

M<sup>3</sup>: THE THREE-MATHEMATICAL MINDS MODEL FOR THE IDENTIFICATION  
OF MATHEMATICALLY GIFTED STUDENTS

by

Ugur Sak

A Dissertation Submitted to the Faculty of the  
DEPARTMENT OF SPECIAL EDUCATION, REHABILITATION,  
AND SCHOOL PSYCHOLOGY

In Partial Fulfillment of the Requirements  
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2005

THE UNIVERSITY OF ARIZONA  
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation, prepared by Ugur Sak and entitled M<sup>3</sup>: The Three-Mathematical Minds Model for the Identification of Mathematically Gifted Students, and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of DOCTOR OF PHILOSOPHY.

C. June Maker \_\_\_\_\_ Date: 04/11/2005

Lawrence M. Aleamoni \_\_\_\_\_ Date: 04/11/2005

Shirley Schiever \_\_\_\_\_ Date: 04/11/2005

Shitala Mishra \_\_\_\_\_ Date: 04/11/2005

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

C. June Maker

04/11/2005

\_\_\_\_\_  
Dissertation Director

\_\_\_\_\_  
Date

### STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: \_\_\_\_\_

## ACKNOWLEDGMENTS

This work has come about through the support and advice of a number of scholarly individuals to whom I am indebted. First, I would like to express my sincere appreciation to Dr. C. June Maker, who not only has advised me academically but also has become a close friend during my four-year doctoral program. She vehemently has supported me to advance in my scholarly life.

Dr. Shirley Schiever has been an unusually great teacher to me. Her interpersonal relationships and skills have inspired me ever since I met her. I learned invaluable teaching skills from her. I also am thankful for her advice and editorial help during the writing of my dissertation.

I would like to thank to my committee members Drs. James Chalfant, Lawrence M. Aleamoni and Shitala Mishra for their advice and their sincere comments that have helped me improve my work.

I also would like to thank to Dr. Jennifer Yates for her editorial help in the writing of my dissertation.

I also am grateful to Drs. Çetin Urtis and Cumali Öksüz for their advice about mathematics.

I also would like to thank Murat Bıyık, Rıdvan Birçiçek, Debbie Chandler and Larry Chandler for their help in data collection.

Finally, I would like to express my gratefulness to the Turkish Ministry of Education and the New York Turkish Educational Attaché for their financial support during my graduate studies.

Dedicated to My Children

Çağrı and Hilâl

## TABLE OF CONTENTS

LIST OF TABLES .....	8
LIST OF FIGURES .....	10
ABSTRACT .....	11
CHAPTER I .....	13
INTRODUCTION .....	13
Significance of the Study .....	15
Purpose .....	27
Research Questions .....	27
CHAPTER II .....	30
LITERATURE REVIEW .....	30
Part I	
The Nature of Mathematics .....	31
The Nature of Mathematical Ability .....	38
Part II	
Searching for Different Mathematical Minds .....	55
Part III	
Psychological and Neuropsychological Investigations about Creative and Analytical Abilities .....	68
Part IV	
The Three-Mathematical Minds Model .....	83
CHAPTER III .....	89
METHOD .....	89
Participants .....	89
Procedure .....	89
Instruments .....	92
Data Analysis .....	106
CHAPTER IV .....	110
RESULTS .....	110
Research Question 1 .....	110
Research Question 2 .....	118
Research Question 3 .....	127
Research Question 4 .....	130
Research Question 5 .....	133
Summary of Findings .....	139
CHAPTER V .....	141
DISCUSSION AND CONCLUSION .....	141
The Three-Mathematical Minds Model .....	141
The Three-Level Cognitive Complexity Model .....	157
Study Limitations .....	161
APPENDIX A .....	163
Crosstabulation of Participants .....	164
Interitem Biserial Correlations .....	164

TABLE OF CONTENTS – *Continued*

Distribution of Correct Answers in the M <sup>3</sup> .....	169
APPENDIX B .....	184
M <sup>3</sup> problems and their orders in the test booklet .....	185
Solutions of the M <sup>3</sup> Problems .....	200
Problem Characteristics .....	207
Cover Page of the M <sup>3</sup> .....	209
APPENDIX C .....	210
Student Questionnaire .....	211
Teacher Instruction .....	212
APPENDIX D .....	213
Epilogue .....	214
REFERENCES .....	216

## LIST OF TABLES

TABLE 1.1,	Integration of the three-mathematical minds and the three levels of The cognitive complexity model for the assessment of mathematical ability to identify mathematically gifted students .....	27
TABLE 2.1,	Number of divisions of a space by the number of dividing planes, of a plane by the number of dividing lines, and of a line by the number of dividing points.....	51
TABLE 3.1,	Frequency of participants by school .....	90
TABLE 3.2,	Frequency of participants by grade .....	90
TABLE 3.3,	Frequency of participants by race .....	91
TABLE 3.4,	Frequency of participants by gender .....	91
TABLE 3.5,	Item development according to the three-mathematical minds model with the three-levels of the cognitive complexity model.....	96
TABLE 3.6,	Estimation of item discrimination indices based on different levels of different types of mathematical ability .....	109
TABLE 4.1,	Unrotated component matrix .....	112
TABLE 4.2,	Total variance explained by each component .....	113
TABLE 4.3,	Varimax rotation for three factor solution for the $M^3$ .....	114
TABLE 4.4,	Bivariate correlations among the $M^3$ components .....	116
TABLE 4.5,	Bivariate correlations among the $M^3$ subcomponents .....	117
TABLE 4.6,	Bivariate correlations among student variables .....	119
TABLE 4.7,	Partial correlations for grade between student variables and the $M^3$ components and subcomponents.....	120
TABLE 4.8,	Mean and standard deviation for student variables and the $M^3$ subcomponents by grade .....	122
TABLE 4.9,	Item-subtest and item-total test biserial correlations .....	124
TABLE 4.10,	Item difficulty, item cognitive complexity and item discriminations based on the $M^3$ composite and its components .....	128
TABLE 4.11,	Bivariate correlations among ICC, ID and D .....	131
TABLE 4.12,	Summary of Multiple Regression Analysis for ICC predicting ID .....	132
TABLE 4.13,	Mean and standard deviation of performance of all ability groups on three levels of the $C^3$ .....	135
TABLE 4.14,	Tukey HSD Multiple Comparisons .....	136
TABLE 4.15,	Chi Square test for group differences on each item and the proportion of the ability groups passing each item .....	137
TABLE A.1,	Crosstabulation of participants by school, grade, gender and race .....	164
TABLE A.2,	Interitem biserial correlations .....	166
TABLE A.3,	Number of correct answers in the total test battery .....	169
TABLE A.4,	Number of correct answers in the knowledge subtest .....	171
TABLE A.5,	Number of correct answers in the analytical subtest .....	173
TABLE A.6,	Number of correct answers in the creativity subtest .....	175
TABLE A.7,	ANOVA for grade differences on the $M^3$ composite.....	182

LIST OF TABLES – *Continued*

TABLE A.8, Post-Hoc comparisons of mean differences among all grades on the M <sup>3</sup> .....	182
TABLE B.1, M <sup>3</sup> problems and their order in the test booklet .....	185
TABLE B.2, Solutions of the problems .....	200
TABLE B.3, Problem complexity levels, and cognitive components and subcomponents measured by each problem.....	207

## LIST OF FIGURES

FIGURE 2.1,	Three-mathematical minds model .....	87
FIGURE 2.2,	Major instruments of mathematical minds .....	88
FIGURE 3.1,	Continuum of Expertise .....	103
FIGURE A.1,	Distribution of scores in the total test battery .....	170
FIGURE A.2,	Distribution of scores in the knowledge subtest .....	172
FIGURE A.3,	Distribution of scores in the analytical subtest .....	174
FIGURE A.4,	Distribution of scores in the creativity subtest .....	176
FIGURE A.5,	Distribution of scores in the algebra subtest .....	177
FIGURE A.6,	Distribution of scores in the geometry subtest .....	177
FIGURE A.7,	Distribution of scores in the statistics subtest .....	178
FIGURE A.8,	Distribution of scores in the linear syllogism subtest .....	178
FIGURE A.9,	Distribution of scores in the conditional syllogism subtest .....	179
FIGURE A.10,	Distribution of scores in the categorical syllogism subtest .....	179
FIGURE A.11,	Distribution of scores in the selection subtest .....	180
FIGURE A.12,	Distribution of scores in the induction subtest .....	180
FIGURE A.13,	Distribution of scores in the insight subtest .....	181
FIGURE A.14,	Distribution of three ability groups on the $M^3$ .....	182
FIGURE A.15,	Distribution of three ability groups on the $M^3$ .....	182
FIGURE B.1,	Cover page of the test booklet .....	209

## ABSTRACT

Views of giftedness have evolved from unilateral notions to multilateral conceptions. The primary purpose of this study was to investigate the psychological validity of the three-mathematical minds model ( $M^3$ ) developed by the author. The  $M^3$  is based on multilateral conceptions of giftedness to identify mathematically gifted students. Teachings of Poincare and Polya about mathematical ability as well as the theory of successful intelligence proposed by Sternberg (1997) provided the initial framework in the development of the  $M^3$ . A secondary purpose was to examine the psychological validity of the three-level cognitive complexity model ( $C^3$ ) developed by the author. The  $C^3$  is based on studies about expertise to differentiate among gifted, above-average and average-below-average students at three levels.

The author developed a test of mathematical ability based on the  $M^3$  and  $C^3$  with the collaboration of mathematicians. The test was administered to 291 middle school students from four different schools. The reliability analysis indicated that the  $M^3$  had a .72 coefficient as a consistency of scores. Exploratory factor analysis yielded three separate components explaining 55% of the total variance. The convergent validity analysis showed that the  $M^3$  had medium to high-medium correlations with teachers' ratings of students' mathematical ability ( $r = .45$ ) and students' ratings of their own ability ( $r = .36$ ) and their liking of mathematics ( $r = .35$ ). Item-subtest-total score correlations ranged from low to high. Some  $M^3$  items were found to be homogenous measuring only one aspect of mathematical ability, such as creative mathematical ability,

whereas some items were found to be good measures of more than one facet of mathematical ability.

