this is the *only* published empirical study investigating Kumon's effectiveness. A Math programme used by millions of students (gifted, average, learning disabled, and otherwise) should have more evidence-based support.

### *Kumon Mathematics Programme*

When Kumon was first established in 1958 it consisted only of a Math programme (Niikura, 2006a).Since then, the basic principles of the Kumon Math programme have remained the same but the levels and worksheets have been revised (KNA, 2008). The 26 levels of the Math programme are listed and described in Table 1.

### *Kumon and Gifted Students*

The Kumon programme is used extensively by gifted students for enrichment in Mathematics and Reading. Kumon meets many of the gifted child's educational needs because the programme is flexible, individualized, and sequential (McKenna et al., 2005). Gifted Math students, for example, can work through Mathematics materials without having to wait for his or her peers to catch up as they often do in general education classrooms (McKenna et al., 2005). The only study conducted on the Kumon programme that is published in a peer reviewed journal investigated the potential of the Kumon Math programme as an enrichment programme that would "attend to the potential giftedness of economically disadvantaged students [and] to give opportunities for Mathematics acceleration …" (McKenna et al., 2005, p. 223). Although McKenna and colleagues (2005) found that Kumon students demonstrated significantly greater gains in Mathematics skill levels than non-Kumon students, the study did not address giftedness in their selection or assessment of participants or in their results or discussion sections. Research on the effectiveness of the Kumon programme for gifted students is clearly in a very early stage.

### *Kumon and Students with Disabilities*

Kumon Mathematics and Reading programmes are primarily used by average-achieving and gifted children; however, there are a significant number of children with disabilities who make use of these programmes. KTRIE (2002) reported that there are approximately 4,700 students with disabilities enrolled in Kumon Centres in Japan. For a description of the types, numbers, and ratios of students with disabilities in Japanese Kumon Centres, see Table 2. These demographics should be interpreted with caution; the diagnostic criteria used to identify individuals with disabilities in the KTRIE (2002) study are unknown.

The Kumon philosophy is that every child has the possibility to grow and that every child should be guided to reach his or her full potential (KNA, 2008; KTRIE, 2002). KTRIE (2002) reported that "expand[ing] and develop[ing] one's strong point … is important when instructing normal children, however, it is more important when instructing disabled children" (p. 2). Kumon cannot be used to treat a child's disability. Rather, Kumon enhances and develops the child's capacity to learn (or strong point) which can make the child's disabilities seem less debilitating in comparison to his or her newfound abilities (KTRIE, 2002).

The Kumon Method also includes the tenet that students should learn at their own pace (KTRIE, 2002). This flexibility is important for students with disabilities as they often learn at a much slower pace than typically developing children (KTRIE, 2002). Even within groups of children with the same diagnosis there will be variability in the rate at which each child learns. In sum, although Kumon was originally designed for average-achieving children, the Kumon Method of Learning also addresses the special needs of children with disabilities.

The effectiveness of the Kumon programmes for children with disabilities has not been investigated empirically. However, there is an anecdotal report of the effectiveness of the Kumon Mathematics programme for an 18-year-old with Down syndrome published in a journal (Haslam, 2007). Haslam (2007), the teenager's Mother, suggested that it was the small steps and practice in the Kumon programme that allowed her son to succeed in Mathematics. Parent reports of the benefits of Kumon add valuable information and support for further use of the program. Nevertheless, objective, quantitative support is needed.

### *Kumon and Students with Learning Disabilities*

To the current author's best knowledge, there is no research or literature pertaining to the effectiveness of the Kumon Mathematics programme for children with any LDs, including MD. The predictions of the current research study, therefore, have been made based on knowledge of behavioural and cognitive learning theories, the characteristic deficits of MD, the treatment strategies and programmes that have previously been proven effective for children with MDs, and the tenets of the Kumon Method of Learning.

## Summary of Literature Review

Learning theory is used to guide understanding of the development of mathematical ability. Behavioural Psychology's learning hierarchy model and Cognitive Psychology's information-processing model are especially useful in studying typical and atypical development in Mathematics. These models are useful in interpreting whether the differences between children with Mathematics Disorder, average achieving children, and gifted children are quantitative, qualitative, or both. Another important area within the literature on Mathematics is the issue of whether differences exist between the genders in terms of mathematical ability.

The knowledge base of effective remedial and enrichment Mathematics programmes is insufficient. Kumon is an afterschool Math programme that claims to be effective for children of varying abilities. There is only one empirical study that has investigated

Kumon's effectiveness in a scientific manner. This is unacceptable considering that the program is used by hundreds of thousands of children across the World.

In sum, the current state of knowledge demands systematic, quantitative research of Mathematics programmes for children of varying abilities. Since Kumon seems to be a promising intervention (based on a review of the available literature and anecdotal evidence) and its' use is so ubiquitous, the current study will begin to add to the current literature by investigating the effectiveness of Kumon as a remedial and enrichment Mathematics programme. Chapter three will describe the research method that will be used to accomplish this goal.

# CHAPTER III: METHOD

Chapter three describes the research design that was used to test the hypotheses of this study. In addition to the research design, this section describes the participants, instruments and materials, procedure, and method of analysis of the current research project.

## Research Design

The optimal research design to determine the effectiveness of the Kumon Math programme is a pretest posttest control-group design with random assignment. Due to financial, time, and logistical constraints, random assignment of participants and inclusion of a control group was not possible. The current study, therefore, had a nonexperimental design. More specifically, this study had a causal comparative research design.

## Participants

Forty six Kumon centres in the Greater Toronto Area were approached to participate in this study. Eighteen Centre Instructors consented. All grades four, five, and six students who enrolled in the Math programme at any of these eighteen centres between January 1, 2009 and February 28, 2009 were invited to participate. Parental consent was obtained for thirty seven students. Fifteen of the largest Kumon centres in the Greater Vancouver Area were contacted in February, 2009 to increase the study’s sample size. Five participants were recruited from the Vancouver centres. The study, therefore, had an initial sample size of forty two students.

## Instruments and Materials

### *Canadian Achievement Test*, *Third Edition*

The Canadian Achievement Test, Third Edition (CAT-3) is a group administered, standardized test designed to assess achievement in Reading, Language, Spelling, Mathematics, and Writing (Anderson, 2005; Soares, 2005). The response format of the CAT-3 basic battery and supplemental tests is multiple-choice. Scores are available in the form of scaled scores, national percentiles, stanines, and grade equivalent scores (Anderson, 2005).

There are 10 test levels that correspond with specific school grades (grade one to postsecondary education) (Anderson, 2005). Within each test level there are three categories: basic battery, supplemental tests, and constructed response (see Table 3) (Soares, 2005). Although administration time varies with the level of the test, administration typically takes 85-115 minutes for the basic battery, 55-75 minutes for the supplemental tests, and 40-230 minutes for the constructed response sections (Spies & Plake, 2005). To administer only the Mathematics subtests, administration takes approximately 110 minutes [45 minutes for the Mathematics subtest (basic battery), 25 minutes for computation and numerical estimation (supplemental battery), and 40 minutes for the constructed response subtest].

The CAT-3 contains three Mathematics batteries. The basic battery (Mathematics subtest) assesses mathematical reasoning (Canadian Testing Centre, 2008). The computation and numerical estimation subtest assesses the ability to solve basic Math fact equations and to make mathematical estimations. The constructed response subtest assesses mathematical reasoning and the ability to communicate mathematical reasoning skills through written response (Canadian Testing Centre, 2008). The written responses are graded subjectively and thus the reliability of scoring comes into question with administration of this subtest. Thus, the constructed response battery was not included in the current study.

Test items were drawn from the Comprehensive Test of Basic Skills, Fourth Edition (CTBS-4) and revisions were made to reflect the curriculum taught in Canadian schools and the metric system (Michalko & Saklofske, 1996). New items created by Canadian teachers as well as Language Arts and Mathematics specialists were also added (Michalko & Saklofske, 1996; Soares, 2005).

The CAT has been used as a measure of academic achievement in several Canadian studies. Dahinten, Shapka, and Willms (2007) used the CAT-2 to assess the mathematical achievement of adolescents and their Mothers. The researchers found several correlations between child-Mother Mathematics scores on the CAT-2 and maternal variables, such as the Mother’s age and education (Dahinten et al., 2007). Unfortunately, these authors did not review the CAT-2, nor did they evaluate the appropriateness of the CAT-2 for their study.

The strongest support for use of the CAT-3 in the present study comes from the fact that the CAT has been used as an exemplar of psychometrics in the evaluation of several other achievement tests. Michalko and Saklofske (1996) administered the Wechsler Individual Achievement Test (WIAT) and the CAT-2 to a sample of Canadian children in order to determine the appropriateness of the WIAT with Canadian subjects. More recently, Saklofske, Caravan, and Schwartz (2000) used the CAT-2 as a comparison measure with the Wechsler Abbreviated Scale of Intelligence (WASI) for a sample of

Canadian children. Both groups of researchers chose the CAT-2 as a comparison measure because it has strong psychometric properties and Canadian norms (Michalko & Saklofske, 1996; Saklofske et al., 2000).

The CAT is considered one of the most psychometrically sound measures of academic achievement in Canada (Bachor & Summers, 1985). Systematic, national norming of the CAT-3 was conducted with a sample of over 44,000 students from 211 Canadian schools (Anderson, 2005). This sample was described by the authors as representative of the national population (Soares, 2005). Based on analysis of the internal structure of the CAT-3, the authors concluded that this test can be used to measure the achievement of specific composites (Language Arts, Mathematics, etc.) across grade levels (Soares, 2005). Factor analysis confirmed that the CAT-3 measures specific constructs (Language Arts and Mathematics) (Soares, 2005). Chronbach alpha internal consistency coefficients ranged from .72 to .92 among subtests (Anderson, 2005). Most coefficients (83%) exceeded .80 and few (9%) exceeded .90 (Anderson, 2005), which means that the items of each individual subtest measure the same skill.

The CAT is one of the most widely used measures of academic achievement in Canada (Bachor & Summers, 1985). The Mathematics subtests of this measure are appropriate for the current study because the sample of children to be tested are Canadian and are representative of the national population; “test scores and clinical interpretations may be misleading if the item content and normative data of a test do not reflect the learning, language, culture, and experiential background of the child or children assessed” (Michalko & Saklofske, 1996, p. 45). Based on format, purpose, development, recent literature, and technical qualities, the CAT-3 is an appropriate measure of the basic academic skills, such as Mathematics, for use with school-aged children (Anderson, 2005; Soares, 2005). As such, the CAT-3 was an appropriate measure of mathematical achievement for the current study.

*Canadian Cognitive Abilities Test*

The Canadian Cognitive Abilities Test (CCAT) is a group administered standardized test designed to assess students' reasoning and problem solving skills in verbal, quantitative, and spatial domains (Thorndike & Hagen, 1998). There are two editions, the Primary Battery and the Multilevel Edition (Thorndike & Hagen, 1998). The Primary Battery is used with children in Kindergarten through grade two and the Multilevel Edition is used with children in grades three through twelve (Thorndike & Hagen, 1998).

Both editions yield Verbal, Quantitative, Non-Verbal and Composite scores (Thorndike & Hagen, 1998). Within the Multilevel Edition there are eight levels (A-H) (Thorndike & Hagen, 1998). The Verbal battery assesses "verbal inductive and deductive reasoning skills as well as flexibility, fluency, and adaptability in working with verbal materials and solving verbal problems" (Thorndike & Hagen, 1998, p. 8). The Quantitative battery assesses similar skills but in relation to quantitative symbols and concepts (Thorndike & Hagen, 1998). The Non-Verbal battery taps the student's general inductive reasoning skills and the ability to use and adapt cognitive strategies (Thorndike & Hagen, 1998). The Composite score incorporates the Verbal, Quantitative, and Non-Verbal scores to provide an indication of the student's overall level of cognitive abilities, or "g" (general intelligence) (Thorndike & Hagen, 1998). Within each battery there are three subtests. For a list of these subtests, see Table 4.

All questions on the CCAT are in multiple-choice format (Thorndike & Hagen, 1998). Each question has five possible answers with the exception of the Quantitative Relations subtest which contains questions with only three answer choices (Thorndike & Hagen, 1998). There are specific time limits for each subtest which make for a total administration time of about two hours (Hattie, 1995).

The CCAT yields four types of scores: Universal Scale Scores (USS), Standard Age Scores (SAS), percentile ranks, and stanines (Thorndike & Hagen, 1998). The USS are standardized scores used for comparing an individual's scores across the various CCAT levels (if administered in successive years) (Anderson, 1995). The SAS are standard scores that allow comparison of individuals to others in the same age group (Thorndike & Hagen, 1998). These scores are normalized with a mean of 100 and a standard deviation of sixteen (Thorndike & Hagen, 1998). Percentile ranks and stanines are available based on both grade and age (Anderson, 1995).

The items within the CCAT are based on those of the Cognitive Abilities Test (CogAT) but contain some modifications such as changes in spellings and systems of measurement to make the test more appropriate for Canadian students (Hattie, 1995). The only major adaptation from the CogAT was the development of Canadian norms (Anderson, 1995).