The  $C^3$  accounted for 41% of variance in item difficulty ( $R^2 = .408$ ,  $p < .001$ ). Item difficulty ranged from .02 to .93 with a mean of .29. The analysis of the discrimination power of the three levels of the  $C^3$  revealed that level-two and level-three problems differentiated significantly among three ability levels, but level-one problems did not differentiate between gifted and above average students. The findings provide partial evidence for the psychological validity of both the  $M^3$  and  $C^3$  for the identification of mathematically gifted students.

## CHAPTER I

### INTRODUCTION

Conceptions of giftedness vary among scholars within a field, as well as among those in different fields. This diversity can be explained in part by remarkable growth in knowledge about human intellectual abilities in the last century because of the work of those with keen intellects. More divergent approaches to the study of giftedness are available, both in general and to the assessment of mathematical ability.

Multidisciplinary studies, in particular, have provided sound evidence about the nature of human ability that partially supports divergent theories of ability in which people can be gifted in domains or in processes. Three models of giftedness, for example, have emerged from divergent theories: general giftedness influenced by the theory of general intelligence-*g* (Spearman, 1904), domain specific giftedness influenced by the theory of multiple intelligences (Gardner, 1983, 1999), and process giftedness influenced by the theory of the triarchic mind (Sternberg, 1988, 1997). The theorists postulated thoughtful principles and found evidence to support their theories. For example, Spearman relied on correlational and factorial evidence to support his theory of general intelligence; Sternberg generally used information processing data to support his process-oriented theory of the triarchic mind; and Gardner made use of evidence from different disciplines to support his domain-oriented theory of multiple intelligences.

Following theoretical developments, other issues related to the conceptualization and assessment of giftedness have emerged. Ideas related to these issues are abundant (Heller, Monks, Sternberg & Subotnik, 2000; Sternberg & Davidson, 1986). Probably

the most prevalent question is whether giftedness emerges in different forms. Although the notion of multiple forms of giftedness has received sufficient egalitarian applause from a sociological point of view, how these forms can be assessed and determined objectively has not been given enough attention. The assessment of multilateral giftedness in mathematics is the main subject of this study.

As suggested by the title of this study, the author believes that the assessment of mathematical giftedness should include three forms of giftedness: experts, analysts and creators. Note that these forms of giftedness, the three mathematical minds, are kinds of giftedness that can be measured at particular levels of performance on an expertise continuum. I do not intend to discover some hidden potential that may show up one day or never. As a matter of fact, the model might constitute a foundation for assessment practices of both mathematical competence and mathematical potential, rather than those of only mathematical potential or only competence. That is to say, the assessment of competence also includes the assessment of potential because competence does not develop without potential. The differentiation is similar to that between buds and blooms; that is, buds have promise to flower while blooms already have done so. In fact, how potentially gifted students are identified also is questionable, because most tests measuring potential employ items that are thought to be free of domain knowledge and experience. Yet, giftedness, particularly expertise and creativity, belongs to a domain (Chi, Glaser & Farr, 1988; Csikzentmihalyi, 1997; Gardner, 1999; Sternberg & Lubart, 1995). My reason for relating giftedness to expertise is that giftedness develops upon learning and exposure in a domain of human performance. Numerous interviews with

and biographies of youngsters who participated in competition at the Athens 2004 Olympics indicated that many years of hard body and mind work were needed to demonstrate superior performance.

I think that mathematics is no different from other disciplines in the development of exceptional performance. Only child prodigies, who are one in a million, are born mathematicians, but most who become expert mathematicians deliberately work many years to develop their skills. As Sternberg (1998) postulated, abilities are forms of expertise, so giftedness is developing an expertise, whether in the form of analysis, of creativity or of domain expertise and whether in mathematics or in other domains.

#### Significance of the Study

The questions the author seeks to answer in this study are of two kinds. One group of questions pertains to theoretical issues in the assessment of mathematical ability to identify mathematically gifted students. The author proposes to assess mathematical ability from a multilateral point of view. Unilateral practices for the assessment of mathematical giftedness are criticized from a multilateral point of view. A multilateral practice involves the assessment of several essential facets of ability. The three-mathematical minds model ( $M^3$ ) postulated by the author is an alternative to unilateral practices to assess the analytical, creative and knowledge aspects of mathematical ability. The second group of questions is related to item development that is used to measure different types of ability at different levels. The three-level cognitive complexity model ( $C^3$ ) offered by the author is a way to assess mathematical ability at different levels. Furthermore, the author offers to use differential item functioning to validate item

constructs and to develop theory-driven ability tests as a supplement to the traditional differential item functioning practices. Note that any problems to be discussed below pertain directly or indirectly to the need for the three-mathematical minds model or the three-level cognitive complexity model.

### *Theoretical Concerns*

One central paradigm in the assessment of human abilities is that what is measured by an ability test must be of essential use in the domain of knowledge if the ability test is to have educational value. Sternberg (1996), for example, stated that performance in mathematics courses and on ability tests usually does not predict effectively who succeeds as a mathematician. The prediction failure of many ability tests and school grades is in part due to the fact that they often measure only some aspects of mathematical ability such as analytical ability, memory or mathematical knowledge. As Sternberg asserted, someone can get away with good memory and analytical skills until one attains the highest level of education. By the same token, Poincare (1952b) believed that many people are thought gifted because of their great mental power for recall; however, they are not gifted in the real sense of productive mathematics because they lack the ability to apprehend the harmonious structure of mathematics. Here, I do not blame assessment practices that have been undertaken by some researchers, but I believe that the problem remains unsolved to some extent. The main source of this enduring problem does not result totally from assessment practices, but from narrowly-defined theoretical underpinnings of these practices. In other words, some researchers have conceptualized mathematical ability in only one way as a unified construct that could be

applied in all branches of mathematics, while they have overlooked its many aspects in their assessment practices.

From this author's vantage point, one theoretical solution for a comprehensive assessment of domain abilities is to view a domain from multiple angles; that is, what cornerstones are those on which the domain is built. The extensive review of the empirical and theoretical literature by this author indicated that there indeed exist cornerstone abilities that play crucial roles in production, reproduction and in problem solving in many knowledge domains. The three-mathematical minds model to be discussed extensively in the next chapter is the effort of the author to integrate psychological, philosophical and mathematical models to assess mathematical ability multilaterally, which includes analytical mathematical ability as a source of analytical minds, creative mathematical ability as a source of creative minds, and mathematical knowledge as a source of knowledge expert minds. The model, based on studies about expertise, on ideas of mathematicians about mathematical ability, and on the theory of successful intelligence (Sternberg, 1997), is an instrument for developing theory-driven tests of mathematical ability to assess students' three primary cognitive abilities for production, reproduction and problem solving in the domain of mathematics.

At this point in the discussion, I should pay some attention to the term "*mathematician*" to pinpoint intellectual tools of mathematicians and to explicate the need for multilateral assessment. Poincare (1952b), who was a genius in mathematics and a philosopher of mathematical reasoning, asserted that the abilities to store and recall information do not necessarily make a person a real mathematician. A mathematician is

the one who can discover mathematical rules and make useful constructions with mathematical entities. The importance of mathematical ability, in this creative form possessed by real mathematicians, was enunciated by the greatest mathematicians. Polya (1954a) extensively quotes such statements from Descartes, Gauss, Kepler, Laplace, and Poincare. These mathematicians stated that creative ability is of the highest importance to a mathematician because discoveries are of the utmost importance for the advancement of mathematics and discoveries are the result of creative ability. In the course of discoveries, mathematicians usually start with making conjectures based upon a feeling of intuition or a priori synthetic judgment (Poincare, 1958). Then, they *induce* mathematical orders or rules, make *generalizations* and *specializations* through demonstrations, transfer rules from known problems to unknown problems through analogy, and end up with *reasonable* conclusions. At the final step, these new conclusions are verified through *mathematical deductions* (Poincare, 1952a; Polya, 1954a, b). This course of discovery involves the ability to think in a flexible fashion. Flexible thinking necessitates the ability to shift cognitive functioning in multiple directions breaking through cognitive blocks and *restructuring thinking* so that a problem is viewed from multiple perspectives.

If creativity is central to the advancement of mathematics, then a test should differentiate between those who will be creative mathematicians and those who will not. An integration of mathematicians' ideas in the assessment of mathematical ability can make an invaluable contribution to the identification of mathematically gifted students who will be creative mathematicians. One way to achieve this integration is to take a

mathematical approach combined with psychometric and information processing approaches. The mathematical approach can help researchers understand how real mathematicians solve complex problems and how they go through discoveries. Psychometric and information processing approaches can guide researchers to develop scientific methods to measure mathematical ability possessed by mathematicians. For example, mathematicians of the productive kind extensively use induction, deduction, analogy and selection. They need extensive mathematical domain knowledge while they need few computational skills or little memory because computation or memory does not make productions (Ironically, technology already has taken the part of computational skills).

An ability, such as creativity, to be measured by some tests is not neutral with respect to whether the level of a person's ability is a result of inherited characteristics, of learning or of a combination of them. Carroll (1996) articulated that the estimate of an individual's level of ability in terms of some tasks is only a documentation of the individual's capability at a given time to perform these tasks. However, it does not say much about how ability has developed or can develop through learning over time. For this reason, and because knowledge and ability often interact in superior cognitive performance (Weisberg, 1999), an assessment of cognitive abilities also should measure the current learning level – or factual domain knowledge – of an individual in a domain. This type of assessment, in turn, provides a more comprehensive evaluation of an individual's retrospective and prospective intellectual performance in the domain. The preponderance of research (Weisberg) also reveals that domain knowledge is associated

significantly with creative achievement. Weisberg further asserted that domain knowledge is the only factor that accounts for creative achievement after statistically controlling for intellectual and personality attributes. This finding, especially, is true for the domain of mathematics because it is very abstract in nature and usually is acquired by schooling.

Researchers who have studied expertise have reported key findings about how domain knowledge is acquired by schooling and by deliberate practices, and how expertise should be measured as acquisition of domain-specific knowledge in students. One of the key findings is that experts have different knowledge structures from novices (Chi., Glaser., & Farr, 1988). In other words, experts differ from novices in the ways they store, recall and use information, not necessarily in the strength of their innate abilities (Ericsson, 2003). Further, researchers who empirically studied students reported that teaching, training or deliberate practices improved students' cognitive abilities, including induction, deduction, domain knowledge and insight (Ansburg & Hill, 2003; Dollinger, Levin, Robinson, 1991; Ferrara, Brown & Campione, 1986; Gray, Pinto, Pitta & Tall, 1999; Klauer, Meiser & Naumer, 2000; Miyazaki, 2000; Vartanian, Martindale & Kwiatkowski, 2003). As I pointed out in the foregoing discussion, these abilities also should be the core elements in the assessment of mathematical ability because the major purpose of the assessment is to inform teaching-learning and thinking practices. Studies on expertise have important implications not only for instruction but also for assessments related to developing proclivities, abilities and competencies. Because studies on expertise are rather new compared to those on other abilities, few researchers have

studied both expertise and abilities together, and few expertise theorists have integrated tests of abilities in their research. Likewise, few abilities theorists have used tests of expertise in their research. Sternberg and Grigorenko (2003) proposed that insufficient communication between these two areas of human studies brought about a lack of comprehensive accounts on how abilities, competencies and expertise relate to each other. Because expertise is more related to the acquisition of domain knowledge and to its representation in long term memory, assessment of mathematical ability also should contain tests that measure developing expertise knowledge, as well as tests that measure analytical and creative abilities.

### *Methodological Concerns*

*Factorial level concerns.* Researchers who have used factor analysis to examine intelligence often have employed subtests to measure a variety of cognitive processes in domains of intelligence, such as verbal or quantitative, but not in process domains, such as analytical or creative. This type of investigation usually yields a second stratum factor, such as quantitative reasoning; that is, mathematical ability was found as an aspect of general ability (Carroll, 1996). By the same token, researchers who have used factor analysis to investigate mathematical ability usually included subtests that measured a mixture of mathematical reasoning and knowledge. This type of investigation, on the other hand, resulted in a measurement of general mathematical ability. Therefore, like the theory of general intelligence “g,” a general theory of mathematical ability has prevailed in the assessment of mathematical ability. I call this overarching mathematical ability by the capital letter “*M*,” and other aspects of mathematical ability, such as

creative mathematical ability, with “*Mc.*” Indeed, this notion of general mathematical ability might be psychologically credible. However, mathematical ability deserves more research to find out whether an assessment of multilateral facets of mathematical ability is more promising as a way to identify mathematically gifted students.

*Item level concerns.* An assessment model should be validated not only on the factorial level but also on the item level. The problem to be discussed here is whether an item functions differently in the assessment of different types of mathematical ability, and whether differential item functioning provides evidence for the validity of the three-mathematical minds model. The point of discussion, thus, is related to item validity and homogeneity. Carroll (1996) maintained that one of the major problems of current item development practices is to determine the homogeneity of items. Item Response Theory researchers (IRT, [Embretson & McCollam, 2000]), for example, assume that all items in a test are homogeneous in the sense that they measure the same ability. Item homogeneity is subjectively judged by the test constructor in the beginning of item development; that is, the test constructor develops a series of items, similar in format and content but that vary in difficulty, and that measure the same ability. Because of no initial item validation, the homogeneity becomes a serious matter; thus, a test might measure unintended abilities, in addition to those abilities it is developed to measure. What needs to be done at this point of analysis is a kind of item construct validity. In nature, this validity looks like the general notion of test construct validity, but should be applied at the item level. The rationale behind my assumption is that every item is developed to measure a construct, just as every test is. One way to accomplish such an

item validation is to carry out further analysis of item discriminations and difficulties based on different ability groups. This examination provides information about whether an item discriminates between those who are high in one type of ability and those who are high in another type of ability. I would call such a situation “*functional deviation*,” if items differentiate between unintended abilities, and such items “*functionally deviated items*,” showing functional deviations. Items of functional deviation can be said to lack construct validity (especially divergent validity) because they measure some abilities they are not developed to measure, as well as abilities they are developed to measure. Other items that measure what they need to measure could be called “*functionally fitted items*” and the situation “*functional fitness*.”

Differential Item Functioning (DIF) has been used to analyze items demonstrating different functions for different groups of individuals. DIF analysis often has been applied to members of such groups as gender, race, ethnicity, region, socioeconomic and second language learners (Linn, 1993). This mode of analysis also can shed light on practices in the assessment of mathematical ability if applied to members of high ability, average ability, high-average ability and low ability in mathematical ability in general as well as in mathematical domain knowledge, analytical mathematical ability and creative mathematical ability in particular. In turn, this mode of analysis also indicates whether the item measures only the intended ability or some other abilities, as well. For example, an item developed to measure analytical ability also can discriminate highly between high and average domain knowledge if it requires some domain knowledge.

Application of DIF to different types of ability is different from the conventional approach, in which the composite score often is used to estimate item discrimination and difficulty indexes (Anastasi & Urbina, 1997), which usually results in overlooking low discriminative items that might be good measures of different aspects of ability if factor scores are used as the base. The point is that one item can have a low discrimination level for an ability such as analytical mathematical ability while the same item can be a good measure of another ability such as creative mathematical ability, depending upon which factor score is used as the base on which the item discrimination index is estimated. The use of different factors as the base can provide evidence about whether an item is a good measure of the ability under measurement. For example, one item can discriminate very well between high analytical ability and low analytical ability; however, the same item does not necessarily discriminate between high creative ability and low creative ability. Indeed, the same item can discriminate against high creative ability because the type of psychological construct the item measures can be very different from that to measure creative ability. Therefore, test developers should disclose what each item measures and the purpose of each item should be aligned with the purpose of the assessment. The point is that if the purpose of an assessment tool is to identify analytically gifted students, then each item must be scrutinized to determine whether it also discriminates between those who are high in other abilities.

*Item cognitive complexity.* As I partially enunciated problems surrounding item homogeneity in the foregoing discussion, another issue is why an item is difficult and what makes an item difficult for some individuals. Sternberg (2002) stated that sources

of item difficulty must be developed psychologically through a systematic approach. Likewise, Lohman (2000) maintained that a good psychometric test is the collection of items of different difficulty levels. From a psychological point of view, an assessment model, such as the one undertaken by the author, should be informative about what it is that makes some items more difficult than others. The author's hypothesis is that item difficulty varies as a result of the performance levels of individuals as demonstrated in a domain. Performance level in a domain means that item difficulty comes from two sources, respectively. The very first one is the function of the content of an item – content domain – corresponding to an intellectual domain. For example, a problem of the analytical mathematical kind can be very difficult for a group of individuals who have strong creative mathematical ability but have weak or average analytical ability, or vice versa; thus, the type of ability plays the major role in item difficulty. The second source is the function of the cognitive complexity level of an item. For example, calculus problems can be too advanced for average middle school students because these students do not have that level of knowledge, but it is appropriate for high school students; thus, the level of an item plays the major role in item difficulty. Further, a few problems that are constructed at significantly different difficulty levels are enough to measure a single construct.

One theoretical and methodological model for item development is cognitive complexity, an approach that includes both performance level and performance domain. Different models of complexity approaches exist. According to levels of cognitive complexity models, for example, thinking can be measured at different levels (McDaniel

& Lawrence, 1990) and according to the learning-development approaches, learning and cognitive development can be measured at different qualitative levels (Biggs & Collins, 1982; Pegg, 2003). However, the most current theories and research on cognitive complexity are concerned with the quality or complexity of students' responses to problems, not that of the item itself. The approach for the development of item complexity the author undertakes in this study is the development of items according to performance level as a cultivation of experience in mathematics, such as novice as first level, developing experts as second level and experts as the third level. In this model, titled three-level cognitive complexity, each item measures a specific level of a particular ability because of the level of cognitive demand the item poses to the problem solver. Cognitive demands are established on some psychological sources, such as demands for knowledge, or for analytical ability or for creative ability. While an individual can be at the third level in analytical ability, the same person can be at the first or second level in other abilities. Further, if a test, such as the one undertaken by the author, measures three different abilities, such as three mathematical minds, at three complexity levels, a 3 x 3 profile of an individual can be obtained from the performance of that particular person on the test. Table 1 shows how this comparison can be made by standings of individuals A, B, and C according to the combination of the three-mathematical minds with the three-level cognitive complexity model.

Table 1.1

Integration of the three-mathematical minds and the three-level cognitive complexity model for the assessment of mathematical ability to identify mathematically gifted students

Cognitive Complexity	Types of Minds		
	Analytical	Creative	Expert
Level 1	A	C	B
Level 2	B	A	C
Level 3	C	B	A

### Purpose

The primary purpose of this study was to investigate the psychological validity of the three-mathematical minds model for the assessment of mathematical ability to identify mathematically gifted students. The secondary purpose was to examine the psychological validity of the three-level cognitive complexity model for the development of psychologically-constructed test items.

### Research Questions and Hypotheses

The first three questions and associated hypotheses are related to the three mathematical-minds model. The last two questions and their associated hypotheses are related to the three-level cognitive complexity model.

1. How theoretically valid is the three-mathematical-minds model (M<sup>3</sup>)?

- a. What is the underlying structure of mathematical ability?
  - b. How are the components of mathematical minds hypothetically constructed as mathematical expertise, analytical mathematical ability, and creative mathematical ability associated?
  - c. How are subcomponents hypothetically constructed as knowledge of algebra, of geometry and of statistics; linear syllogism, conditional syllogism and categorical syllogism; and as induction, insight, and selective attention associated?
2. What are the psychometric properties of the M<sup>3</sup> test battery?
  - a. How reliable is the M<sup>3</sup>?
  - b. What convergent validity does the M<sup>3</sup> have when students' liking of mathematics, their rating of their own mathematical ability and teachers' rating of students' mathematical ability are used as converging variables?
  - c. Does the M<sup>3</sup> differentiate between students of various grade levels?
  - d. How valid is the internal consistency of the M<sup>3</sup> for item-total score, item-subtest and subtest-total score correlations?
3. Which M<sup>3</sup> items are good measures of mathematical knowledge, analytical mathematical ability and creative mathematical ability?
4. How psychologically valid is the three-level cognitive complexity model?
  - a. What relations, if any, exist among item cognitive complexity (ICC), item difficulty (ID) and item discrimination (D)?
  - b. Hypothesis – ICC significantly predicts ID.