The Canadian norms for the CCAT are psychometrically sound. Approximately 30,000 students nation-wide were used to standardize and create these norms (Thorndike & Hagen, 1998). Reliability of the CCAT is also said to be respectable. The internal consistency estimates for the three subtests at all levels are between .81 and .94 (Anderson, 1995). Researchers also measured the stability of scores over time and found that the test does generate consistent results (Anderson, 1995). Both Anderson (1995) and Hattie (1995), reviewers of the CCAT, state that validity is inadequately addressed in the CCAT technical manual. However, there is reported evidence of the CCAT's criterion-related validity (Anderson, 1995). Correlations between scores on the CCAT and the Canadian Test of Basic Skills (CTBS) are between .40 and .80, providing moderately positive evidence of convergent validity (Anderson, 1995). CCAT Verbal scores were also correlated with the Henmon-Nelson Ability Test (between .78 and .84), a measure of general ability (Anderson, 1995). In sum, the CCAT has desirable psychometric properties as a measure of cognitive ability.

The CCAT has been used is past research as a measure of children's cognitive ability. Johnson, Im-Bolter, and Pascual-Leone (2003) selected the sample for their study by reviewing students' scores on the CCAT as it had already been administered by their classroom teachers. Johnson and colleagues (2003) categorized children as *gifted* or *mainstream* based on their Wechsler Intelligence Scale for Children, Third Edition (WISC-III) or CCAT scores. Unfortunately, Johnson and colleagues (2003) did not define the cutoff point for inclusion in the *gifted* group. Schneider, Clegg, Byrne, Ledingham, and Crombie (1989) recruited gifted children for their study from schools that routinely administer group IQ tests such as the CCAT and the Henmon-Nelson Test of Mental Ability. The researchers reviewed records of students' past scores and selected children scoring 129 or higher on the Henmon-Nelson or above the 97$^{th}$ percentile on the Verbal section of the CCAT (Schneider et al., 1989). The Verbal section scores of the CCAT were used (as opposed to the Quantitative, or Non-Verbal scores) "because of its high correlation with the Henmon-Nelson" (Schneider et al., 1989, p. 50). The CCAT has also been administered within the study, as opposed to using CCAT scores obtained prior to the study. Westbury (1994) administered the CCAT to two groups of students, those who were held back one school grade and those who advanced typically. The CCAT served as a pre- and post-test measure of cognitive ability and allowed Westbury (1994) to suggest that grade retention did not improve children’s abilities in three years time. In sum, the CCAT has been used in past research as a measure of students’ general cognitive functioning. Past researchers as well as CCAT reviewers agree that the CCAT is a useful instrument for determining students’ cognitive ability and is a valuable research tool (Anderson, 1995; Hattie, 1995).

## Procedure

### *Recruitment of Participants*

After receiving permission from the University of Lethbridge Ethics Committee and the individual Kumon Centre Instructors, Instructors were asked to give consent forms to the parents of all eligible students. Pretesting was scheduled after parents returned signed consent forms.

### *Preparation*

Protocols were prepared prior to the testing date. The participant’s grade determined which CAT-3 and CCAT levels were administered and thus which protocols were prepared (see Table 5). To protect participant’s anonymity, each student was assigned a six digit code. The first two digits represented the child’s centre, the third digit represented the student’s grade, and the last three digits indicated the order in which the participant was tested, relative to the total number of participants at his or her centre (i.e., 001, 002, etc.). This code was the only identifying information on the protocols.

*Pretesting*

Pretesting occurred between January 14th, 2009 and April 1, 2009. The vast majority of pretesting in the Greater Toronto Area was completed by the lead researcher. When multiple pretesting sessions were scheduled on the same day, Kumon Field Staff helped in test administration. The lead researcher trained the Field Staff on administration procedures prior to commencing pretesting. All testing in the Greater Vancouver Area was completed by Kumon Field Staff members.

Testing took place at the student's Kumon centre during his or her regularly scheduled Kumon class time. The CAT-3 subtests were written on the first day of testing. The Mathematics subtest was written first, followed by the computation and numerical estimation subtest. The CCAT Verbal subtests were written on the second day of testing in the following order: Verbal Classification, Sentence Completion, and then Verbal Analogies. For a summary of the order of testing, see Table 6. The average length of time between the first and second day of testing was 14 days ($SD$ = 11).

There was some variability in testing conditions due to the fact that there is variability in the amount of space available at each Kumon centre. Most Instructors were able to designate a table to testing in the corner of the classroom. In a few cases, the testing table was shared with Kumon students who were not participating in the study. The examiner and examinee(s) were the only people sitting at this table when possible. When there were multiple participants completing testing at the same table, seating was arranged so that students could not see each other's answers.

Noise levels also varied across testing. For the most part, noise was kept to a low whisper but sometimes the noise level escalated to a medium volume (people talking near the testing table using their regular speaking voices).

Before administering any assessment measures, the examiner explained the purpose of the study by saying, “I’m doing a study to see how Kumon helps grade 4, 5, and 6 students to learn Math. Here is a form that explains this study [examiner presented assent form]. Would you like to read this form yourself or would you like me to read it to you? [read form if asked]. Do you have any questions?” After all questions were answered, the examiner asked the student to print his or her name on the assent form if he or she was willing to participate.

The examiner then briefly explained the purpose of administering the CAT-3. The examiner said, “This test is called the Canadian Achievement Test and it will be used to measure your Math skills.” The examiner then read and followed the directions for test administration as outlined in the CAT-3 examiner’s manual for both CAT-3 subtests. Upon completing the test, the examiner thanked the participant for his or her time and effort and reminded him or her of the upcoming CCAT testing.

On day two of pretesting, the examiner briefly reviewed the purpose of the study with the participant(s) and obtained verbal assent to proceed. The examiner then introduced the CCAT by saying, “Today you are going to take a test called the Canadian Cognitive Abilities Test. There are many different kinds of problems on this test. We want to find out how well you can solve them.” Directions for CCAT administration were then followed as outlined in the test manual. Once the participants finished writing the three CCAT Verbal subtests, the examiner informed them that they have completed all the requirements for the first part of the research study. The examiner explained that he or she would be back in six months to complete the second part of the study. Last, students were thanked for their time and effort.

### *Posttesting*

Posttesting began on July 14, 2009 (six months after pretesting began) and ended on October 22, 2009. Posttesting and pretesting conditions were very similar. Testing occurred at participants' Kumon centres during their regularly scheduled class time, participants were secluded in the corner of the Kumon classroom while they completed testing when possible, and noise levels fluctuated between low and medium.

The CAT-3 subtests administered at posttesting were the same (subtests and level) as those administered at pretesting. Participants were not re-administered the CCAT. See Table 7 for a list of the posttesting measures administered. Verbal assent was obtained prior to beginning testing. The examiner then followed the CAT-3 administration directions. All posttesting in the Greater Toronto Area was completed by the lead researcher. All posttesting in the Greater Vancouver Area was completed by Kumon Field Staff.

Last, the examiner determined how many worksheets each participant completed in the Kumon programme between pretesting and posttesting. A 'successfully completed worksheet' was defined as a single worksheet completed with adequate speed and accuracy as determined by the child's Kumon Instructor. In other words, a successfully completed worksheet was not re-assigned by the Instructor.

## Method of Analysis

The purpose of the study is to investigate the effectiveness of the Kumon Mathematics programme for children of varying abilities. Specifically, the research questions guiding this study are:

1. Is there a difference between male and female participants in terms of their pretest scores, posttest measures, or the gains they made on the CAT-3?
2. Is there a difference in the magnitude of gains made in Math skills among children of different achievement groups (below average, average, above average, and "gifted")?
3. Is there a difference in the magnitude of gains made on a measure of computation skills versus mathematical reasoning skills after six months of participating in the Kumon Math programme?
4. What is the relationship between children's participation in the Kumon Mathematics programme and their achievement in Mathematics?

### *Hypothesis One: Gender Differences in Mathematical Ability*

The null hypothesis is that there is no difference between males and females in terms of mathematical ability as measured by the CAT-3. More specifically, there is no significant difference between genders in terms of pretest scores (CAT-3 scaled scores or CCAT standard age scores), number of Kumon worksheets completed, posttest scaled scores on the CAT-3 subtests or gains made on the CAT-3 subtests between pre- and posttesting (difference in scaled scores). An independent samples t-test was used to test the null hypothesis.

## *Hypothesis Two: Children of Varying Mathematical Ability*

### *Categorization of Participants*

In order to test the second null hypothesis, participants had to be categorized into achievement groups. The participants were first categorized based on their pretest CCAT standard age score. Children who scored between 69 and 131 were categorized as having average cognitive abilities. Children who had a score greater than 131 were categorized in the above average cognitive abilities group.

The average cognitive abilities group was then further divided based on their percentile ranking on the CAT-3 Total Math subtest. There were four achievement groups: below average, average, above average, and “gifted” (see Table 8 for the corresponding percentile ranking ranges).

The below average group includes students who may have a MD. However, participants in this group were labeled ‘below average” because children cannot be formally identified as having a MD based on the data collected in the current study. Similarly, students at the upper extreme of achievement at pretesting were labeled as “gifted”, as they cannot definitively be identified as gifted using the information gathered in the current study.

### *Method of Analysis*

The null hypothesis is that there is no difference among achievement groups in the gains made on the CAT-3 Mathematics subtest or the CAT-3 computation and numerical estimation subtest. The directional hypothesis is that “gifted” students have the greatest gains, followed by the above average students, followed by the average students, who are followed by the below average students. A one-way Analysis of Variance (ANOVA) was used to test the null hypothesis.

A limitation of using the CAT-3 as the measure of mathematical ability is that it does not have a high enough ceiling to capture gains made by gifted students. As such, achievement in Mathematics was compared among achievement groups using a second measure, number of Kumon worksheets completed. The null hypothesis is that there is no difference in the number of steps completed among the four achievement groups. The directional hypothesis is that “gifted” students complete the greatest number of worksheets, followed by above average students, followed by average students, followed by below average students. A one-way ANOVA was used to test the null hypothesis.

*Hypothesis Three: Kumon Mathematics Programme*

The null hypothesis is that there is no difference in the gains made in computation skills versus gains made in mathematical reasoning skills. The directional hypothesis is that gains made on the CAT-3 computation and numerical subtest will be significantly greater than gains made on the CAT-3 Mathematics subtest (a measure of mathematical reasoning). A paired samples t-test was used to test the null hypothesis.

*Hypothesis Four: The Relationship between Kumon and Achievement in Mathematics*

The null hypothesis is that gains made in Mathematics skills cannot be predicted based on a measure of participation in the Kumon Math programme (number of worksheets completed). The directional hypothesis is that gains made in Mathematics can be predicted based on the number of Kumon Math worksheets a child completes. Since Kumon worksheets focus on computation exercises, it was predicted that the results from the current study will only support the directional hypothesis for gains made in computation skills.

The second null hypothesis is that the relationship between gains made in computation skills cannot be predicted based on pretest achievement scores. The directional hypothesis is that there is a positive relationship between gains made on the computation and numerical estimation and pretest scores on the same subtest. Furthermore, a model that examines the predictive ability of the two variables (number of Kumon worksheets completed and pretest computation scores) together will explain more variance in gains made in computation skills than models that examine the relationships between the predictor and criterion variables separately.

Multiple regression analyses were used to test the null hypotheses. Separate analyses were required, one for each criterion variable (gains made on the Mathematics subtest and gains made on the computation and numerical estimation subtest). The predictor variables in both analyses were the number of Kumon worksheets completed and pretest score on the respective CAT-3 subtest.

## Summary of Method

The current study had a nonexperimental research design. Forty two children met the eligibility criteria for participation in the study and were given parental consent to participate. Students completed the CAT-3 Mathematics subtest, the CAT-3 computation and numerical estimation subtest, and the CCAT Verbal subtests at the time of pretesting. Six months later, at posttesting, children were re-administered the CAT-3 subtests. The number of Kumon Math worksheets completed between pre- and posttesting was also recorded for each participant. T-tests, ANOVAs and multiple regression were used to answer the study’s research questions. Chapter four presents the results from these statistical analyses.

# CHAPTER IV: RESULTS

Chapter four presents the research findings of the current study. First, the descriptive statistics for the pre- and posttest samples are presented. Next, the results of the statistical analyses are presented for each of the study's four hypotheses; each hypothesis is restated and followed by all relevant statistical findings. PASW Statistics GradPack version 18.0 computer software was used for all statistical analyses.

## Sample at Pretesting

The sample at the time of pretesting consisted of 17 males and 25 females. Nine of the participants were in grade 4, 15 were in grade 5, and 18 were in grade 6. Participants' ages ranged from nine years, two months to 12 years, one month. The majority of participants (30 students) in the sample were enrolled in the Kumon Math programme, while twelve students were enrolled in both the Math and Reading programmes.

All but one participant, who was in the gifted range, fell in the average range on the measure of cognitive ability. On the measure of mathematical achievement, 12 participants were in the below average range, 23 in the average range, six in the above average range, and one participant was in the "gifted" range (this is the same participant who was in the above average range on the CCAT). See Table 9 for a summary of the descriptive statistics for the pretesting sample.

## Sample at Posttesting

There was an extremely high attrition rate in this study. Over a six-month period, 50% of participants dropped out of Kumon. Instructors reported various reasons for participants leaving Kumon, such as parents having to refocus their financial and/or time commitments. Other students left Kumon in June for summer break and never returned in

September. The two groups (students who remained in Kumon versus students who left Kumon) were compared based on their pretest variables to determine if there are any group differences.