5. How do the three ability groups, gifted (above 95%), above average (85-94%) and average and below-average (below 85%) as identified by the composite score, differ in their performance on the items at different levels of cognitive complexity?
- a. Null hypothesis– No significant difference exists between the performance of gifted students and that of above average students on items at the third level of cognitive complexity only.
  - b. Null hypothesis – No significant difference exists between the performance of above average students and that of average and below average students on items at the second and third levels of the cognitive complexity.
  - c. Null hypothesis – No significant difference exists among the performance of the three ability groups on items at the first level of the cognitive complexity.

## CHAPTER II

### LITERATURE REVIEW

This chapter is an overview of theories and research about the nature of mathematical ability in general and of mathematical giftedness in particular. In the first part, I will discuss the nature of mathematical ability, mostly referring to mathematicians' ideas about mathematical reasoning to set the theoretical ground for this study. In the second part, I will discuss different mathematical minds as a way to study and assess mathematical ability and to identify mathematically gifted students. The third part includes a discussion of psychological and neuropsychological theories and research about analytical ability, creativity and expertise to provide support to the study of different mathematical minds. Finally, at the end of the third part, I will propose the three-mathematical minds model ( $M^3$ ) for the assessment of mathematical ability.

Mathematical ability can be studied from multiple perspectives including branches of mathematics as a discipline of knowledge and cognitive processes of mathematical ability as a discipline of thought. The road a researcher takes to study mathematical ability often is influenced by the paradigms of the discipline in which the researcher works. I find three disciplines essential for the study of this ability: philosophy, psychology and mathematics. My attempt will be to integrate these disciplines to study mathematical ability. I find no restrictions but do find directions and positions about how to study mathematical ability. Others can choose other ways depending on their convictions.

## Part I

### The Nature of Mathematics

Mathematics is intriguing, in that researchers from a variety of disciplines have been keen on the study of mathematics. The involvement of a wide range of disciplines in the study of mathematics, the viewing of mathematics from multiple angles, has brought about a rich and wide body of knowledge. Researchers from any discipline who attempt to study mathematical ability, however, are expected to ask how to define the domain of mathematics, how knowledge is produced in this domain, and what tools mathematicians use to produce knowledge. I will use Shaw's classification (1918) to frame the points of my discussion about the nature of mathematics while enriching my argument with the teachings of the mathematicians Poincare and Polya about mathematical ability. According to Shaw, the study of mathematics can be classified in many ways such as the content of mathematics, the central principles of mathematics, and the methods of mathematics. The scope of my discussion will be the central principles that define the structure of mathematics and the methods that characterize the processes of knowledge production and problem solving.

First, it is essential to pinpoint the subject-matter of mathematics to understand how the principles and methods of mathematics apply to the subject matter. The content of mathematics includes static mathematics and dynamic mathematics. Static mathematics is composed of numbers leading to arithmetic, of figures leading to geometry, of arrangements leading to tactic, and of propositions leading to logic. These divisions are of the static kind because any objects being studied in this way are

given fixed entities or some collection of entities. Dynamic mathematics, on the other hand, includes operators leading to operational calculus, hypernumbers leading to algebra, processes leading to transmutations, and systems leading to general inference. These divisions are of the dynamic kind because objects of this type are transitions rather than states.

### *Central Principles of Mathematics*

In each subject-matter of mathematics, Shaw (1918) argued that four central principles of mathematics appear: forms, invariance, functionality, and inversion. Form or structure is the particular character or property of constructions. Kempe (as cited in Shaw, 1918) claimed that the study of mathematical properties of any subject-matter is only a study of form. Forms appear in numbers, figures, arrangements or any other constructions. Regularity and harmony, for example, are important characteristics of mathematics and can be seen in numerical or geometric constructions. Consider, for example, natural numbers. Every even number is followed by an odd number: 1, 2, 3, 4, 5... and vice versa. There is more in this regularity. Every number is increased by one that is  $n + 1$ . This simple example indicates to us some evidence of how mathematical entities are ordered by some rules, which, indeed, are particular characteristics of natural numbers. The rule is  $n + 1$  in the above example; so, when an individual studies natural numbers, s/he studies the form of natural numbers composed of particular features or rules.

How is regularity related to the other principles of mathematics? Mathematics is far more than simple facts; rather, it is structured on relations between mathematical facts

and on relations between relations. Hardy (1940) asserted mathematics is much deeper than what nonmathematicians think. According to Hardy, a mathematician works like a painter or a poet. A painter makes patterns with color, a poet makes patterns with words, and a mathematician makes patterns with ideas (pp.25):

The mathematician's patterns, like the painter's and the poet's,  
must be beautiful; the ideas, like the colors or the words, must fit together  
in a harmonious way. Beauty is the first test: There is no permanent place  
in the world for ugly mathematics.

While the structure or the form defines the particular character of constructions, such as natural numbers, invariance is the common characteristic of any class. The structure and invariance can be observed not only in relations between mathematical entities that form a class, but also between relations of relations, showing functionality. That is, forms of each kind correspond to each other in one to one, one to many, and many to many ways. A relation, for example, can be analogous to another relation. Consider this example on relations of relations. In mathematics, something always is analogous to another that is higher or simpler in structure. A square is analogous to a cube in that the relationship between them is that a cube is composed of six squares in surface; and the area of a square ( $n$ ) is analogous to that of a cube, that is  $6n$ . Now, consider a triangle and a prism. The relationship between a square and a cube is exactly analogous to that of the relationship between a triangle and a prism. A prism consists of four triangles in surface, and the area of a prism is that of four times a triangle, that is  $4n$ .

### *Methods of Mathematics*

Methods mathematicians use to carry out investigations are of four kinds. The first is the scientific method from which mathematicians borrowed methods of observation, experimentation, analysis and generalization. Sylvester (as quoted in Shaw, 1918, pp. 169), for example, stated that most of the great ideas of modern mathematics emanated from observations. The arithmetical theory of forms, for instance, is rooted in observations of such geniuses as Euler and Jacobi. Likewise, mathematicians use generalizations, just like natural scientists do, to apply an idea or an equation to solve a variety of problems. Poincare's discovery of Fuchsian functions (1958) is a good example for mathematical generalizations in that Fuchsian functions could be used to solve differential equations and algebraic equations, and to express coordinates of algebraic curves. Generalization is crucial in mathematics in that mathematics proceeds from the particular to the general (Poincare, 1952a). According to Hardy (1940), a serious mathematical idea or a theorem should be general that can be applied in many mathematical constructs. The theorem should be able to be extended and be typical of other theorems of its kind. Therefore, mathematicians always attempt to generalize propositions they have obtained, such as from the particular instance  $a + 1 = 1 + a$ , to a more general statement  $a + b = b + a$ . The process is proof by recurrence according to Poincare such that we show that a theorem is true for  $n = 1$ ; then if it is true for  $n - 1$ , it is true for  $n$ . Then we conclude that it is true for all integers.

The second method mathematicians use to carry out investigations is the intuitional method. This method, very often criticized by strict logicians, is an *a priori*

synthetic from Poincare's perspective (1952b). In essence, a person immediately discerns what the problem is and what missing elements are crucial for the solution. For example, some people say, aha, I know what the answer is," when they are given a perplexing or puzzle-like problem. They often are not able to explain where or how the answer pops into their heads, but they say, "I just know it." Therefore, the intuitional findings or solutions do not come about through some reasoning processes, such as dissecting, comparing, contrasting, or relating; instead, they occur through gestalts. Intuitive people usually think with pictures, diagrams, or other visuals. Indeed, Poincare called creative mathematicians geometers, postulating that creative mathematicians are those who work with geometry, use tables, diagrams and other visuals in their work. Poincare further argued that it is almost impossible to find even one diagram in the work of many mathematicians considered uncreative by Poincare. Furthermore, intuition is not restricted to the visual representation of problems, according to Shaw; rather, intuition expands the capability for insightful thinking. Riemann, for example, connected the deformation of surfaces and the theory of algebraic functions through insights just like the insight of Poincare who argued the curves by an intuitive study. Shaw (1918, pp.175) pointed out the importance of intuition in mathematics by saying that:

It is the intuition method that enables mathematicians to pass in the direction just opposite to that of logic, namely, from the particular to the general. It is primarily a method of discovery and often starting from a few particular cases is able to see in them theorems that are universally true. It must be accompanied by a keen power of analysis and ready

perception of what is essential. It often happens that hasty generalizations would lead to results that are valid for many cases, for the analytical power must be very keen.

Another method mathematicians always use in problem solving is called the “deductive method.” Because all mathematicians have to use deduction to prove their postulates, mathematics has come to be viewed as a deductive science. In the deductive method, individuals deduce conclusions from given premises or information. The conclusion is drawn from the general to the particular. According to Shaw (1918), deduction is used in two ways: first, it is the method of exposition of results, and secondly, it is the method of verification. The deductive method, however, usually is used as a means of verifying theorems or confirming conjectures.

Although the deductive method is a tool for certainty, it is not a tool for creativity. Consider this syllogism. All humans are mortal. Alice is human so Alice is mortal. The first sentence is a true statement, a premise. The second sentence is a reality and a valid conclusion. Nothing is new in this statement. Further, consider the following syllogism. All prime numbers are dividable only by themselves and by one. Seven is a prime number; therefore, seven is dividable only by itself and by one. Here, an axiom or a theorem exists about prime numbers that was postulated by its discoverer. The logic we followed to prove this theorem is the product of deduction. We did not discover something new or did not add anything to the original theory. Therefore, the most essential use of the deductive method is to verify our knowledge. A new theory always is

in need of such verifications because it often is discovered through induction or intuition, or through our common sense, whose products are not always correct.

The fourth method, but perhaps not the last, is the creative method. As enunciated before, mathematics usually is construed as a demonstrative discipline of science that relies on certainty or only proofs by pure deduction. However, deduction is only one aspect of mathematics (Poincare, 1952a; Polya, 1954a). Polya maintained that mathematics has two facets. One is the deductive facet that is the rigorous side of mathematics. The second facet is induction that is related to inventions in mathematics.

In the making of knowledge, mathematics is no different from other disciplines. Consider this general process of a discovery: A mathematician has to make conjectures about a mathematical theorem before he proves it. Successively, he conjectures the method of proof before he works out the details. He has to combine observations or information selectively and use analogies. To Descartes every problem he solved became a rule that served as an analogy to solve other problems (Polya, 1962). He enunciated that any new mathematical truths he discovered depended on a few principal problems. What happens at the end of a mathematician's discovery or any creative work is generalization; that is, a mathematical discovery should be applicable to a wide range of domains, which is the actual discovery. Because I will discuss mathematical inventions in detail in the following sections, I shall end my discussion of the creative method with what Shaw (1918) said about it. According to Shaw, the stamp of the great mathematician is what he creates as a new set of mathematical entities. These entities usually arise as a response to the need for solutions that are applicable in many domains.

## The Nature of Mathematical Ability

### *Conceptions of Mathematical Ability*

Conceptions of mathematical ability vary among scholars who have been involved in the psychological and philosophical study of mathematics. No doubt a psychological study of human abilities has had much influence on the conceptions because the controversy of “can do” or “has done” (potential vs competence) is seen in these conceptions. For example, Thurstone (1950) defined ability as what an individual *can do* while Werdelin (1958) defined it as what an individual *has done* during the measurement. Carroll (1996) suggested that the definition of ability must be related to variations over individuals and their threshold levels of difficulty in successfully performing some defined class of tasks. According to this suggestion, ability is highlighted by the tasks used in the measurement.

Conceptual variations among psychologists come partially from different weights attached to the different facets of mathematical ability. These conceptions can be at the practical or theoretical level. In these conceptions, both the aspects of mathematics as a domain of knowledge and the aspects of mathematical thinking as a discipline of thought can be seen. Therefore, the conceptions may vary depending on beliefs about the nature of mathematics as a knowledge domain or as a thought domain. For example, Griffin (2000) considered math as a set of conceptual relations between quantities and numerical symbols. In this definition, the branch of numbers is emphasized as a crucial aspect or branch of mathematics and the definition is rather theoretical.

Thomas (the definition is cited in Werdelin, 1958) differentiated between different aspects of mathematical ability at a practical level. He emphasized abstraction, logical reasoning, spatial perception, and intuitive power, as well as the ability to use formulas, mathematical imagination, and the ability to construct mathematical gestalts. Note that Thomas' definition includes thought processes. Cameron's definition (1925), however, is more practical. He defined the most essential facets of mathematical ability as the power of analysis of combinations and reconstructions of its elements in a new way, as the power of comparison and classification of numerical and spatial data, as the power of concrete imagination and facility in mechanical operations, and as the ability to apply general principles and to manipulate abstract quantities. Notice that Cameron emphasized both processes and branches of mathematics in the definition. Werdelin (1958) also addressed multiple aspects of mathematical ability in the definition at a practical level. He defined mathematical ability as the ability to understand the nature of the mathematical problems, symbols, methods and proofs; to learn them, to retain them in the memory and to reproduce them; to combine them with other problems, symbols, methods and proofs; and to use them when solving mathematical tasks. A contemporary psychologist Howard Gardner (1999) used logic and mathematics together in his conception of mathematical ability as logical-mathematical intelligence. According to him, logical-mathematical intelligence is the capacity to analyze problems of the mathematical kind logically, perform mathematical operations and study problems scientifically. Obviously, this definition is much more general than the others.

Conceptions of mathematicians and of philosophers of mathematics differ from those of psychologists. Logic and intuition are main concepts in mathematicians' conceptions, and power of memory has little importance in mathematical ability. Poincare (1952b), for example, enunciated that it is not a strong memory or attention or mental calculation that makes people mathematicians, but it is the intuition that enables them to feel, to see and to conceive the structure or relations among mathematical entities. Poincare believed that this ability does not belong to everyone, but to those who are in a condition of discovery. Poincare stated that people with great memory and attention and the capacity for analysis also can be gifted in mathematics. They can learn every detail of mathematics, but they lack the ability to create or to discover.

Another type of mathematical mind is the mental calculator who usually is uneducated, but can make very complicated calculations very quickly (Hadamard, 1954). According to Hadamard, such talent is very different from mathematical ability, and only a few eminent mathematicians possessed such talent. Hadamard asserted that exceptional calculators can show remarkable characteristics. They carry out remarkable calculations without willful effort that is activated in their unconscious. Ferrol, for example, was able to do such complex calculations. Ferrol's statements indicated that answers to problems came to his mind suddenly as if someone had whispered in his ear. However, Ferrol is an exceptional case, and he was good at algebra as Hadamard said. According to Poincare, however, mental calculation is not a characteristic of mathematicians of the productive kind. Many known mathematicians, including Poincare himself, were not good

calculators. Indeed, Poincare confessed that he forgot the multiplication tables many times.

In addition, Polya (1954a, 1954b) proposed two kinds of reasoning underlying mathematical ability. One is *demonstrative reasoning*, usually called deductive reasoning or logic. This mode of thinking belongs to analysts. The principal function of demonstrative reasoning is to distinguish a proof from a conjecture or a valid argument from an invalid argument; thus, demonstrative reasoning ensures certainty in mathematics. The other type of reasoning is *plausible reasoning*, usually called inductive reasoning in psychology; though plausible reasoning is a more comprehensive phenomenon that includes induction, analogy, and other similar constructs. Polya considered these particular reasoning tools as particular cases of plausible reasoning or the entire reasoning process. The primary function of plausible reasoning is to differentiate a more reasonable conjecture from a less reasonable conjecture by providing logical evidence. This mode of thinking belongs to creators. The two types of reasoning are not in a polar fashion, but they complete each other in mathematical problem solving.

#### *Factorial Structure of Mathematical Ability*

Factor analytic studies of mathematical ability are as old as the history of factor analysis itself. In this section, I briefly discuss what factor analyses of cognitive abilities have revealed about the nature of mathematical ability. Mathematical ability takes many forms, depending on the nature of the mathematical task. Mathematical tasks also vary, depending on the branches of mathematics such as arithmetic, algebra, geometry,

numbers or statistics, and on the cognitive processes such as induction, deduction or computation.

One of the earliest quantitative investigations of mathematical ability, before factor analytic methods, is the one by Rogers (1918). Rogers constructed a variety of tests, which covered many branches and processes in mathematics as well as productive and reproductive aspects of mathematics, to examine different aspects of mathematical ability. The mode of the analysis was correlation coefficients. The coefficients of correlation ranged from .01 to .59. After further analysis of combined test scores, Rogers concluded that geometric, algebraic and even verbal abilities were equally important to mathematical ability. She maintained that mathematical ability is a complex confluence of a number of loosely connected capacities. However, Spearman's and Werdelin's reanalysis of Rogers' data, using factor analysis, (as reviewed in Werdelin, 1958) revealed a different picture. In the reanalysis, a geometric, an algebraic (mostly numerical) and a verbal factor were found. Note that all factors are related to the domain of knowledge, not to that of cognitive processes.

Following the advent of factor analysis, studies of mathematical ability seemed to be more informative. At least four or five factors underlying mathematical ability were found repeatedly. A numerical factor was found in major factor analytic studies of mathematical ability (Spearman, 1927; Thurstone, 1938; Werdelin, 1958) consisting mostly of addition, multiplication and other arithmetical problems. Spearman stated that the numerical factor was one factor common to arithmetic over and above g. Another common factor found was related to visual or spatial tasks that required the manipulation

of objects in two and three dimensional space, implying a connection between the ability to visualize and geometric ability. However, the relationship between the numerical factor and the spatial factor is controversial in that the correlations between these two factors were found positive in some studies and negative in others (Werdelin). Other common factors in factor analytic studies related to mathematical ability were the reasoning factors, consisting mostly of induction (number series) and deduction tasks (syllogisms) (Thurstone, 1938; Werdelin, 1958). Werdelin also found a deduction factor separate from the general reasoning factor underlying mathematical ability. Other factors, such as verbal and scholastic factors reflecting school grades and achievement tests, also were found in these studies.

Carroll (1993) reanalyzed four hundred eighty studies using exploratory factor analysis. Many of the studies also had datasets relevant to mathematical ability. His reanalysis indicated a hierarchical structure in cognitive abilities, similar to other theories proposed earlier (Cattell, 1971). Carroll suggested a three-stratum theory whereby cognitive abilities can be classified hierarchically in terms of their generality into general, broad and narrow factors. At the highest level is the general ability (g); broad abilities are located in the second stratum, such as crystallized and fluid intelligence, and narrow factors are in the first stratum. In his reanalysis, Carroll found quantitative ability under the second-stratum factor, fluid intelligence. No unique mathematical ability existed according to the reanalysis. However, some first-level factors were found. These first level factors were general sequential reasoning, quantitative reasoning and induction. However, Carroll (1996) suggested that fluid intelligence was related to mathematical

ability because most reasoning activities under the fluid intelligence were associated with logical and quantitative concepts. Further, he concluded that fluid intelligence could be estimated separately from “g” because fluid intelligence did not correlate perfectly with “g” in many analyses. Although Carroll never suggested the existence of a unique mathematical ability, his conclusion implicitly communicates that what usually distinguishes fluid intelligence from other factors is the existence of many mathematical tasks under fluid intelligence factor.

Factor analytic studies have enlightened researchers about many aspects of mathematical ability. However, as happens in many statistical analyses, many findings may be artificial, depending on the number, type and difficulty of tasks. Another drawback of most factor analytic studies is how the reasoning tasks were different from the knowledge tasks used in prior studies were not clear. Most tasks used in these studies seem to measure aspects of analytical mathematical ability, leaving out essential tasks underlying mathematical creativity. Mathematical creativity never was mentioned with the exception of Sternberg’s study, which indicated that separate analytical and creative abilities existed in mathematics (Sternberg, 2002). However, in Carroll’s reanalysis (1993) under the second-stratum broad retrieval ability, first level factors representing some aspects of creativity, such as originality and fluency, were found. Nonetheless, these factors were conceptualized as retrieval ability, far from referring to mathematical creativity. Therefore, processes and branches in mathematics deserve more factorial investigations with a focus on aspects of mathematical creativity, along with other mathematical reasoning components. As a final point of this section, I would like to

express that although factor analytic studies help understand the psychological structure of mathematical ability, they are not a tool for understanding how mathematicians solve problems and go through discoveries. A deeper understanding can be gained from mathematicians themselves about their ability.