*Analysis for Group Differences*

Independent samples t-tests revealed that there are no significant differences between the two groups (students who stayed in Kumon versus students who quit Kumon mid-study) in terms of pretest CAT-3 scores. More specifically, pretest mathematical reasoning scores are not significantly different between children who stayed in the programme ($M$ = 510, $SD$ = 86) and children who left the programme ($M$ = 487 $SD$ = 57), $t(40) = 1.02$, $p = .32$. Pretest numerical operation scores are not significantly different either [children who stayed in the programme ($M$ = 492, $SD$ = 82); children who left the programme ($M$ = 479, $SD$ = 51); $t(40) = .58$, $p = .56$]. Last, the difference between groups when total Mathematics scores are compared is not significant [children who stayed in the programme ($M$ = 501, $SD$ = 77); children who left the programme ($M$ = 483, $SD$ = 48); $t(40) = .88$, $p = .38$]. There are no significant group differences when CCAT scores are compared either [children who stayed in the programme ($M$ = 104, $SD$ = 17); children who left the programme ($M$ = 96, $SD$ = 13); $t(40) = 1.52$, $p = .14$].

Chi-square tests were used to determine if the children who stayed in the programme and children who left the programme differed significantly in terms of grade, sex, and/or Kumon programme (enrollment in Math only versus enrollment in both Math and Reading). Results indicate that the two groups do not differ from each other significantly on any of these variables (see Table 10).

### *Sample at Posttesting*

The final sample consisted of 11 males and 11 females. Five of the participants were in grade 4, eight were in grade 5, and nine were in grade 6. The youngest participant in this sample was nine years, two months old at the time of pretesting, and the oldest student was 12 years old, one month at pretesting. The majority of participants (18 students) were enrolled only the Kumon Math programme while four students were enrolled in both the Math and Reading programmes. There were six participants in the below average achievement group, 10 in the average group, five in the above average group, and there was one participant in the “gifted” group. See Table 11 for a summary of the descriptive statistics for the final sample.

## Statistical Analyses

### *Hypothesis One: Gender Differences in Mathematical Ability*

An independent samples t-test was used to test the null hypothesis that there are no significant differences between males and females in terms of mathematical or cognitive ability. The results suggest that there are no significant differences between the two groups for any of the variables investigated (see Table 12) and that the null hypothesis should be accepted.

### *Hypothesis Two: Children of Varying Mathematical Ability*

The “gifted” group was excluded from analysis for two reasons. First, the CAT-3 does not have a high enough ceiling to accurately capture gains made by students who had a high number of correct responses at pretesting. Second, there was only one case in this group.

The null hypothesis is that there are no significant differences between achievement groups in terms of gains made in Mathematics over a six-month period. ANOVA results suggest that the CAT-3 scores are not significantly different among achievement groups (Table 13). However, the difference among achievement groups on the CAT-3 total Math subtest approaches significance.

Descriptive statistics show a notable trend in gains made on the computation and numerical estimation subtest (see Table 14). Below average students tend to make greater gains in this area than above average students. However, it is also notable that there is significant variance (*SD* = 78) in above average students' gains made in computation and numerical estimation. The null hypothesis is accepted.

*Kumon Worksheets Completed*

A one-way ANOVA was used to test the hypothesis that there would be differences among achievement groups in the number of Kumon levels advanced over a six-month period. The null hypothesis is that there is no difference in the number of worksheets completed among achievement groups.

Descriptive statistics suggest that there is a minor trend toward below average students completing the fewest and above average students completing the greatest number of worksheets (see Table 15). However, the ANOVA results indicate that the number of levels advanced in the Kumon programme is not significantly different among achievement groups, $F(2, 18) = .181$, $p > .05$. The null hypothesis is accepted.

*Hypothesis Three: Kumon Mathematics Programme*

A paired samples t-test was used to determine if the gains made by all participants in mathematical reasoning are significantly different from the gains students made in their computation and numerical estimation skills. The null hypothesis is that there is no significant difference between the gains made in the two domains of Mathematics.

Descriptive statistics suggest that there is a trend towards students demonstrating greater gains in computation skills than mathematical reasoning skills (see Table 16). These results also show that there is a lot of variance in the results (wide range of gains made in both subtests and large *SD*s). The paired samples t-test results indicate that the difference between gains made in the two domains of Mathematics is not significant, $t(20) = -1.660$, $p > .05$. Thus, the null hypothesis is accepted.

*Hypothesis Four: The Relationship between Kumon and Achievement in Mathematics*

Multiple regression was used to investigate how the number of Kumon levels advanced, pretest CAT-3 scores, and pretest CCAT scores are related to gains made in Math skills. The null hypothesis is that none of these variables significantly predict gains made in Mathematics. Grade, age, and sex were not investigated as predictor variables due to the small sample size.

Analyses were conducted to determine the strength of the relationships between the predictor variables and the criterion variables. Three multiple regression analyses were completed, one for each criterion variable: gains made in mathematical reasoning, gains made in computation and numerical estimation, and gains made in total Mathematics (see Tables 17-19).

*Mathematical Reasoning*

Results indicate that the neither the number of Kumon levels advanced, $t(17) = -.170$, $p > .05$, CAT-3 pretest scores, $t(17) = .147$, $p > .05$, nor scores on the measure of cognitive ability, $t(17) = -1.425$, $p > .05$, are reliable predictors of gains made in mathematical reasoning. Even together, these predictor variables to not explain much of the variance in the predictor variable (adjusted R square = .057). The null hypothesis is accepted.

*Computation and Numerical Estimation*

The number of Kumon levels advanced, $t(17) = 2.077$, $p = .053$, is not a statistically significant predictor of gains made in computation and numerical estimation skills; however, it is extremely close to reaching significance. CAT-3 pretest scores on the computation and numerical estimation subtest are significant predictors of gains made in the same subtest, $t(17) = -3.206$, $p < .01$. Cognitive ability scores, in contrast, are not significant predictors of gains made in computation and numerical estimation, $t(17) = .727$, $p > .05$. This model explains about one-third of the variance in gains made in computation and numerical estimation (adjusted R square = .340). The null hypothesis is rejected in this case.

*Total Mathematics*

Results suggest that the number of Kumon levels advanced, $t(17) = 1.288$, $p > .05$, CAT-3 total Math pretest scores, $t(17) = -.907$, $p > .05$, and CCAT scores, $t(17) = -.506$, $p > .05$, are not significant predictors of gains made in total Mathematics. Thus, the alternate hypothesis is accepted. Overall, this model does not explain very much variance in the gains made in total Math (adjusted R square = .087).

## Summary of Results

The sample at the time of posttesting was considerably smaller than the original sample. Statistical analyses were conducted to determine if there were any differences between children who stayed in the Kumon Math programme and children who left. No differences were found in any of the variables investigated.

The results of the current study suggest that there is no significant difference between males and females (in grades 4-6) in terms of mathematical or cognitive ability. This is true before and after participating in the Kumon Math programme for six months.

Statistical analysis also suggested that there is no significant difference in gains made in Mathematics after participating in the Kumon Math programme among children of varying abilities. However, there is a trend that suggests that children who fall in the below average group make greater gains in computation skills than children who fall in the above average group.

There was no significant difference in the number of Kumon worksheets completed among children of varying abilities. There was a weak trend, though, that suggested that below average students may complete the fewest number of sheets while above average children complete the greatest number.

The current results suggest that there is no significant difference in gains made in computation versus mathematical reasoning skills after six months of participation in the Kumon Math programme. The mean for gains made in computation skills is greater than the mean gains made in mathematical reasoning, but the difference is not statistically significant.

Last, statistical analysis was used to investigate the relationship between gains made in Mathematics and several predictor variables (number of Kumon worksheets completed, pretest Math scores, and cognitive ability scores). None of the variables investigated are significant predictors of gains made in mathematical reasoning or total

Mathematics. In contrast, the number of Kumon worksheets completed approaches significance as a predictor of gains made in computation skills. There is a significant, negative relationship between pretest computation scores and gains made in computation skills. Cognitive ability scores do not significantly predict gains made in computation skills.

These results are interpreted in the chapter that follows. Limitations, such as the small final sample size, will not only be discussed in chapter five, but also incorporated into the interpretation of chapter four’s results.

# CHAPTER V: DISCUSSION

Chapter five discusses the results of the current study. First is a summary of the study's research problem and method, followed by an interpretation of each important result noted in chapter four, a discussion on the limitations of the current study, and implications of the study's findings for current practice. Last is a section containing suggestions for future research.

## Summary of Research Problem and Method

The research question of the current study was formulated after a review of the current literature and discovery of a gap in the scientific knowledge base of effective Mathematics remedial and enrichment programmes for children. Kumon was selected as the programme for investigation because it is widely used, yet has hardly been evaluated scientifically.

An experimental design is needed to determine if Kumon's Math programme is effective. The resources needed to conduct such an experiment were not available, thus the research problem became: what is the relationship between participation in the Kumon Mathematics programme and achievement in Mathematics for children of varying ability? Children in grades 4-6 were administered a measure of mathematical achievement and cognitive ability when they first started the Kumon Math programme. Six months later, children were re-administered the same measure of mathematical ability and information was gathered on the number of Kumon worksheets the child completed between pre- and posttesting. Statistical analyses were used to answer the research question and to test the study's null hypotheses.

## Interpretation of Results

### *Sample*

The results from statistical analyses suggest that the group of children who quit Kumon over the course of the study and the group of children who remained in Kumon are not different from each other in terms of pretest scores in Mathematics, cognitive ability, grade, sex, or whether they were enrolled in Kumon Math or Kumon Math and Reading. The final sample size was small (22 participants), but since the final sample is not distinguishable from the considerably larger initial sample (43 participants) on any major variables investigated, one can have greater confidence that the final sample is representative of the target population.

Still, it is possible that the final sample is different from the original sample on some unmeasured variable(s). For example, parental income and education likely influence the length of time that a child spends in Kumon. These variables could also influence a child's achievement in Mathematics.

### *Hypothesis One: Gender Differences in Mathematical Ability*

The first null hypothesis was that there would be no significant differences in mathematical ability or achievement between male and female participants. The null hypothesis was accepted; there are no significant differences between the genders in terms of CAT-3 scores (pretest, posttest, or gains made in mathematical reasoning, computation and numerical estimation, or total Math), CCAT Standard Age scores, or number of Kumon worksheets completed.

These results are consistent with past findings that gender differences in mathematical ability are small and insignificant (Bjorklund, 2005; Georgiou et al., 2007; Hedges &

Nowell, 1995; Hyde et al., 1990; Leahey & Guo, 2001; Maccoby & Jacklin, 1974). When mathematical ability is separated into subskills, some researchers have found that females have stronger computation skills than males (Hyde et al., 1990). These results were not replicated in the current study; there was no difference in pretest, posttest, or gains made in computation skills between males and females. However, it is important to keep in mind that the sub-sample sizes were quite small (11 males and 11 females). Thus, the results of the current study may not reflect the true abilities of grade 4-6 males and/or females.

Researchers have also found that there are differences in the number of males and females at the upper extremes of mathematical ability. In the current study both of the participants who scored above the 95$^{th}$ percentile in total Mathematics were males. Again, generalizations to the target population must be made with caution due to the small sample and sub-sample sizes. However, the trend found in the current study is consistent with past research.

### *Hypothesis Two: Children of Varying Mathematical Ability*

It was predicted that students who demonstrate above average and "gifted" level achievement in Mathematics upon entering the Kumon Math programme will accelerate at a faster pace than average and below average children because of the qualitative and/or quantitative differences in their abilities. More specifically, it was predicted that because gifted children are quicker at solving problems, have more elaborate procedural knowledge, are more flexible in using strategies, and are more sophisticated in meta-cognition (Threlfall & Hargreaves, 2008), they will gain more from the Kumon programme and accelerate through its' levels at a faster rate.

In a related vein, it was predicted that children in the below average group will advance at the slowest rate in the Kumon programme because of their qualitative and/or quantitative differences in their ability to learn Mathematics. Difficulties at each stage in the learning hierarchy and psychological processing deficits were expected to slow their rate of growth in the Kumon Math programme.

Surprisingly, the null hypothesis was accepted. Statistical analysis reveals that there is no significant difference in gains made in Mathematics among achievement groups (below average, average, and above average). These results may suggest that rate of growth Mathematics while participating in the Kumon Math programme is not influenced by the child's mathematical ability at the time of entering the programme; children of varying abilities progress at the same rate in Kumon.

However, trends within the results and past research suggest alternative explanations. First, there was a strong trend in the results toward below average students demonstrating greater gains in computation skills than above average students. This trend may reflect a ceiling effect on the CAT-3. Above average children may not be able to demonstrate the magnitude of their growth in Mathematics via the CAT-3. All children are administered the same number of questions on the CAT-3 subtests, unlike on individually administered tests, where the number of questions administered is adjusted based on the individual's performance. As such, children who had a higher percentage of correct responses at pretesting had a smaller window to demonstrate growth than children who initially scored in the lower ranges.

Still, the difference among achievement groups was not statistically significant, contrary to what was predicted based on a literature review. Perhaps the effect of pre-intervention ability is masked by the small sample size and even smaller sub-group (below average, average, above average) sample sizes. As a result, the mean scores may not be representative of the target population.

The criteria used to classify children into achievement groups may not be appropriate. Percentile rankings on the Total Mathematics subtest were used to categorize participants. This subtest's percentile rankings are determined based on the individual's scores on the Mathematics and the Computation and Numerical Estimation subtests, and is not a subtest in itself. If a large discrepancy exists between an individual's Math and computation scores, his or her total Mathematics percentile ranking may not accurately reflect his or her overall mathematical ability. In this case, total Math percentile ranking may not be an appropriate measure to categorize the participant into an achievement group.