*Mathematical Ability from the Mathematician's Point of View*

The focus of this part of the chapter is on the most essential components of mathematical ability from the vantage points of Poincare and Polya. Particularly important is the argument I will carry out about how creative mathematical problem solving is achieved using inductive and insight tools; therefore, the discussion starts with mathematical induction and ends with insightful thinking.

*Mathematical induction.* Polya (1954a) defined mathematical induction as a process of discovering general laws by observation and combination of particular instances. According to Polya, inductive reasoning is a component of *plausible reasoning*. In very general terms, plausible reasoning implies a series of inductive processes followed by demonstrative phases. Induction is carried out mostly by conjectures. Polya maintained that knowledge is produced primarily by conjectures. Some conjectures are highly credible and reliable, such as those in general laws of science, while some conjectures are neither credible nor reliable, such as those in newspapers. While mathematical knowledge is ensured by demonstrative reasoning, we strengthen our conjectures by plausible reasoning. A mathematical proof, for example, is demonstrative reasoning, but a weather broadcast belongs to plausible reasoning because variations might happen in weather estimates.

Let us see first how the process of plausible reasoning is carried out during mathematical problem solving; then, we can realize better the full importance of inductive reasoning as an underpinning to mathematical reasoning. Let us first analyze the solution of a number problem, consisting of several phases as proposed by Polya (1954a). Most people are able to discern properties of integers, such as 1, 2, 3, 4, 5.... They can distinguish odd and even numbers, and know the squares, such as 1, 4, 9, and the primes, such as 2, 3, 5, 7, and 11. Some people may be able to observe intriguing relations between numbers that others do not see.

Consider these equations  $3 + 7 = 10$ ,  $13 + 17 = 30$ , and  $23 + 37 = 60$ . A good mathematical mind discerns the relations between these numbers. The numbers, 3, 7, 13, 17, 23 and 37 all are odd numbers, and a deeper analysis indicates that they also are primes; thus they are odd primes. An analysis of their sums indicates that they all are even numbers (10, 30 and 60). This first step of analysis revealed some relations or similarities between these numbers. A second-step analysis indicated that the equations are analogous to each other in terms of their structural relations. Then, a conjecture can be made as a third step based on the initial analysis of our observation; that is, *the sum of two odd primes is an even number*. Poincare deemed this entire initial process or the discernment of harmonious relations a “prior synthetic judgment,” which is fundamental to mathematical creativity.

The initial attempt yielded a *conjecture*. We still need to go further and further, and find particular cases that verify the conjecture. When we take more even numbers as sums, such as 12, 14, 16 and 18, there still are odd primes that verify the conjecture. An

analysis of the numbers 2 and 4, however, indicates that they cannot be split into a sum of two odd primes; therefore, we need a more specific statement: *Any even number that is neither a prime nor the square of a prime is the sum of two odd primes.* This last step was a generalization derived from the particular examples; namely, the examples proceeded from the particular cases of 3, 7, 13, and 17 to all odd primes, and from 10, 20, and 30 to all even numbers, and then to a general relation: even number = prime + prime.

This statement or conclusion is a conjecture in essence, indicated by particular cases and obtained by induction. It is, by no means, proved yet. Thus, it deserves further investigation. What we have carried out so far is the first stage of the inductive process according to Polya and a prior synthetic judgment followed by an incomplete induction according to Poincare. Now, we shall continue our analysis further to complete a total induction or plausible reasoning.

A systematic examination of even numbers as sums of two primes, for example, from 6 to 100 can provide further support for the conjecture or prior synthetic judgment. A tabulation of all these numbers such as,  $6 = 3 + 3$ ,  $8 = 3 + 5$ ...  $98 = 37 + 61$ ,  $100 = 43 + 57$ , indicates that our conjecture is true for all even numbers from 6 to 100. However, such instances still do not prove the theorem because numbers are limitless. Although a single verification does not prove the theorem, each verification increases confidence in the conjecture by providing further support; that is, each verification renders the conjecture more *credible* and makes it more *plausible* according to Polya.

The entire process discussed above indicates that a conjecture was conceived by prior synthetic judgment from some relations between numbers. Particular instances

were investigated to verify the conjecture; demonstrations were used for other instances, and the conjecture was found true in all cases; as a result a general statement was made. Mathematical discoveries, therefore, involve several phases that include generalization, specialization, analogy, and induction. These processes often are used together in the discovery. The previous example discussed above indicated the analogy of three relations:  $3 + 7 = 10$ ,  $3 + 17 = 20$  and  $13 + 17 = 30$ ; then a generalization from the particular odd numbers to all primes, which were obtained by induction. Finally, a specialization from the conjectural statement to particular cases such as 6 and 8 was verified by deductions.

*Analogy in mathematics.* Laplace enunciated that in mathematical science the principal instruments to discover the truth are induction and analogy (as cited in Polya, 1954a). An essential part of mathematical knowledge is stored in the form of formerly proven theorems. According to Polya, almost no problems are unrelated to formerly solved problems. Indeed, he claimed that if such problems do exist, they would be insoluble. When people encounter a problem, they make use of previously solved problems, using their results or their methods, or the experience they acquired while solving them. However, in the case of problems too ill-structured to choose or to relate to a problem, people usually search for links by analogy. An analogy might mean many similar aspects between objects, which in turn brings the idea of regularity or order in the mathematical world. As stated before, this is a tool for the discoverer in mathematics.

When solving most mathematical problems, people use previous experience to guide their thinking and solutions for problems just as mathematicians do. Use of prior

learning can take several forms. People can recall an appropriate formula from memory, such as that for computing the area of a circle, and apply it to the solution of the new problem. If a proper formula cannot be found, students most likely will try two other options (Novick, 1992). One option is to develop a solution or solution strategies on their own by simply using their general knowledge of mathematical facts and procedures. Another option is to relate the new problem to a specific problem, previously encountered, and to transfer some aspects of the prior problem-solving experience to the new situation. Notice that this is not a mere transfer of prior knowledge; rather, much more thinking exists in this last option, an analogy. However, according to Poincare (1952b), an ordinary analogy does not guide discoveries; rather, the kinds of analogies leading to discoveries are much more deeply hidden. Like Gestalt psychologists, Poincare believed that such analogies involve an uncommon penetration into the problem situation while giving up old habits of thinking.

Now we shall see how the use of analogies can be a major instrument for discoveries. We shall start with an example in solid geometry related to the partition of space on which Polya worked (1954a): Into how many pieces, at most, is space divided by five planes, provided that the planes are in a general position (no two planes are parallel)? At a glance, geometric visualization of all the partitions affected by five planes is too difficult to see, and is more difficult or even impossible if the planes are more than five. In fact, finding the number of partitions even is very difficult if the planes are more than ten. In such circumstances, the main tool of a mathematician is to devise easy analogous problems to develop a model. This method of analogous model development

often is used in mathematical science (Polya, 1954a). Polya articulated that using simpler problems that are similar to the present problem could be a promising analogy. These simpler problems also are accessible to our geometric intuition. Going back to our original problem, a single plane splits a space into two pieces; two planes cut it into a maximum of four pieces; three planes divide it into eight parts and four planes divide it into fifteen parts at most. This underlying similarity can help us to estimate the number of partitions in a space made by five or more planes. However, more analogy exists in spatial geometry.

An analogous relationship exists between points, lines, planes and spaces. This analogy can be used to solve problems related to points on a line, lines on a plane and, lastly, planes on a space. Let's start with the easiest, point-line relation, and then line-plane relation. One point splits 1 line into 2 parts; 2 points divide it into 3 parts; 3 points cut it into 4 parts; 4 points divide it into 5 parts; and 5 points split it into 6 parts; and  $n$  points split it into  $n + 1$  different parts. Now we shall look at lines on planes to continue our analogy one step further. Again, no lines are parallel as in a space-plane problem. A plane is divided by 1 line into 2 parts, by 2 lines into 4 parts, by 3 lines into 7 parts, and by 4 lines into 11 parts. Here, a pattern appears in all cases.

Table 2.1, similar to the one designed by Polya, shows this pattern. This table containing some mathematical regularity is a challenge to creative ability according to Polya. The task is to induce the relation and to discover the rule. Note that we have come to this point, the table, by a variety of little analogies. More discoveries can be made from the table by the use of inductive reasoning and selective combinations of

particular elements. The result is not a simple juxtaposition of elements; rather each combination (number of divisions in the table) communicates some regularity or connection among the elements of the juxtaposition.

Table 2.1

Number of divisions of a space by the number of dividing planes, of a plane by the number of dividing lines, and of a line by the number of dividing points.

---

Number of dividing elements	Divisions of line By points	Divisions of plane By lines	Divisions of space by planes
0	1	1	1
1	2	2	2
2	3	4	4
3	4	7	8
4	5	11	15
5	...	...	...
N	$N + 1$	N	N

---

Our original problem was to search for the number of divisions in a space made by 5 planes. The last column presents regularity: 1, 2, 4, and 8. This regularity is the successive powers of 2. The next term is 15. It does not fit into this regularity. Then, we look at the third column, divisions of plane by lines, as it is analogous to the last column, divisions of space by planes. We induce that every entry in the third column is the sum

of two other entries; that is the number above it and the number to the left of it, such as  $4 + 7 = 11$  (italicized in the table). This pattern continues. After finding the order, we can verify it in more cases. A mathematician often uses recursive reasoning (a branch of mathematical induction) to prove the rule because he cannot work out all numbers, for they are infinite. If we proceed further by recursive reasoning, we can find the rule and develop the formula:  $1 + n + n(n - 1)/2$ , which gives the number of divisions of a plane by lines. We already found that  $1 + n$  is the number of divisions of a line by points.