In addition, the percentile ranking ranges outlined to categorize students may not accurately define achievement groups. The literature review conducted in the current study revealed extremely inconsistent cutoff criteria among researchers and practitioners for identifying below average and gifted children. The cutoff ranking for Mathematics Disorder varies between the $8^{th}$ and $48^{th}$ percentile (Swanson & Jerman, 2006) and criteria for identifying gifted children fluctuates by approximately 10 percentile rankings. The ranges selected for use in the current study were made based on the most frequently used ranges found in the literature search, but may not capture the true ranges that define achievement groups.

Last, statistical analysis may have revealed insignificant differences among achievement groups because the relationship between gains made in Mathematics and pretest ability (below average, average, or above average) is influence by additional variables. Multiple regression was used to analyze the relationship between gains made in Mathematics and several predictor variables (see Hypothesis four: The relationship between Kumon and achievement in Mathematics).

*Kumon Worksheets Completed*

Based on the same literature review on the qualitative and quantitative differences between below average, average, above average, and gifted children, it was predicted that there would be a significant difference in the number of worksheets completed in the Kumon Math programme between achievement groups. Above average students were expected to complete the greatest number of worksheets and below average students were expected to complete the fewest.

Statistical analysis led to acceptance of the null hypothesis. The number of worksheets completed by students was not significantly different across achievement groups (below average, average, and above average). This could suggest that the rate of advancement in Kumon is independent of the child's abilities at the time of entering the programme.

The accuracy of interpretations based on this study's results is tempered by the small sample size of the current study. There are only a few cases in each achievement group, reducing the likelihood that the obtained results reflect the results of the general population.

The results may be contrary to what was predicted because the classification criteria used to define achievement groups are invalid. Another possible explanation for the unexpected findings is that there may be extraneous variables influencing children's rate of worksheet completion. For example, number of worksheets completed is likely influenced by student motivation and parent encouragement.

Another factor to consider is that there is subjectivity in determining mastery of a worksheet. Kumon has set objective criteria for evaluating mastery of worksheets, but instructors are permitted to use subjective judgment in determining whether or not a child should advance to the next worksheet. In other words, instructor variables influence the number of worksheets completed. If there is a lot of variability in 'mastery criteria' between instructors, achievement group differences may have been obscured. Therefore, the results in this section may best be interpreted as evidence that unmeasured variables (such as student motivation or instructor differences) are stronger predictors of the number of Kumon worksheets completed than pretest achievement scores or that certain variables must be considered in tandem to explain the number of worksheets completed by a Kumon student.

### *Hypothesis Three: Kumon Mathematics Programme*

The vast majority of Kumon's Math worksheets focus on computation skills and only a few questions specifically exercise mathematical reasoning skills. As such, it was predicted that Kumon would have a differential effect on Mathematics skills; children will make significantly greater gains in computation skills than mathematical reasoning skills. The results of the current study show the difference is not significant. One interpretation of these results is that Kumon has an equal effect on both types of Math skill.

The computation skills that are exercised in Kumon worksheets may be generalizing to mathematical reasoning skills. Although the computation skills needed to succeed on the Math reasoning subtest are more basic than those required on the numerical computation subtest (for example, single-digit addition versus multiple-digit addition), computation skills are needed nonetheless.

Kumon students begin the programme at a low start point (a level that would correspond to several grade levels below their current scholastic grade level). As a result, students work on lower level computation skills for the first few months after enrolling in Kumon. The basic computation skills needed to excel on the Math reasoning subtest are thus practiced repeatedly and mastered in the first few months of the child's participation in the Kumon Math programme. The more advanced computation skills (such as those needed to succeed on the computation subtest) are targeted only after lower level skills are mastered. In summary, gains made in the two areas may be comparable because in a six month period, the child spends more time practicing lower level skills than higher level skills and thus is at different levels in the learning hierarchy. Basic computation skills are practiced repeatedly in the first six months and thus may have reached the mastery level. As such, students can generalize and adapt their skills to solve Math reasoning problems. Higher level computation skills are practiced less, and thus may only be in the fluency or even acquisition stage at the time of posttesting.

Alternatively, the comparable gains in both domains of Mathematics could reflect development in an overarching skill. Kumon states that its' programmes develop not only Math and Reading, but concentration, study habits, and self- confidence (Izumi, 2001; KNA, 2008). Information-processing theory suggests that psychological processes, such as attention and concentration, support and enhance the transfer of information to long-term memory (Hetherington et al., 2005). Therefore, participation in the Kumon Math programme may lead to equal growth in computation and mathematical reasoning skills as a result of strengthened psychological processing skills.

McKenna and colleagues' (2005) also found that the Kumon group had significantly higher test scores than the non-Kumon group on measures of both mathematical concepts (computation) and problem solving (mathematical reasoning). The following year, retention effects were measured and again the difference between experimental and control groups was significant for both mathematical procedures (computation) and problem solving (mathematical reasoning) (McKenna et al., 2005). Therefore, Kumon may have a significant, positive, and comparable effect on both Math reasoning and computation skills.

Descriptive statistics of gains made in computation and Math reasoning skills also yielded notable results. The range of gains made shows that some students scored lower at posttesting than pretesting. Four participants demonstrated deficits in mathematical reasoning and three students demonstrated deficits in computation and numerical estimation.

One interpretation of these results is that some students regress in their mathematical achievement after participating in Kumon. It isn't likely that students would actually lose Math skills, as long as practice is relatively consistent (as is mandated by the Kumon programme). Instead, it could be that students' *performance* on a measure of mathematical ability is compromised.

Speed is a main goal in the Kumon programme and this is emphasized by instructors and in Kumon's evaluation of student progress. Perhaps some participants sacrificed accuracy for speed at posttesting. At pretesting, 21% of students were not able to complete the subtests (answer every question in the allotted time). At posttesting, only one participant did not finish before the time expired (5% of the final sample). This suggests that students are working more quickly on Math problems after participating in the programme for six months.

However, it is not likely that Kumon students learned to value speed over accuracy, as accuracy and speed are emphasized equally in the evaluation of worksheets and achievement tests. This may highlight a theoretical flaw in the Kumon philosophy, and thus why most of the null hypotheses of this study have been accepted thus far.

Behavioural psychologists believe that skills are acquired by systematically progressing through the learning hierarchy: acquisition, fluency, generalization, and then adaptation (Daly & Martens, 1994; Haring et al., 1978). Kumon's intervention starts at the fluency level, expecting students to master accuracy and speed simultaneously. According to behavioural psychology, if instruction does not follow the learning hierarchy, the student's foundation will be shaky and the skill will not be mastered. This could explain why some participants demonstrated limited growth and others even demonstrated negative growth.

Still, many students demonstrated considerable, positive gains in both types of Math skills. Negative and small gains made may be attributable to problems with the measure used to assess mathematical ability. Two of the five students who demonstrated skill deficits were fell in the above average achievement group based on their pretest scores. This means that these students had a smaller window to demonstrate growth at posttesting on the CAT-3. However, one would expect these students to score at least as high as they did at pretesting. It is possible that some students' pretest scores were on the positive side of the standard error of measurement from their true score while their posttest score is on the negative side. This could result in a negative growth score.

There are two other alternative interpretations for the finding that there is no significant difference between gains made in the computation and mathematical reasoning domains. First, this study has a small sample size and thus all interpretations should be tentative. With a larger sample, the mean gains made scores may shift and the difference between means come become significant. Second, there could be extraneous variables influencing children's gains made in mathematical reasoning and computation skills or certain variables could be interacting, having a complex effect on gains made. These relationships are not accounted for in a paired samples t-test.

*Hypothesis Four: The Relationship between Kumon and Achievement in Mathematics*

*Mathematical Reasoning*

The results from multiple regression analysis suggest that gains in mathematical reasoning cannot be predicted by the number of Kumon worksheets completed, pretest Math reasoning scores, or cognitive ability. These findings are consistent with what was predicted. Since Kumon Math worksheets are composed almost entirely of computation exercises, it was expected that children would not make significant gains in mathematical reasoning as a result of participating in Kumon. In other words, the number of worksheets students completed or their mathematical or cognitive abilities at the time of entering the programme can not be used to predict growth in the area of mathematical reasoning because Math reasoning is not specifically exercised by the Kumon programme. Alternative explanations for these null findings follow in the sections below (see "Computation and numerical estimation").

*Computation and Numerical Estimation*

A separate multiple regression analysis reveals that the number of Kumon worksheets completed has a strong (but not statistically significant) relationship with gains made in computation and numerical estimation. There is a significant relationship between gains made in computation skills and pretest scores in the same domain. Contrary to what was predicted, children’s computation skills at the time of beginning Kumon and gains made after six months in the programme are inversely related.

*Number of Kumon worksheets completed.* Although the number of Kumon worksheets completed is not a statistically significant predictor of gains made in computation skills, there is a strong relationship between these variables. These results could suggest that the more Kumon worksheets a child completes, the greater his or her gains will be in computation skills.

Support for this interpretation comes from a comparison of the relationship between number of Kumon worksheets completed and gains made in the two domains of Math skill measured. Multiple regression analyses show that the number of Kumon worksheets completed has a stronger relationship with gains made in computation skills than gains made in Math reasoning skills. This is consistent with the fact that Kumon worksheets target computation skills more so then they target Math reasoning/problem solving skills. However, all interpretations of these results must be conservative at this point in the research as the relationship between worksheets completed and gains made in the computation domain is not statistically significant and also because there are many limitations to this study, such as having a small sample size (see “Limitations”).

Last, it is important to note that the relationship between the number of worksheets completed and gains made in computation skills is influenced by the individual's computation skills at the time of entering Kumon. In fact, thirty percent more of the variance in gains made on the computation subtest is accounted for when pretest computation scores are included in the multiple regression analysis. In contrast, pretest Math reasoning scores only explain an additional 5% of the variance in gains made on the Math reasoning subtest. In sum, these results may suggest that the relationship between Kumon and achievement in Mathematics is complex and influenced by numerous variables. One such variable may be mathematical ability at the time of enrolling in the Kumon programme.

*Pretest mathematical ability.* There is a statistically significant relationship between pretest computation scores and gains made in the same subtest, as was predicted. However, the relationship between pretest scores and gains made in computation skills is not in the predicted direction. The results of the current study suggest that children with lower scores at pretesting make greater gains in computation skills than children with stronger skills at the time of pretesting.

Below average students tend to have more difficulties in all stages of the learning hierarchy (quantitative differences) and weaker processing abilities (qualitative differences) than average and above average students (Augustyniak et al., 2008; Daly & Martens, 1994; Haring et al., 1978; Shalev, 2004; Stroving & Rourke, 1985). As such, it was hypothesized that below average students would demonstrate the smallest gains (when compared to average and above average students) in Math skills following six months of participating in the Kumon Math programme. The results of this study may suggest the opposite; students who enter the Kumon Math programme with weaker computation abilities gain more computation skills than students who enter the programme with higher scores. Perhaps the Kumon Method of Learning appeals more and/or is better suited to below average students than average or above average students.

One of the components of the Kumon Method of Learning is that students begin at a low starting point so that initial work is easily completed (KNA, 2008). This may appeal more to struggling students than students looking for enrichment. Struggling Math students have had fewer successes in Mathematics and as such may be more reluctant to work on Math problems. These experiences of early success in the Kumon Math programme may be more meaningful to below average students, resulting in greater self-confidence and confidence in the programme’s effectiveness.

In contrast, above average students prefer complex, challenging problems (Threlfall & Hargreaves, 2008). These students may become disinterested and lose confidence in the programme’s usefulness in the first few weeks of the programme. Ma (2005) suggests that an under-stimulating Math curriculum will lead to loss of motivation and under-achievement among talented Math students. Perhaps the comfortable starting point in the Kumon programme is advantageous to below average students but hinders above average students’ achievement.

The Kumon Method of Learning may also have a differential effect on students’ achievement because the programme’s strategies better address the needs of below average students than those of advanced students. For example, below average students, specifically students with a MD, tend to have memory deficits (Geary, 1993; Geary & Hoard, 2001; Ginsburg, 1997; Jordan & Montani, 1997; Shalev & Gross-Tsur, 2001).

These students' memory deficits may be remediated by the repeated, daily practice that is central to the Kumon programme. Also, students often struggle in Mathematics because of difficulties in organizing mathematical information spatially (Augustyniak et al., 2005; Geary, 2003; Strang & Rourke, 1985). Kumon worksheets are highly structured both in content and spatial organization so that students can progress incrementally and independently (Izumi, 2001). This likely helps below average students improve their performance in Mathematics.

Further, Mills and colleagues (1994) evaluated the effectiveness of a Math programme comprised of strategies similar to those that are central to Kumon's Method of Learning for children who demonstrated giftedness in Mathematics. The strategies employed include a linear curriculum, individualized learning, and instructional placement based on the child's level at the time of entering the programme (Mills et al., 1994). Participants demonstrated greater gains in Mathematics than controls (Mills et al., 1994). This is further support for the hypothesis that Kumon's strategies are effective teaching techniques for children who are at an above average level in Mathematics.

Despite the many specific, positive effects that the Kumon programme may hold for below average students, there is not enough evidence to conclude that the Kumon Method of Learning is *more* beneficial to below average students than average or above average students or further that Kumon's Method of Learning *hinders* the mathematical achievement of above average students. With the exception of a low starting point, the aforementioned Kumon strategies that help below average students (repeated practice, well-organized curriculum, and independent learning) would most likely enhance all children's development in Mathematics. At the very least, these strategies are not likely detrimental to above average students' learning. An alternative explanation for the negative relationship between pretest scores and gains made is more likely.

Gains made by below average students could be superficially greater than gains made by average and above average students because of the structure of the CAT-3. Both the Mathematics and the computation and numerical estimation subtests have a set number of questions (40 and 26, respectively). As such, the ceiling for this measure is determined by the number of questions and not the students' abilities for high achieving students.

To determine the gains a student made in either subtest, pretest scale scores were subtracted from posttest scale scores. Scale scores are determined based on the number of correct responses the individual made within the subtest. This means that participants who had a high percentage of correct responses at pretesting were more restricted in their ability to demonstrate gains than students who initially had a lower percentage of correct responses. This could explain why multiple regression shows a negative relationship between pretest scores and gains made in computation skills.

However, the negative relationship between pretest scores and gains made only exists for computation skills (the relationship between pretest scores and gains made in mathematical reasoning is insignificant). This could be due to differences in the structure of the two subtests.

The CAT-3 Norms Book states that scale scores "are comparable across test levels (but not across tests, e.g., Reading and Mathematics)" (Canadian Test Centre, 2001, p. 1). The Norms Book does not explicitly state whether comparisons can be drawn between subtests that measure different domains of the same skill (computation and mathematical reasoning skills).

Examination of the CAT-3 subtests and norms reveals that the same number of correct responses on the computation subtest and the Math reasoning subtest corresponds with different scale scores, with number of correct responses corresponding to higher scale scores on the computation subtest. In other words, if a grade four student answers 15 more questions correctly at posttesting than he/she did at pretesting on both subtests, his/her gains made score would be greater in the computation and numerical estimation domain than it would be in the mathematical reasoning domain. However, statistical analysis is required to determine if this difference is significant and thus would make gains made in the Math reasoning and computation domains not directly comparable. Rather, they would be attributable to the structure of the mathematical achievement measure administered.

Alternatively, there may be differences in the relationship between gains made and pretest scores for the two domains of Math skills because of the influence of one more other variables. These other variables may affect gains made in computation skills more than gains made in mathematical reasoning skills. In fact, multiple regression shows that many predictor variables are needed to account for the variance in the criterion variables (gains made in computation and Math reasoning).

One of the variables influencing the relationship between gains made in computation skills and pretest computation scores is the number of Kumon worksheets completed by the student. Multiple regression shows that pretest computation scores explain 24.6% of the variance in gains made in computation skills. Adding the number of Kumon worksheets completed to this model accounts for nine percent more of the variance. In contrast, adding the number of Kumon worksheets completed to the mathematical reasoning model does not help explain any additional variance. This may suggest that Kumon's relationship with achievement in computation skills is stronger than its relationship with growth in mathematical reasoning skills.

Still, approximately two-thirds of the variance in gains made in computation skills (and even more of the variance in gains made in Math reasoning) has yet to be explained. There are other, unmeasured variables influencing children's gains in both mathematical reasoning and computation.

In sum, the most plausible explanation, in the current author's opinion, for the finding that the relationship between pretest computation scores and gains made in computation skills is negative is that the results were skewed by limitations of the CAT-3. The structure of the CAT-3, specifically the ceiling on the CAT-3, limits above average students in their ability to demonstrate growth. Above average students are more limited in their ability to demonstrate gains made on the computation subtest versus the mathematical reasoning subtest because the former has significantly fewer questions.

*Cognitive ability.* It was hypothesized that children with higher cognitive ability scores at the time of pretesting would demonstrate greater gains in Math skills than children who scored in the lower ranges on the cognitive measure. Results indicate that cognitive ability scores do not significantly predict gains made in computation and numerical estimation or mathematical reasoning. This could suggest that a child's cognitive ability does not influence his or her ability to achieve in the Kumon Math programme.

Another explanation for these results is that children with high or above average cognitive abilities are restricted in their ability to achieve in Mathematics by the structure of the Kumon programme. It is possible that the programme progresses so gradually that advanced children are under-stimulated and lose motivation, resulting in under-achievement (Ma, 2005).

In the current author's opinion, this is not the most plausible explanation because individualized learning and advanced study are two of the main strategies behind Kumon's programme (Izumi, 2001; KNA, 2008). The pace of progression is dictated by the individual student's abilities and all students are encouraged to reach their highest potential as quickly as possible (Izumi, 2001; KNA, 2008).

Children's scores on the Verbal section of the CCAT have been used by past researchers as the measure of cognitive ability because of its high correlation with the Henmon-Nelson (Schneider et al., 1989). The Verbal section of the CCAT has also been selected in the past as a measure of the cognitive ability because the skills measured are those most closely related to the abilities needed to achieve academically. Thus, the CCAT likely provides an appropriate measure of cognitive ability.

The literature clearly states that cognitive ability and the ability to achieve in Mathematics are positively correlated. Therefore, it is more likely that the null findings of the current study reflect a limitation of the study. Specifically, the results likely reflect the study's small sample size. In the pretest sample of 42 participants, only one participant fell outside the average range of cognitive ability. This lack of variance in the sample supports the claim that the study's sample was not representative of the general population. The lack of variance in the pretest sample could be due to non-random sampling; perhaps children of below and above average intelligence do not typically enroll in Kumon. However, the sample used to measure the relationship between cognitive ability scores and gains made in Mathematics consisted of only 22 participants. It is unlikely that this small sample is representative of the general population in terms of cognitive ability.

*Total Mathematics*

Total Mathematics is a composite score derived from averaging an individual's scores on the Mathematics subtest and his or her scores on the computation and numerical estimation subtest. As such, multiple regression analysis of the relationship between gains made in total Math and the three predictor variables (number of Kumon worksheets completed, pretest total Math scores, cognitive ability scores) provides an addition measure of the strength of the relationship between participation in Kumon and achievement in Mathematics.

It was hypothesized that the gains made in computation skills would be so great in magnitude as a result of participating in Kumon that gains made in total Math would also have a significant, positive relationship with the number of Kumon worksheets completed. Multiple regression shows that there are no relationships between gains made in total Math and any of the predictor variables. This is consistent with the earlier finding that the relationship between gains made in computation skills and participation in Kumon is not statistically significant.

## *Limitations*

The current study has many limitations. Specifically, there are notable limitations in the study's sampling procedures, instrumentation, and data collection.

*Sampling Procedures*

A major limitation of the current study is that it has a non-experimental design. Without a control group, cause and effect cannot be determined. For example, all significant results may have occurred as a result of maturation. Alternatively, relationships between the variables in this study could reflect the fact that participants spent extra time studying, and not specifically because they were studying with Kumon materials. Also, participants were not randomly assigned to the Kumon group. The participants in the current study may therefore represent a subgroup of the general population that is different in some way. The children who enroll in Kumon voluntarily may come from families that have more money, parents who have achieved higher education, parents that are more involved in their children's academics, etc. All of these factors could influence students' achievement in Mathematics and could have influenced the results of this study. Experimentation is needed to determine whether there is a cause and effect relationship between children's participation in Kumon and achievement in Mathematics.

Noted throughout the discussion section is that the study had a small sample size. With only 22 participants, statistical analyses do not yield meaningful results. This study's results are not representative of the general population and therefore the results are not generalizable to all grade four, five, and six students. As a result, inferences and implications cannot be drawn from the current study's results; statements regarding the usefulness of Kumon for children cannot be made at this point in the research. The statistical analyses used in the current study were for exploratory purposes only and discussion of the study's results exists to fulfill the requirements of this project.

Last, the sample included children enrolled in the Kumon Math programme (18 participants) as well as children enrolled in the Kumon Math and Reading programmes (four participants). Students who are enrolled in both programmes were not excluded as this would have further reduced the study's sample size. This is a potential limitation of the study because the sample may not be homogeneous; students involved in an addition academic intervention could be at an advantage over students who are only participating in Kumon Math. Kumon claims that both of their programmes enhance concentration, study habits, and self-confidence (Izumi, 2001; KNA, 2008). If this is true, then participation in the Reading programme would enhance students' ability to learn in general, including their ability to achievement in Mathematics. This further limits how much gains in Mathematics can be attributed to participation in the Kumon Math programme.

*Instrumentation*

The limitations associated with using the CAT-3 as the measure of mathematical achievement have been discussed at length. The results of the current study may suggest that students were limited in their ability to demonstrate gains made in Mathematics by the ceilings of the CAT-3. Another potential limitation of using the CAT-3 as the measure of gains made in Mathematics is that there is only one version of the test for each level. This could lead to practice effects at posttesting and artificially greater scale score differences.

The Verbal section of the CCAT is best used as a screener for cognitive ability as it does not capture all domains of intelligence and as a group administered test, the ceiling is determined by the number of questions and not each individual's ability. An individually administered measure of cognitive ability, such as the Weschler Intelligence Scale for Children, Fourth Edition (WISC-IV), would provide a more accurate estimate of participants' cognitive ability.

There are also limitations in using the number of Kumon worksheets completed as the measure of participation in Kumon. The criteria for determining mastery of a worksheet are not standardized. Kumon has set guidelines for instructors to follow in evaluating worksheets. However, each instructor is permitted and encouraged to use their personal judgment to determine whether or not a child is ready to advance to the next worksheet. Participants in the current study were recruited from over 15 centres. Across 15 different instructors, there is almost certainly at least minor variation in how 'mastery of a worksheet' is defined.

Another potential limitation of using the number of Kumon worksheets completed as the independent measure is that this measure does not take into account the consistency of participation in the programme. Some students completed approximately the same number of worksheets each week over the six months in question while others had periods of high completion rates and periods of inactivity (for example, some students took several weeks off from Kumon in the summer). It is possible that these two approaches result in different patterns of development and achievement in Mathematics.

*Data Collection*

Data was collected in the students' regular Kumon classrooms. This means that at times, testing conditions were less than ideal. Noise levels within the classroom sometimes escalated and there were a few distractions during testing; for example, people talking to the participant while he or she was completing testing. Since these conditions are different from those during standardization of the CAT-3 and CCAT, it may not be appropriate to use their norms. This is not considered a major limitation of the study as testing conditions were ideal (or very close to ideal) for the majority of testing.

Vancouver participants' pre- and posttest data were collected by Vancouver Kumon field staff due to logistical restraints. Kumon field staff also helped collect some of the pretesting data in the Greater Toronto Area. Toronto field staff members were recruited to help administer the CAT-3 and CCAT to reduce the amount of time between enrollment and pretesting and to reduce the amount of time between CAT-3 and CCAT testing. All field staff were trained by the lead researcher on how to administer these tests in a one hour tutorial. Still, all attempts were made to limit the amount of testing administered by Kumon staff; tests were only administered by Kumon staff if the lead researcher was occupied (administering tests in another location). All posttest data was collected by the lead researcher (save for the data from Vancouver centres). Of note, all participants who were tested by Kumon field staff at pretesting dropped out of the Kumon programme before posttesting.

Kumon staff may have monetary interest in the results of this study. As such, Vancouver field staff may have advertently or inadvertently had a negative influence on students' pretest scores. Students would be more likely to demonstrate significant gains in Mathematics if their pretest scores are lower. Post-hoc analysis was conducted to determine if there were significant differences between pretest scores collected by Kumon staff and pretest scores collected by the independent researcher. There are no significant differences between the two groups. However, these two groups contain very few cases; with larger subsamples differences may emerge.

Instructors were asked to invite all students who enrolled in January and February of 2009 to participate in this study within one week of their enrolment. Instructors have many other priorities and as a result there were often delays in the distribution of consent forms. Some parents took an extended time period to return the consent form. Once parental consent was received, there was often difficulty reaching instructors to schedule pretesting. As a result, the length of time between enrolling in the programme and pretesting was often not ideal. The average number of days between enrollment and pretesting was 45. This could have affected the results of the current study. Pretesting scores may not represent children's mathematical ability prior to starting Kumon. It is possible that children's Math scores change significantly in the first month of participating in Kumon. If so, the pretest scores in the current study do no represent pre-intervention ability and thus inferences about the relationship between children's pre-intervention ability, participation in the Kumon Math programme, and achievement in Mathematics cannot be made using this type of data.

The time period between pre- and posttesting was approximately six months. Students may need more than six months of exposure to Kumon in order to demonstrate significant gains. However, McKenna and colleagues (2005) found that Kumon students made significant improvements in their Math skills after only seven months of receiving the intervention. This project's timeline was approved by a general manager at Kumon Canada, who, of course, has interest in ensuring that the research design allows students the opportunity to demonstrate any gains incurred as a result of participating in the Kumon Math programme. Nonetheless, the timeline for the current study was based more on the researcher's resources than literature stating the recommended length of intervention, as there is no such research. The results of the current study may have been different if the time period of the study was longer (or shorter).

*Implications*

The implications that can be drawn from the current study's results are extremely restricted by the aforementioned limitations. Most notable are the limitations imposed by the small sample size of the study. There is a strong likelihood that the results obtained from the sample in the current study are different from the results that would have been obtained from a larger, more representative sample.

The most important and valid implications of this study's results pertain to researchers. This study makes a contribution to the scientific knowledge base in that it promotes awareness of the dearth of knowledge in the areas of MD, giftedness in Mathematics, and effective remedial and enrichment Math programmes (the Kumon Math programme in particular). This study brings to light many areas in need of research. Specific questions and suggestions for future research are also noted below.

Implications for practitioners and teachers are few as there are so many limitations in this study. Ultimately what some practitioners and teachers will gain from this study is awareness of a Math intervention that they were previously unaware of and reasons why it may or may not be effective based on learning theory. This does not allow them to make recommendations to parents seeking enrichment or remedial Math programmes for their children, but it allows counsellors and educators to provide parents with a more comprehensive understanding of what is available.