If we proceed by recursive reasoning and apply it to the last column, under which divisions of space by planes are represented, we also can discover the pattern, the rule and the formula to find the number of divisions of space. For instance, four planes divide space into 15 parts as seen in the table. Our inductive reasoning indicates that 15 is the sum of 8 and 7; that is the entry above 15 and the entry left of 8. When we continue such reasoning, we discover that five planes divide space into 26 parts; that is  $11 + 15 = 26$ . Indeed, if we continue further, we will see that this pattern continues, as well. In sum, we used several branches of mathematical inductive reasoning to solve a problem that initially seemed too challenging to our minds.

*Selection, attention, and insight in mathematics.* According to Polya (1954a) and Poincare (1952b), induction and analogy are major instruments for mathematical discoveries. Also important instruments for discoveries are selection, attention and insight. Poincare (391) proclaimed that, “*discovery is selection.*” Polya underscored “*selective attention*” during a problem solution. Selection occurs, first, in understanding the problem. A good mathematical mind becomes *selective* during problem solving

(Polya, 1954b). A filter can be a good analogy to a selectively working mathematician. S/he filters out what is related to the problem under investigation and what is unrelated to it, and s/he identifies what is known and what is unknown. Then, s/he works with the information that is promising to the solution.

Poincare (1952a) deemed that the ability to discern what information is promising for the solution is a prior synthetic judgment. According to him, a creative mathematician, by intuition, feels the interconnectedness of elements in a situation. What does Poincare really mean by intuition? It is the insight ability to discern, at a glance, harmonious relations between mathematical entities. According to Poincare (1952b), this intuition helps people to see harmonies and relations hidden among juxtaposed elements. Some people do not have this ability of apprehension of relations. He claimed that one could be a mathematical creator without a great power of attention and memory, but he could not be a creator without this feeling of intuition. A mathematical demonstration is not a simple juxtaposition of syllogisms; it consists of syllogisms placed in a certain order, and the order is more important than the elements themselves. Poincare believed that if an individual has the ability to discern this order, s/he does not need to recall all the elements of the juxtaposition because each element places itself in the position prepared for it in the mind of a creator.

Second, selection also occurs in mathematical constructions. A mathematician must work selectively while constructing mathematical combinations. Mathematical discovery is not the making of combinations with mathematical entities that are already known; rather it is constructing new combinations that are useful (Poincare, 1952b). A

real mathematical discoverer has to discern the useful combination among numerous others. In mathematics, the number of samples or cases is so numerous that the entire life of a mathematician would not be enough to examine all samples to prove a conjecture or to make combinations that are promising. What a mathematician really needs to do to discover is to choose among numerous samples or combinations with a view to eliminating those that are useless. According to Poincare, this selective combination must be felt with insight and not be formulated; only then comes proof as a result of insightful thinking.

Another distinguishing feature of mathematical problem solving that leads to discoveries is the phase in which a problem solver questions himself by selecting the direction he takes to solve a problem (Polya, 1954a). In such problems, a problem solver often has to redirect his attention and his strategies selectively, restructure the problem, and sometimes interpret it from very different angles. According to Poincare (1952b) useful combinations come to a creator's mind as a result of a preliminary shift after restructuring the problem. In the same line of thinking, Gestalt psychologists of insightful thinking proposed that solutions to perplexing problems often come after restructuring a problem (Ducker, 1945).

Selection also plays an important role in finding analogies. Promising analogies come from the use of selective working. In other words, a mathematician selectively looks for underlying structures between a target problem and a source problem. For example, the use of dominant functions as an analogy has served to solve numerous problems. Poincare's discovery of Theta-Fuchsian is an example of selective analogy.

He was guided by the analogy with elliptical functions. The remainder was just a justification of his discovery with a series of deductive processes.

## Part II

### Searching for Different Mathematical Minds

From the foregoing discussion, two types of mathematical ability emerged; in other words, two kinds of ability characterize two types of mathematical minds – one that is analytical and one that is creative. In this part, I will discuss these two kinds of mathematical minds in detail. The third mind that I believe is important for understanding and assessing mathematical ability is the expert mind. I will discuss expert minds later. However, I point out at the onset that mathematical expertise, although more related to mathematical knowledge than cognitive processes, might be much more related to mathematical analysis and creativity than I consider. That is to say, analysts and creators differ in their use of cognitive processes while experts differ in their knowledge structure, having specialized knowledge.

#### *Analysts and Creators*

Poincare (1958) thought of two kinds of mathematical minds: *Analysts or logicians* and *Creators or intuitionists*. These two minds work differently and contribute to science differently. Creators make discoveries of theorems. Analysts, on the other hand, usually do microscopic work by analyzing mathematical rules, axioms, combinations or theorems the creator already has discovered. The primary thinking tool of analysts is logic. Syllogisms or deductive reasoning are the particular cases of this logic. The primary thinking tool of creators is mathematical induction by rule discovery,

analogy or mathematical constructions. Let's read this distinction from Poincare's words (pp.197):

It is impossible to study the works of the great mathematicians, without noticing...two entirely different kinds of minds. The one sort are above all preoccupied with logic; to read their works, one is tempted to believe they have advanced only step by step, after the manner of a Vauban who pushes on his trenches against the place besieged, leaving nothing to chance. The other sort are guided by intuition and at the first stroke make quick but sometimes precarious conquests, like bold cavalymen of the advance guard.

According to Poincare (1958), the nature of their minds makes mathematicians either analysts or creators, and this can be seen in the way they approach a novel problem. That is, not only do the two minds work differently, but also the ways these two minds deal with a problem make them different. Analysts approach a problem by their logic whereas creators approach the same problem by their intuition. That is to say, the nature of the problem does not change the nature of thinking of the two minds. Let's read it more precisely from Poincare's language (pp.197):

The method is not imposed by the matter treated. Though one often says of the first that they are analysts and calls the others geometers, that does not prevent the one sort from remaining analysts even when they work at geometry, while the others are still geometers even when they occupy themselves with pure analysis....Nor is it education which has

developed in them one of the two tendencies and stifled the other. The mathematician is born, not made, and it seems he is born a geometer or an analyst.

Further, the former is weak in visualizing space, and the latter is weak in long calculations and soon can become tired of perplexing calculations. Although Poincare believed that a mathematician is born either an analyst or a creator, this notion is very controversial from the psychological point of view, if not from the philosophical standpoint, given the fact that human cognition is prone to so much development if nurtured. For example, Polya (1957, 1962) enunciated that students can learn mathematical problem solving in many ways through purposeful practices, whether analysis or synthesis in nature.

Poincare's contrast (1958) between ancient mathematicians' intuitional approach and contemporary mathematicians' logical approach is even more intriguing. The works of ancient mathematicians indicate that they treated problems by intuition, so they can be classified as intuitionists. However, an analysis of their work shows the work of a logician in their discoveries. According to Poincare, an evolution from the intuitive approach to the analytical approach has emerged. What makes contemporary mathematicians logicians is what the field requires of them; that is to say, the field requires rigorous and certain proofs, which cannot be accomplished only by intuition. In fact, intuition sometimes can deceive people on their hypotheses unless intuition is supported by strict logic or analysis. However, pure logic can lead to tautologies, not to discoveries.

Intuition is not limited to only one kind, and it is not just feeling. Consider these mathematical axioms: 1) Any numbers of quantities that are equal to another quantity also are equal to each other; 2) if a theorem is true for the number 1, and true for  $n + 1$  and for  $n$ , then the theorem can be proven for all whole numbers. The first axiom is a result of formal logic or, so to speak, a microscopic analysis. How about the second axiom? It cannot be solved by pure logic or analysis, neither by intuitive imagination. It is a synthetic a priori judgment and the intuition of pure number, which, according to Poincare, can give a rigorous proof and is real mathematical reasoning. Then, what is the role of logic in inventions? The major role of logic is the analysis of mathematical constructions to prove them. Thus, both logic and intuition have their essential roles in mathematics: “Logic, which alone can give us certainty, is the instrument of demonstration; intuition is the instrument of invention” (Poincare, 1958 p. 207).

### *Experts as Knowledge Masters*

In the forgoing section, I discussed analysts and creators as two types of minds in mathematics and compared how they differ in mathematical problem solving, purposefully excluding experts from the scope of the discussion. I find expertise more related to domain knowledge, knowledge representation in memory and experience in the domain; whereas, I consider analysts and creators more related to the way they approach a problem and the way their minds deal with problems and novelty. Here, I will discuss experts as masters of domain knowledge, the third kind of mind in mathematics.

First, what is expertise? The electronic version of Encyclopedia Britannica Dictionary 2003 includes the following definition: having, involving or displaying

special skill or knowledge derived from training or experience. An expert is one with the special skill or knowledge representing mastery of a particular subject. This definition implies at least the following: 1) expertise involves extensive training or experience in the form of relevant knowledge. 2) It must be demonstrated. 3) It must be activated in the relevant domain. Grigorenko (2003) also made such an analysis of expertise in the definition of expertise. Expert knowledge is linked to the knowledge base and building this knowledge base is associated to extensive training. In the definition, as Grigorenko also stated, individuals encounter two limiting factors. First, an individual usually demonstrates expert performance in one domain or in a limited number of domains when compared to the individual himself-within individual comparison. Second, expert performance is compared to other individuals in the same domain to draw some distinguishing line between experts, developing experts and novices. However, the last point is not free of controversy. For example, rejecting the idea of expertise as intrinsic to a person, Connell, Sheridan and Gardner (2003) asserted that someone cannot have expertise; rather, an individual can have competencies of coordinated skills and factual knowledge that help her/him to solve problems or create products that are valued in a culture. Therefore, the field or the culture defines expertise about meeting some predetermined criteria set by the field.