*Questions for Future Research*

The research question in the current study (what is the relationship between participation in Kumon and achievement in Mathematics?) needs to be posed again by future researchers. Once the limitations of the current study are addressed, researchers can find a more definitive answer to this question and thus more valuable implications can be drawn for practitioners, teachers, parents, and children. Future researchers should use an experimental design and recruit a greater number of participants. More accurate measures of achievement and cognitive ability should be utilized if possible. For example, future researchers may wish to use individually administered tests as the measures of achievement and cognitive ability, such as the Weschler Individual Achievement Test, Second Edition (WIAT-II) and the WISC-IV, respectively. The researcher should be more involved in participant recruitment to limit reliance on Kumon instructors for data collection.

The majority of variance in gains made in Math skills was unexplained in all multiple regression analysis models within this study. Research is needed to determine what other variables influence Kumon's effectiveness and of these variables, which is the most influential? How do these variables influence development in Mathematics in general? Specifically, researchers should investigate the influence of parental education/income/involvement, student motivation, consistency of participation in the programme, and instructor variables (teaching skills, philosophy on learning and education, demeanor with students).

It would also be interesting to include measures of psychological processing skills in future analyses. How do a child's psychological processing skills, such as working memory, concentration, and visuo-spatial processing, influence a child's ability to benefit from the Kumon Math programme? Do any of these skills improve as a result of participating in Kumon?

Future research may show that the Kumon programme is not an effective intervention for mathematical reasoning skills, computation skills, or both. As such, researchers should also pose questions about the effectiveness of other Mathematics programmes. Is the KeyMath Teach and Practice an effective remediation programme? Is it an effective enrichment programme? Is Lightspan an effective Mathematics intervention for children of varying abilities? Is there a difference in the effectiveness of specific teaching strategies for children of different mathematical ability? If so, what strategies are the most effective for teaching children with MDs, typically developing children, and gifted children?

In sum, there are many gaps in the literature regarding Math interventions for below average, average, and above average students. Based on the results from the current study and this literature, several research questions have been suggested to orient future researchers to the most pressing issues.

### *Conclusion*

The purpose of this study was to determine the relationship between participation in the Kumon Math programme and achievement in Mathematics for children of varying abilities. A nonexperimental, causal comparative research design was employed to answer the research question. Results suggest that there may be a stronger relationship between participation in Kumon and development in computation skills versus Math reasoning skills. Results may also suggest that the relationship between Kumon and achievement in computation skills is stronger for children who start the programme at a below average level than children who start at an average or above average level. However, there were so many limitations to the study, namely in sampling, instrumentation, and data collection, that inferences and implications cannot be drawn from the study's results. The main contribution of the current study then, is to raise awareness of the Kumon Math programme and to examine its structure and strategies in comparison to learning theory. This study also contributes several questions and direction for future research in this area.

# REFERENCES

Alarcon, M., Defries, J., Gillis Light, J., & Pennington, B. (1997). A twin study of mathematics disability. *Journal of Learning Disabilities, 30*, 617-623.

Albers, A. E., & Greer, R. D. (1991). Is the three-term contingency trial a predictor of effective instruction? *Journal of Behavioral Education, 1*, 337-354.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed. text revision). Washington, DC: Author.

Anastasi, A. (1958). *Differential psychology* (3rd ed.). New York, NY: Macmillan.

Anderson, J. O. (1995). Review of the Canadian cognitive abilities test, form 7. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp.153-154). Lincoln, NE: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Anderson, J. O. (2005). Review of the Canadian achievement tests, third edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 180-186). Lincoln, NE: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Ardila, A., & Rosselli, M. (2002). Acalculia and dyscalculia. *Neuropsychology Review, 12*(4), 179-231.

Arthur Steen, L. (1999). Twenty questions about mathematical reasoning. In L. Stiff (Ed.), *Developing mathematical reasoning in grades K-12* (pp.270-285). Reston, VA: National Council of Teachers of mathematics.

Augustyniak, K., Murphy, J., & Kester Phillips, D. (2005). Psychological perspectives in assessing mathematics learning needs. *Journal of Instructional Psychology, 32*(4), 277-286.

Aunio, P., Aubrey, C., Godfrey, R., Pan, Y., & Liu, Y. (2008). Children's early numeracy in England, Finland and People's Republic of China. *International Journal of Early Years Education, 16*(3), 203-221.

Bachor, D. G., & Summers, G. (1985). The Canadian achievement test and the Canadian test of basic skills: A critical review. *BC Journal of Special Education, 9*, 175-179.

Badian, N. A. (1983). Dyscalculia and nonverbal disorders of learning. In H. R. Myklebust (Ed.), *Progress in learning disabilities* (Vol. 5, pp. 235-264). New York, NY: Stratton.

Benbow, C. P., & Stanley, J. C. (1983). *Academic precocity: Aspects of its development.* Baltimore, MD: Johns Hopkins University Press.

Billington, E. J., Skinner, C. H., & Cruchon, N. M. (2004). Improving sixth-grade students' perceptions of high-effort assignments by assigning more work: Interaction of additive interspersal and assignment effort on assignment choice. *Journal of School Psychology, 42*, 477-490.

Bjorklund, D. F. (2005). *Children's thinking: Cognitive development and individual differences.* USA: Thomson Wadsworth.

Burns, M. K., VanDerHeyden, A. M., & Jiban, C. L. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review, 35*(3), 401-418.

Business Editors. (2000). *Lightspan announces new version of its Achieve Now programme to further increase student achievement*. Retrieved March 17, 2009, from http://findarticles.com/p/articles/mi_m0EIN/is_2000_August_15/ai_64192844/

Busse, J., Berninger, V. W., Rury Smith, D., & Hildebrand, D. (2001). Assessment for math talent and disability: A developmental model. In J. J. W. Andrews, D. H. Saklofske, & H. L. Janzen (Eds.), *Handbook of psychoeducational assessment* (pp.225-253). San Diego, CA: Elsevier Press.

Butterworth, B. (1999). *The mathematical brain.* London, England: Macmillan.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46*(1), 3-18.

Byrnes, J. P., & Takahira, S. (1993). Explaining gender differences on SAT-math items. *Developmental Psychology, 29,* 805-810.

Canadian Test Centre. (2001). *Norms book*. Markham, ON: Canadian Test Centre Inc.

Canadian Testing Centre. (2008). *Canadian Achievement Tests, third edition.* Retrieved October 11, 2008, from http://www.canadiantestcentre.com/CAT3/CAT3.asp#Math

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem solving skills. In T. P. Carpenter, J. M. Moser & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (vol. LEA, pp. 9-24). Hillsdale, NJ: LEA.

Cawley, J. F., Kahn, H., & Tedesco, A. (1989). Vocational education and students with learning disabilities. *Journal of Learning Disabilities, 22*(10), 630-634.

Codding, R. S., Eckert, T. L., Fanning, E., Shiyko, M., & Solomon, E. (2007). Comparing mathematics interventions: The effects of cover-copy-compare alone and combined with performance feedback on digits correct and incorrect. *Journal of Behavioral Education, 16*(2), 125-141.

Codding, R. S., Shiyko, M., Russo, M., Birch, S., Fanning, E., & Jaspen, D. (2007). Comparing mathematics interventions: Does initial level of fluency predict intervention effectiveness? *Journal of School Psychology, 45*, 603-617.

Cumming, J. J., & Elkins, J. (1999). Lack of automaticity in the basic addition facts as a characteristic of arithmetic learning problems and instructional needs. *Mathematical Cognition, 5*, 149-180.

Dahinten, V. S., Shapka, J. D., & Willms, J. D. (2007). Adolescent children of adolescent mothers: The impact of family functioning on trajectories of development. *Journal of Youth and Adolescence*, *36*, 195-212.

Daly, E. J., & Martens, B. K. (1994). A comparison of three interventions for increasing oral reading performance: Application of the instructional hierarchy. *Journal of Applied Behavior Analysis*, *27*, 459-469.

Daly, E. J., Witt, J. C., Martens, B. K., & Dool, E. J. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review, 26*, 554-574.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487-506.

Demie, F. (2001). Ethnic and gender differences in educational achievement and implications for school improvement strategies. *Educational Research, 43*(1), 91-106.

Diezmann, C. M., & English, L. D. (2001). Developing young children’s multidigit number sense. *Roeper Review, 24*(1), 11-13.

Ding, C. S., Song, K., & Richardson, L. I. (2007). Do mathematical gender differences continue? A longitudinal study of gender difference and excellence in mathematics performance in the U.S. *Educational Studies, 40*(3), 279-295.

Ellis, H. (1894). *Man and woman: A study of human secondary sexual characters.* London, England: Walter Scott.

Erford, B. T., & Klein, L. (2007). Technical analysis of the Slosson-Diagnostic Math Screener (S-DMS). *Educational and Psychological Measurement*, *67*(1), 132-153.

Ericsson, K. A., Krampe, R., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 199*, 363-406.

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies, 18*(1), 3-56.

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research, 62*, 61-84.

Feldhusen, J. F. (1989). Synthesis of research on gifted youth. *Educational Leadership, 46*(3), 6-11.

Felson, R. B., & Trudeau, L. (1991). Gender differences in mathematics performance. *Social Psychology Quarterly, 54*(2), 113-126.

Fox, L. H. (1976). Sex differences in mathematical precocity: Bridging the gap. In D. P. Keating (Ed.), *Intellectual talent: Research and development.* Baltimore, MD: Johns Hopkins University Press.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1994). Measurement of change: Assessing behavior over time and within a developmental context. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities* (pp. 29-58). Baltimore, MD: Paul H. Brookes Publishing Company.

Fuchs, L. S. & Fuchs, D. (2005). Enhancing mathematical problem solving for students with disabilities. *The Journal of Special Education, 39*(45), 45-57.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 1*, 30-47

Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology, 49*, 363-383.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345-362.

Geary, D. C. (1994). *Children's mathematical development.* Washington, D.C.: American Psychological Association.

Geary, D. C. (2003). Learning disabilities in arithmetic: Problem-solving differences and cognitive deficits. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of Learning Disabilities* (pp. 199-212). New York, NY: The Guilford Press.

Geary, D. C., & Brown, S. C. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. *Developmental Psychology, 27*, 787-797.

Geary, D. C., Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology, 27*, 787-797.

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77*, 236-263.

Geary, D. C., & Hoard, M. K. (2001). Numerical and arithmetical deficits in learning-disabled children: Relation to dyscalculia and dyslexia. *Aphasiology, 15*, 635-647.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with math disability. *Journal of Experimental Child Psychology, 88*, 121-151.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology, 74*, 213-239.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number.* Cambridge, MA: Harvard University Press.

Gelman, R., & Meck, E. (1983). Preschoolers counting: Principles before skill. *Cognition, 13*, 343-359.

Georgiou, S. N., Stavrinides, P., & Kalavana, T. (2007). Is Victor better than Victoria at maths? *Educational Psychology in Practice, 23*(4), 329-342.

Ginsburg, H. P. (1997). Mathematics learning disabilities: A view from developmental psychology. *Journal of Learning Disabilities, 30*, 20-33.

Goldman, S. R., & Pellegrino, J. W. (1987). Information-processing and educational microcomputer technology: Where do we go from here? *Journal of Learning Disabilities, 20*, 144-154.

Gorard, S. G., Rees, G., & Salisbury, J. (2001). Investigating the patterns of differential attainment of boys and girls at school. *British Educational Research Journal, 27*(2), 125-139.

Gray, P. (1999). *Psychology.* New York, NY: Worth Publishers.

Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology, 38*, 25-33.

Halpern, D. F. (1986). *Sex differences in cognitive abilities.* Hillsdale, NJ: Erlbaum.

Hammill, D. D. (1990). On defining learning disabilities: An emerging consensus. *Journal of Learning Disabilities, 23*, 74-84.

Haring, N. G., Lovitt, T. C., Eaton, M. D., & Hansen, C. L. (1978). *The fourth r: Research in the classroom.* Columbus, OH: Charles E. Merrill.

Haslam, L. (2007). Sam's progress with learning mathematics. *Down Syndrome Research and Practice, 12*(1), 32-33.

Hattie, J. (1995). Review of the Canadian cognitive abilities test, form 7. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 154-155). Lincoln, NE: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*, 41-45.

Hetherington, E. M., Parke, R. D., & Schmuckler, M. (2005). *Child psychology: A contemporary viewpoint.* Canada: McGraw-Hill Ryerson.

Hyde, J. S. (2007). New directions in the study of gender similarities and differences. *Current Directions in Psychological Science*, *16*(5), 259-263.

Hyde, J. S., Fennema, E. H., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107,* 139-155.

Isaacs, E. B., Edmonds, C. J., Lucas, A., & Gadian, D. G. (2001). Calculation difficulties in children of very low birthweight: A neural correlate. *Brain*, *124*, 1701-1707.

Izumi, L. T. (2001). *Facing the classroom challenge: Teacher quality and teacher training in California's schools of education*. San Francisco, CA: Pacific Research Institute for Public Policy.

Johnson, J., Im-Bolter, N., & Pascual-Leone, J. (2003). Development of mental attention in gifted and mainstream children: The role of mental capacity, inhibition, and speed of processing. *Child Development, 74*(6), 1594-1614.

Jolivette, K., Lingo, A. S., Houchins, D. E., Barton-Arwood, S. M., & Shippen, M. E. (2006). Building math fluency for students with developmental disabilities and attentional difficulties using Great Leaps Math. *Education and Training in Developmental Disabilities*, *41*(4), 392-400.

Jordan, N., Hanich, L., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development, 74,* 834-850.

Jordan, N., & Montani, T. (1997). Cognitive arithmetic and problem solving: A comparison of children with specific and general mathematics difficulties. *Journal of Learning Disabilities*, *30*, 624-634.

Kail, R. V. (2000). Speed of information-processing: Developmental change and links to intelligence. *Journal of School Psychology, 38*, 51-61.

Klein, A., & Starkey, P. S. (1987). The origins and development of numerical cognition: A comparative analysis. In J. A. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 1-25). Oxford, England: Clarendon Press.

Koontz, K. L., & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. *Mathematical Cognition, 2*, 1-23.

Kosc, L. (1974). Developmental dyscalculia. *Journal of Learning Disabilities, 7*, 159-162.

Kronenberger, W. G., & Dunn, D. W. (2003). Learning disorders. *Neurologic Clinics, 21*, 941-952.

Kumon North America. (2008). *About Kumon.* Retrieved June 27, 2008, from http://www.kumon.com/about/default.asp?language=USA.

Kumon Toru Research Institute of Education. (2002). *Expanding one's strong point even further*. Japan: Kumon Toru Research Institute of Education.

Leahey, E., & Guo, G. (2001). Gender differences in mathematical trajectories. *Social Forces, 80*, 713-732.

Learning Disabilities Association of Ontario. (2009). *About learning disabilities.* Retrieved April 8, 2009, from http://www.ldao.ca/aboutLDs/Definitions_of_LDs.php

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10- year old boys and girls. *Journal of Child Psychology and Psychiatry, 35*, 283-292.

Lewis, R. B. (1998). Assistive technology and learning disabilities: Today's realities and tomorrow's promises. *Journal of Learning Disabilities, 31*, 16-26.

Lohman, D. F. (2005). The role of nonverbal ability tests in identifying academically gifted students: An aptitude perspective. *Gifted Child Quarterly, 49*, 111-136.

Lubinski, D., & Persson Benbow, C. (2006). Study of mathematically precocious youth after 35 years. *Perspectives on Psychological Science, 1*(4), 316-345.

Ma, X. (2005). A longitudinal assessment of early acceleration of students in mathematics on growth in mathematics achievement. *Developmental Review, 25*, 104-132.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences.* Stanford, CA: Stanford University Press.

Mazzocco, M. (2001). Math learning disability and math ld subtypes: Evidence from studies of Turner syndrome, fragile X, and neurofibromatosis type 1. *Journal of Learning Disabilities*, *34*, 520-533.

Mazzocco, M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia, 53*, 218-253.

McKenna, M. A., Hollingsworth, P. L., & Barnes, L. L. B. (2005). Developing latent mathematics abilities in economically disadvantaged students. *Roeper Review, 27*(4), 222-227.

McLean, J. F., Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240-260.

Michalko, K. T., & Saklofske, D. H. (1996). A psychometric investigation of the Wechsler individual achievement test with a sample of Saskatchewan schoolchildren. *Canadian Journal of School Psychology, 12*(1), 44-54.

Miller, L. T., & Vernon, P. A. (1997). Developmental changes in speed of information-processing in young children. *Developmental Psychology, 33*, 549-554.

Miller, P. H., & Weiss, M. G. (1981). Children's attention allocation, understanding of attention, and performance on the incidental learning task. *Child Development, 52*, 1183-1190.

Miller, S. P., & Mercer, C. D. (1997). Educational aspects of mathematics disabilities. *Journal of Learning Disabilities, 30*, 47-56.

Mills, C. J., Ablard, K. E., & Gustin, W. C. (1994). Academically talented students' achievement in a flexibly paced mathematics programme. *Journal of Research in Mathematics Education, 25*(5), 495-511.

Mississippi Department of Education. (2006). *Endorsed list of scientifically-research based comprehensive, supplemental, and intervention mathematics programmes in grades K-12$^{th}$*. Retrieved February 15, 2008, from http://www.mde.k12.ms.us/ACAD/ID/Curriculum/Math/endorsed_list_2.pdf

Molko, N., Cachia, A., Rivière, D., Mangin, J-F., Bruandet, M., Le Bihan, D., et al. (2003). Functional and structural alterations of the intraparietal sulcus in a developmental dyscalculia of genetic origin. *Neuron, 40*, 847-858.

Montague, M., & van Garderen, D. (2003). A cross-sectional study of mathematics achievement, estimation skills, and academic self-perception in students of varying ability. *Journal of Learning Disabilities, 36*(5), 437-448.

Morocz, I. A., Gross-Tsur, A., von Aster, M., Manor, O., Breznitz, Z., Karni, A., et al. (2003). Functional magnetic resonance imaging in dyscalculia: Preliminary observations. *Annals of Neurology, 54*(S7), S145.

National Center for Educational Statistics. (1996). *Pursuing excellence.* Washington, DC: U.S. Department of Education.

Nelson Education. (2008). *CCAT form k (1998).* Retrieved October 30, 2008, from http://www.assess.nelson.com/group/ccat-k.html

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science, 297*(5587), 1708-1711.

Niederer, K., Irwin, R. J., Irwin, K. C., & Reilly, I. L. (2003). Identification of mathematically gifted children in New Zealand. *High Ability Studies, 14*(1), 71-84.

Niikura, H. (2006a). *The basic design of the standard completion time.* Japan: Kumon Toru Research Institute of Education.

Niikura, H. (2006b). *The diagnostic test and starting point and the achievement test*. Japan: Kumon Toru Research Institute of Education.

Ostad, S. E. (1998). Comorbidity between mathematics and spelling difficulties. *Logopedics Phoniatric Vocology, 23,* 145-154.

Patton, J. R., Cronin, M. E., Bassett, D. S., & Koppel, A. E. (1997). A life skills approach to mathematics instruction: Preparing students with learning disabilities for the real-life math demands of adulthood. *Journal of Learning Disabilities, 30,* 178-187.

Pearson (2006). *KeyMath teach and practice: TAP*. Retrieved September 1, 2008, from http://ags.pearsonassessments.com/group.asp?nGroupInfoID=a6880

Penner, A. M. (2003). International gender X item difficulty interactions in mathematics and science achievement tests. *Journal of Educational Psychology, 95*, 650-656.

Persson Benbow, C., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science, 222*, 1029-1033.

Phelps, L. A., & Hanley-Maxwell, C. (1997). School-to-work transitions for youth with disabilities: A review of outcomes and practices. *Review of Educational Research, 67*(2), 197-226.

Poncy, B. C., Skinner, C. H., & Jaspers, K. E. (2007). Evaluating and comparing interventions designed to enhance math fact accuracy and fluency: Cover, copy, and compare versus taped problems. *Journal of Behavioral Education, 16*, 27-37.

Preckel, F., Goetz, T., Pekrun, R., & Kleine, M. (2008). Gender differences in gifted and average-ability students: Comparing girls' and boys' achievement, self-concept, interest, and motivation in mathematics. *Gifted Child Quarterly, 52*, 146-159.

Reichel, H. C. (1997). Identifying and promoting mathematically gifted pupils and students (12-20 years). *High Ability Studies, 8*(2), 223-232.

Rosenberger, P. B. (1989). Perceptual-motor and attentional correlates of developmental dyscalculia. *Annals of Neurology, 26*, 216-220.

Saklofske, D. H., Caravan, G., & Schwartz, C. (2000). Concurrent validity of the Wechsler abbreviated scale of intelligence (WASI) with a sample of Canadian children. *Canadian Journal of School Psychology*, *16*(1), 87-94.

Schliemann, A. D., & Carraher, D. W. (2002). The evolution of mathematical reasoning: Everyday versus idealized understandings. *Developmental Review, 22*, 242-266.

Schneider, B. H., Clegg, M. R., Byrne, B. M., Ledingham, J. E., & Crombie, G. (1989). Social relations of gifted children as a function of age and school programme. *Journal of Educational Psychology, 81*(1), 48-56.

Scott, K. S. (1993). Multisensory mathematics for children with mild disabilities. *Exceptionality, 4*(2), 97-111.

Shalev, R. S. (2004). Developmental dyscalculia. *Journal of Child Neurology, 19*, 765-771.

Shalev, R. S., Auerbach, J., & Gross-Tsur, V. (1995). Developmental dyscalculia behavioral and attentional aspects: A research note. *Journal of Child Psychology and Psychiatry, 36*, 1261-1268.

Shalev, R. S., Auerbach, J., Manor, O., & Gross-Tsur, V. (2000). Developmental dyscalculia: Prevalence and prognosis. *European Child and Adolescent Psychiatry, 9*, 58-64.

Shalev, R. S., & Gross-Tsur, V. (1993). Developmental dyscalculia and medical assessment. *Journal of Learning Disabilities, 26*, 134-137.

Shalev, R. S., & Gross-Tsur, V. (2001). Developmental dyscalculia. *Pediatric Neurology, 24*, 337-342.

Shalev, R. S., Manor, O., Auerbach, J., Gross-Tsur, V. (1998). Persistence of developmental dyscalculia: What counts? *Journal of Pediatrics, 3*, 358-362.

Shalev, R. S., Manor, O., & Gross-Tsur, V. (1997). Neuropsychological aspects of developmental dyscalculia. *Mathematical Cognition, 3*, 105-120.

Shalev, R. S., Manor, O., Kerem, B., Ayali, M., Badichi, N., Friedlander, Y., & Gross-Tsur, V. (2001). Developmental dyscalculia is a familial learning disability. *Journal of Learning Disabilities*, *34*, 59-65.

Shore, B. M., & Kanevsky, L. S. (1993). Thinking processes: Being and becoming gifted. In K. A. Heller, F. J. Monks, & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp.133-147). Oxford, England: Pergamon Press.

Siegler, R. S. (1998). *Children's thinking* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Siegler, R. S. (2001). Cognition, instruction, and the quest for meaning. In S. M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 195-203). Mahweh, NJ: Erlbaum.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Silverman, L. K. (1989). The highly gifted. In J. F. Feldhusen, J. Van Tassel-Baska, & K. Seeley (Eds.), *Excellence in educating the gifted* (pp. 71-84). Denver, CO: Love Publishing.

Skinner, C. H. (2002). An empirical analysis of interspersal research: Evidence, implications and applications of the discrete task completion hypothesis. *Journal of School Psychology, 40,* 347-368.

Skinner, C. H., McLaughlin, T. F., Logan, P. (1997). Cover, copy, and compare: A self-managed academic intervention effective across skills, students, and settings. *Journal of Behavioral Education, 7*(3), 295-306.

Skinner, C. H., Pappas, D. N., & Davis, K. A. (2005). Enhancing academic engagement: Providing opportunities for responding and influencing students to choose to respond. *Psychology in the Schools, 42*, 389-403.

Soares, L. M. (2005). Review of the Canadian achievement tests, third edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 180-186). Lincoln, NE: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Spies, R. A., & Plake, B. S. (2005). Canadian achievement tests, third edition. In R. A. Spies & B. S. Plake (Eds.), *The sixteenth mental measurements yearbook* (pp. 180-186). Lincoln, NE: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Stading, M., Williams, R. L., & McLaughlin, T. F. (1996). Effects of copy, cover, and compare procedure on multiplication facts mastery with a third grade girl with learning disabilities in a home setting. *Education and Treatment of Children, 19,* 425-434.

Stage, E. K., Kreinberg, N., Eccles, J. R., & Becker, J. R. (1985). Increasing the participation and achievement of girls and women in mathematics, science, and engineering. In S. S. Klein (Ed.), *Handbook for achieving sex equity through education* (pp. 237-269). Baltimore, MD: Johns Hopkins University Press.

Stanley, J. C. (1973). Accelerating the educational progress of intellectually gifted youths. *Educational Psychologist, 10*, 133-146.

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

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science, 210*(4473), 1033-1035.

Steiner, H. H., & Carr, M. (2003). Cognitive development in gifted children: Towards a more precise understanding of emerging differences in intelligence. *Educational Psychology Review*, *15*(3), 215-246.

Sternberg, R. J. (2001). Giftedness as developing expertise: A theory of the interface between high abilities and achieved excellence. *High Ability Studies*, *12*(2), 159-179.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10*, 349-367.

Stokes, T. F., & Osnes, P. G. (1989). An operant pursuit of generalization. *Behavior Therapy, 20*(3), 337-355.

Strang, J. D., & Rourke, B. P. (1985). Arithmetic disabilities subtypes: The neuropsychological significance of specific arithmetic impairment in childhood. In B. P. Rourke (Ed.), *Neuropsychology of learning disabilities* (pp. 87-101). New York, NY: Guildford Press.

Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology, 56*, 87-114.

Swanson, H. L. (2006). Cognitive processes that underlie mathematical precociousness in young children. *Journal of Experimental Child Psychology*, *93*, 239-264.

Swanson, H. L., & Beebe-Frankenberger, M. E. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 96*, 471-491.

Swanson, H. L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research, 76*(2), 249-274.

Swiatek, M. A., & Benbow, C. P. (1991). A 10-year longitudinal follow-up of participation in a fast-paced mathematics course. *Journal for Research in Mathematics Education, 22*, 138-150.

Temple, C. M., & Sherwood, S. (2002). Representation and retrieval of arithmetic facts: Developmental difficulties. *Quarterly Journal of Experimental Psychology, 55A*(3), 733-752.

Thorndike, R. L., & Hagen, E. P. (1998). *Canadian cognitive abilities test: Teacher's handbook.* Scarborough, ON: ITP Nelson.

Threlfall, J., & Hargreaves, M. (2008). The problem-solving methods of mathematically gifted and older average-attaining students. *High Ability Studies, 19*(1), 83-98.

Tsui, J. M., & Mazzocco, M. M. M. (2007). Effects of mathematics anxiety and perfectionism on timed versus untimed mathematics testing in mathematically gifted sixth graders. *Roeper Review, 29*(2), 132-139.

van Garderen, D., & Montague, M. (2003). Visual-spatial representation, mathematical problem solving, and students of varying abilities. *Learning Disabilities Research & Practice, 18*(4), 246-254.

Westbury, M. (1994). The effect of elementary grade retention on subsequent school achievement and ability. *Canadian Journal of Education, 19*(3), 241-250.

Winner, E. (2000a). Giftedness: Current theory and research. *Current Directions in Psychological Science, 9*(5), 153-156.

Winner, E. (2000b). The origins and ends of giftedness. *American Psychologist, 55*(1), 159-169.

Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: Numerical discrimination and its signature limits. *Developmental Science, 8*(2), 173-181.

Woodward, J., & Montague, M. (2002). Meeting the challenge of mathematics reform for students with LD. *The Journal of Special Education, 36*, 89-101.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749-751.

Zhou, Z., Peverly, S. T., & Lin, J. (2005). Understanding early mathematical competencies in American and Chinese children. *School Psychology International*, *26*(4), 413-427.

# TABLES

Table 1

*Math Curriculum Levels and Descriptions*

| Level | Description |
|---|---|
| 7A | Counting to 10 |
| 6A | Counting to 30 |
| 5A | Line drawing, number puzzles to 50 |
| 4A | Reciting and writing numbers to 220 |
| 3A | Adding up to 5 |
| 2A | Adding up to 10, subtracting from numbers up to 10 |
| A | Horizontal addition, subtraction of larger numbers |
| B | Vertical addition, subtraction |
| C | Multiplication, division |
| D | Long division, introduction to fractions |
| E | Fractions |
| F | Four operations, decimals |
| G | Positive/negative numbers, introduction to algebra |
| H | Linear equations, inequalities and graphing |
| I | Factorization, square roots, quadratic equations |
| J | Advanced algebra |
| K | Functions – quadratic, fractional, irrational, exponential |
| L | Logarithms, calculus |
| M | Trigonometry, graphs, and inequalities |
| N | Loci and quadratic inequalities |
| O | Advanced differentiation and differential calculus |
| XT (Optional) | Triangles |
| XV (Optional) | Vectors, equations of lines |
| XM (Optional) | Matrices, mapping, and transformations |
| XP (Optional) | Permutations and probability |
| XS (Optional) | Statistics |

Table 2

*Disability Demographics of Children in Japanese Kumon Centres*

| Disability | Number of students |
|---|---|
| Autism | 1418 |
| Mental Retardation | 1316 |
| Down Syndrome | 606 |
| LD/ADHD | 560 |
| Anacusis (deafness) | 164 |
| Cerebral Paralysis | 142 |
| Epilepsy | 114 |
| Physical Disability | 118 |
| Emotional Disorder | 99 |
| Amblyopia (visual impairment) | 29 |
| Other | 117 |

*Note*. From KTRIE (2002).

Table 3

*Subtests of the CAT-3*

| Skill | Basic Battery | Supplemental Tests | Constructed Response |
|---|---|---|---|
| Reading | Reading/Language | Word Analysis/Vocabulary | |
| Language | Reading/Language | Language/Writing | |
| Spelling | | Spelling | Dictation |
| Mathematics | Mathematics | Computation and Numerical Estimation | Math Reasoning |
| Writing | | Language/Writing | Narrative/ Informational/Letter |

*Note*. From Soares (2005).

Table 4

*CCAT Structure for Levels A-H*

| Battery | Subtest | Number of items |
|---|---|---|
| Verbal | Verbal Classification | 20 |
| | Sentence Completion | 20 |
| | Verbal Analogies | 25 |
| Quantitative | Quantitative Relations | 25 |
| | Number Series | 20 |
| | Equation Building | 15 |
| Non-Verbal | Figure Classification | 25 |
| | Figure Analogies | 25 |
| | Figure Analysis | 15 |

*Note*. From Thorndike & Hagen (1998, p. 8).

Table 5

*Winter CAT-3 and CCAT Testing Levels by Grade*

| Grade | CAT-3 level | CCAT level |
|---|---|---|
| 4 | 14 | B |
| 5 | 15 | C |
| 6 | 16 | D |

*Note*. From Canadian Testing Centre (2008) and Nelson Education (2008).

Table 6

*Testing Order*

| Day | Order | Test |
|---|---|---|
| 1 | 1 | CAT-3 Mathematics |
| | 2 | CAT-3 Computation and Numerical Estimation |
| 2 | 3 | CCAT Verbal Classification |
| | 4 | CCAT Sentence Completion |
| | 5 | CCAT Verbal Analogies |

Table 7

*Posttesting Measures Administered*

| Grade at pretesting | CAT-3 level | CAT-3 subtests |
|---|---|---|
| 4 | 14 | Mathematics and Computation and Numerical Estimation |
| 5 | 15 | Mathematics and Computation and Numerical Estimation |
| 6 | 16 | Mathematics and Computation and Numerical Estimation |

Table 8

*Participant Groups Based on Pretest Scores*

| CCAT Score | CAT-3 Total Math Percentile Ranking | Qualitative Descriptor |
|---|---|---|
| $\geq$ 132 | | Above Average Cognitive Ability |
| 69-131 | $0.1^{st}$ – $25^{th}$ | Below Average |
| | $26^{th}$ – $74^{th}$ | Average |
| | $75^{th}$ – $96^{th}$ | Above Average |
| | $97^{th}$ – $99.9^{th}$ | “Gifted” |

Table 9

*Descriptive Statistics for Pretest Sample*

| | | CAT-3 Math | | CAT-3 CNE | | CAT-3 Total | | CCAT | |
|---|---|---|---|---|---|---|---|---|---|
| Group | *n* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Grade | | | | | | | | | |
| 4 | 9 | 473 | 64 | 444 | 56 | 459 | 49 | 103 | 17 |
| 5 | 15 | 492 | 76 | 493 | 88 | 493 | 76 | 102 | 14 |
| 6 | 18 | 518 | 76 | 500 | 51 | 509 | 60 | 97 | 16 |
| Gender | | | | | | | | | |
| Male | 18 | 513 | 86 | 492 | 79 | 503 | 76 | 102 | 16 |
| Female | 17 | 490 | 66 | 482 | 63 | 486 | 58 | 99 | 15 |
| Programme | | | | | | | | | |
| Math | 30 | 507 | 74 | 485 | 74 | 496 | 68 | 104 | 14 |
| Math/Reading | 12 | 481 | 76 | 489 | 58 | 485 | 60 | 91 | 15 |
| Achievement Group | | | | | | | | | |
| "Gifted" | 1 | | | | | | | | |
| Above Average | 6 | 568 | 46 | 583 | 23 | 575 | 25 | 112 | 11 |
| Average | 23 | 514 | 53 | 488 | 40 | 501 | 37 | 103 | 12 |
| Below Average | 12 | 424 | 46 | 419 | 47 | 421 | 35 | 86 | 12 |
| Total | 42 | 500 | 74 | 486 | 69 | 493 | 66 | 100 | 15 |

*Note*. CNE = computation and numerical estimation; Total = Total Math score on the CAT-3; CCAT = standard age score on the Canadian Cognitive Abilities Test; Math/Reading = students enrolled in both the Kumon Math and Reading programmes.

Table 10

*Chi-square Analysis*

| Variable | $\chi^2$ | *p* |
|---|---|---|
| Grade | 0.692 | 0.708 |
| Sex | 2.888 | 0.089 |
| Kumon Programme | 1.163 | 0.281 |

Table 11

*Pretest Descriptive Statistics for Final Sample*

| Group | *n* | CAT-3 Math | | CAT-3 CNE | | CAT-3 Total | | CCAT | |
|---|---|---|---|---|---|---|---|---|---|
| | | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Grade | | | | | | | | | |
| 4 | 5 | 505 | 52 | 454 | 75 | 479 | 54 | 110 | 14 |
| 5 | 8 | 485 | 65 | 484 | 101 | 485 | 79 | 103 | 13 |
| 6 | 9 | 532 | 92 | 506 | 60 | 519 | 72 | 101 | 17 |
| Gender | | | | | | | | | |
| Male | 11 | 529 | 79 | 489 | 78 | 509 | 74 | 104 | 13 |
| Female | 11 | 493 | 70 | 483 | 82 | 488 | 68 | 104 | 17 |
| Programme | | | | | | | | | |
| Math | 18 | 505 | 76 | 476 | 81 | 491 | 71 | 104 | 15 |
| Math/Reading | 4 | 532 | 76 | 529 | 55 | 530 | 61 | 102 | 14 |
| Achievement Group | | | | | | | | | |
| Gifted | 1 | | | | | | | | |
| Above Average | 5 | 574 | 48 | 587 | 23 | 580 | 25 | 114 | 11 |
| Average | 10 | 530 | 53 | 486 | 34 | 508 | 30 | 106 | 13 |
| Below Average | 6 | 425 | 44 | 402 | 57 | 413 | 17 | 91 | 14 |
| Total | 22 | 510 | 75 | 486 | 28 | 498 | 70 | 104 | 15 |

*Note*. CAT-3 = Canadian Achievement Test, Third Edition; CNE = computation and numerical estimation; Total = Total Math score on the CAT-3; CCAT = standard age score on the Verbal battery of the Canadian Cognitive Abilities Test.

Table 12

*Pretest, Posttest, and Gains Made Comparisons between Males and Females*

| Dependent Variable | *t* | *p* |
|---|---|---|
| Pretest Measures | | |
| CAT-3 Mathematics | -1.126 | .274 |
| CAT-3 CNE | -.177 | .861 |
| CAT-3 Total Mathematics | -.693 | .497 |
| CCAT Standard Age Score | -.011 | .991 |
| Posttest Measures | | |
| Kumon Levels Advanced | -.999 | .330 |
| CAT-3 Mathematics | -1.261 | .223 |
| CAT-3 CNE | -.949 | .355 |
| CAT-3 Total Mathematics | -1.164 | .259 |
| Scaled Score Differences | | |
| CAT-3 Mathematics | -.183 | .857 |
| CAT-3 CNE | -.823 | .421 |
| CAT-3 Total Mathematics | -.799 | .434 |

*Note.* CNE = computation and numerical estimation.

Table 13

*Summary of ANOVA Results Comparing the Gains Made on the CAT-3 Mathematics Subtests among Achievement Groups*

| CAT-3 Subtest | *F* | *p* |
|---|---|---|
| Mathematics | 1.634 | .223 |
| Computation and Numerical Estimation | 2.576 | .104 |
| Total Mathematics | 2.841 | .085 |

Table 14

*Summary of Means and Standard Deviations for Scaled Score Differences on the CAT-3 Subtests for each Achievement Group*

| | Mathematics | | Computation and Numerical Estimation | | Total | |
|---|---|---|---|---|---|---|
| Achievement Group | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Below Average | 48 | 31 | 88 | 32 | 68 | 25 |
| Average | 15 | 40 | 59 | 59 | 37 | 32 |
| Above Average | 35 | 34 | 9 | 78 | 22 | 44 |

Table 15

*Summary of Means, Standard Deviations, and Range of the Number of Kumon Worksheets Completed among Achievement Groups*

| | | | Range | |
|---|---|---|---|---|
| Group | *M* | *SD* | Minimum | Maximum |
| Below Average | 455 | 115 | 339 | 601 |
| Average | 501 | 153 | 270 | 740 |
| Above Average | 551 | 494 | 230 | 1426 |

Table 16

*Summary of Means, Standard Deviations, and Range of Gains Made on CAT-3 Subtests*

| | | | Range | |
|---|---|---|---|---|
| Subtest | *M* | *SD* | Minimum | Maximum |
| Mathematics | 29 | 37 | -44 | 102 |
| Computation and Numerical Estimation | 55 | 62 | -76 | 170 |

Table 17

*Multiple Regression Analysis Predicting Gains Made in Mathematical Reasoning from Number of Kumon Levels Advanced, Pretest Math Scores, and Cognitive Ability Scores*

| Predictor | *b* | *SE b* | β | *t* | *p* |
|---|---|---|---|---|---|
| Kumon Levels Advanced | -.006 | .037 | -.043 | -.170 | .867 |
| CAT-3 Mathematics Score | .026 | .174 | .051 | .147 | .885 |
| CCAT Standard Age Score | -1.223 | .858 | -.485 | -1.425 | .172 |

Table 18

*Multiple Regression Analysis Predicting Gains Made in Computation and Numerical Estimation from Number of Kumon Levels Advanced, Pretest Math Scores, and Cognitive Ability Scores*

| Predictor | *b* | *SE b* | β | *t* | *p* |
|---|---|---|---|---|---|
| Kumon Levels Advanced | .096 | .046 | .389 | 2.077 | .053 |
| CAT-3 CNE Score | -.544 | .170 | -.683 | -3.206 | .005 |
| CCAT Standard Age Score | .936 | .898 | .223 | 1.042 | .312 |

*Note.* CNE = Computation and Numerical Estimation

Table 19

*Multiple Regression Analysis Predicting Gains Made in Total Mathematics from Number of Kumon Levels Advanced, Pretest Math Scores, and Cognitive Ability Scores*

| Predictor | b | SE b | β | t | p |
|---|---|---|---|---|---|
| Kumon Levels Advanced | .043 | .033 | .300 | 1.288 | .215 |
| CAT-3 Total Math Score | -.143 | .157 | -.276 | -.907 | .377 |
| CCAT Standard Age Score | -.371 | .733 | -.152 | -.506 | .619 |