The theories of expertise have been influenced by studies of masters, experts and novices in chess. For example, using a think-aloud procedure in a pioneering study, DeGroot (1965) investigated expert chess players' strategies. DeGroot concluded that the knowledge experts acquired playing chess over a prolonged time enabled them to

outperform their opponents. DeGroot drew a direct line between knowledge base, experience and expertise. Experts were better at recognizing meaningful chess configurations. DeGroot thought that the better recall ability of experts was related to their extensive knowledge base while Chase and Simon (1973) asserted that experts' better recall ability could be explained simply by their better memories. According to theories of expertise (Chase & Simon, 1973; DeGroot, 1965), expertise is characterized by domain-specific knowledge. Consider the major features of experts' knowledge (Bransford, Brown & Cocking, 2000): Experts have mastered extensive content knowledge in organized ways that enable them to understand the deep structure of the subject matter; however, experts' knowledge cannot be characterized as isolated facts. Their knowledge is organized around principles and big ideas and reflects practical applicability in specific situations. This knowledge organization helps them to discern meaningful patterns of information and to retrieve relevant knowledge fluently. The superior recognition ability of experts was explained by the way they chunk information into meaningful, relational and familiar patterns in their memory (Bransford, Brown & Cocking). In an analysis of DeGroot's study and Chase and Simon's study, Grigorenko (2003) concluded that the organization of experts' knowledge base rather than only superior recall ability was the major factor that distinguished experts from novices. Further, based upon my own analysis of studies on expertise, expertise also can be explained by the interaction effect of recall and the knowledge base.

Other points of view and lines of research that need to be discussed are related to deliberate practice and expertise. Individuals having these views usually undervalue the

importance of innate abilities, but underscore the importance of deliberate practice. Expert performance is accounted for by such factors as early involvement in domain-related activities, early training and the amount of relevant experience (Ericsson, 2003). Based on empirical data and extensive literature review, Ericsson stated that expertise is primarily acquired and that learning mechanisms primarily explain this acquisition, not the innate abilities because changing an individual's innate potential associated with learning is impossible. Although this is a very extreme view that undervalues an individual's genetic endowment in explaining expert performance, this view also is very important in pointing out the importance of deliberate practice associated with expertise. According to Ericsson, the key principle to attain expert performance is not mere practice, but continuously increased challenges that improve the performance beyond the current level.

*Expertise in mathematics.* My discussion of expertise, hitherto, has been related to general views of expertise. How, then, is expertise related to mathematics? An expectation of superior performance from someone who lacks knowledge in mathematics is rather unimaginable. First, I will discuss the kinds of knowledge an individual needs to solve mathematical problems as Mayer proposed (1991, 2003). Then, I will discuss expertise in mathematics and how it is related to the kinds of knowledge in mathematical problem solving.

According to Mayer (1991, 2003) successful mathematical problem solving requires knowing the kind of knowledge needed in each phase for solutions. The first phase is *problem representation*; that is, the problem solver translates the problem and

integrates problem elements into mental representations. The second phase is *problem solution* wherein the problem solver plans a solution strategy and, then, executes this strategy to produce solutions. In problem representation an individual needs linguistic, semantic and schematic knowledge to understand the problem and translate it into internal representations. *Linguistic knowledge* helps one to understand the meaning of words or sentences in a problem. However, if the problem statement does not contain any words or phrases, a problem solver may not need linguistic knowledge to solve the problem. An algebraic problem such as  $2y = x$ ;  $y = 1$ ;  $x = ?$  does not require linguistic knowledge. *Semantic knowledge* is based upon the kind of facts, rules or theorems in problem solving. An example is that one straight line divides a space into two equal halves. Semantic knowledge is important in mathematical problem solving, in that many mathematical problems require at least simple facts for solutions. For example, an individual who has no knowledge of graphs will spend a great deal of time understanding and solving time-rate-distance problems provided that he has not learned the formula  $\text{distance} = \text{time} \times \text{rate}$ . Needless to say, semantic knowledge sets up the foundation for high performance in mathematics. Without this kind of factual knowledge, one hardly performs at a higher cognitive level. While semantic knowledge constitutes the foundations in mathematical problem solving, another kind of knowledge is needed to make connections between chunks of information or mathematical facts and distinguish irrelevant from relevant information to integrate the relevant. This type of knowledge is called *schematic knowledge*; namely, knowledge of problem types (Mayer, 1991). Schematic knowledge constitutes the bridge between single facts. For example, knowing

that all area problems are based on a length  $\times$  width relationship helps a problem solver carry out further or harder area problems, such as the area of a cube or a prism rather than the area of a square or a rectangle. When a problem solver recognizes that a problem involves area, s/he activates schemas related to area formulas.

In problem solution an individual needs at least two types of knowledge as proposed by Mayer (1991). One is *strategic knowledge*, which the individual uses to plan and monitor a problem solution. Strategic knowledge usually plays its role after the individual understands and builds the mental representation of a problem. The individual can set up some plans or strategies to execute the solution, such as determining subgoals, reducing or isolating elements. Polya (1957) also suggested a set of strategies for effective mathematical problem solutions, such as finding similar but easier problems and going backward in the solution process. However, neither model of strategic knowledge is related to factual mathematical knowledge; rather, they involve skills or knowledge of problem solving processes. Another type of knowledge in problem solution is called *procedural knowledge*, the knowledge of how to carry out a sequence of operations such as computations. For example, in finding the area of a rectangle, a problem solver needs to know how to multiply two numerical values, say 3 and 4, to get the correct answer. Notice that this stage of problem solving involves a rather basic cognitive process, computation.

*Experts' knowledge in mathematics.* In the foregoing sections I discussed ideas of mathematicians about how mathematics is a science of structure and relations and how the knowledge of structural features of problems is much more important than those of

surface features. Therefore, only isolated, factual, mathematical knowledge does not suffice in demonstrating expert performance; rather, knowledge needs to be structured in a way that enables the recall of chunks of information in an interconnected manner, which helps the problem solver discern structural relations among elements of a problem. Mathematics experts, for example, were able to recognize quickly patterns of information in situations that entailed specific classes of mathematical solutions. Physicists were able to recognize problems of river currents and problems of headwinds and tailwinds in airplanes as involving similar mathematical principles, those of relative velocity (Hinsley, Hayes & Simon, 1977).

The kind of knowledge experts possess that enables them to recognize problem types is characterized as involving an organized, conceptual structure or what is called schematic representation of knowledge in memory. This type of representation of knowledge, as opposed to isolated facts, enables experts to think around principles when encountering a problem while novices tend to solve problems by attempting to recall specific formulas that could be applicable to the problem. For example, Hinsley, Hayes and Simon (1977) investigated schematic representation of high school and college students. They presented students with a problem that was a distance-rate-time problem, but they included irrelevant information about a triangular relationship. Hinsley and colleagues found that the participants used either distance-rate-time schema or triangle schema to solve the problem. They concluded that the participants' schema influenced what they looked for in the problem statement and, thus, their choice of problem solution.

Consider the following line of research on knowledge representation differences among experts, developing experts and novices in mathematics. Silver (1981) asked seventh graders to classify 16 arithmetic word problems into groups. Each problem overlapped with one or more problems both in its mathematical structure and in its story context. Students were grouped into three levels of proficiency representing three levels of expertise: good, average and poor based on their performances in solving arithmetic story problems. Expert students classified problems according to similarities in their mathematical structures while novices grouped problems based on their story contexts. Schoenfeld and Herrmann's experimental study (1982) provided evidence that representation differences between experts and novices were related to knowledge differences, not ability differences. Students' representations of a set of problems before and after instruction and also compared their representations to college professors' representations. The findings indicated that students' representations shifted from grouping problems according to their surface features to structural features as a result of the instruction, though their representations still were not similar to those of professors after the instruction. Further, Aaronson and So (as cited in Novick, 1992) investigated how much time expert problem solvers allocated to specific components of problems and compared their allocation to novice problem solvers. Novice problem solvers spent most of their time on words that represented surface features of problems, such as objects and actors, while expert problem solvers spent most of their time on the components that represented structural features of problems, such as units and operations.

In addition to knowledge representation, mathematics experts also differ in their level of semantic knowledge. Concerning students' developing expertise and their level of knowledge, Ni (1998) investigated 7 and 8 year-old children's performance on classificatory reasoning tasks that required sorting dinosaurs into class-memberships and into class inclusion. Children were tested to determine their cognitive levels as pre-operational or concrete operational using Piagetian conservation and classification of quantitative tasks. Children's expertise level was determined by measuring their knowledge level about dinosaurs. The findings indicated that levels of expertise were found to have a significant effect on children's sorting performance. Children with high expertise were more able to base their sorting on domain-relevant solutions, including both perceptual (surface) and implicit features (structural) while children with low expertise based their sorting on perceptual-domain relevant features and domain-irrelevant features. Also, my own research (Sak & Maker, 2003) involved first through 6<sup>th</sup> grade students' level of factual or semantic knowledge and their use of numerical strategies in mathematics. After statistically controlling the effect of age, semantic knowledge still had a significant effect on the students' use of numerical strategies. Students with high knowledge were better able to produce more strategies and were more likely to elaborate on the strategies they produced. Likewise, Ostad (1998) investigated 2<sup>nd</sup>, 4<sup>th</sup>, and 6<sup>th</sup> grade mathematically normal and mathematically disabled children's performance on arithmetic word problems and on number fact problems and their use of task-specific strategies to solve the problems. Mathematically normal children used verbal and material strategies in the early grades, and they advanced in mental strategies

as they advanced in school. Mathematically disabled children primarily used material and verbal strategies at all grades. The author concluded that the poor performance of mathematically disabled children might be explained by retrieval problems and working memory problems, along with their insufficient conceptual knowledge. These three aforementioned studies indicate that a domain-specific knowledge base helps children to process and recall task-specific information more efficiently and helps them apply strategies more effectively.

As a final remark, let's assume some secondary students are developing experts. A student, for example, with strong memory can recall many facts related to a given problem. This student, having only factual knowledge, can be thought at the first level of expertise or as a novice. What else s/he needs is the knowledge to relate each piece of information to other chunks of information stored in her or his memory. For example, knowing the formula to calculate the area of a square drawn in a triangle, whose sides are defined, may not be enough for this student to find the area. Obviously, the solution of this problem requires some knowledge of relationships between a square and a triangle. The student who can solve this problem may be thought to be at the second level of expertise, or developing expertise, when compared to others who have similar experience. A third level may be beyond the relational knowledge. This level can be characterized as schematic or conceptual knowledge as elaborated by theorists of expertise. Here, the student has a knowledge representation that enables her or him to solve mathematical problems that require inferring relations of relations between chunks of information or applying to a mathematical theorem that looks irrelevant at first glance.

Reconsider the above square-triangle example. If more demand is added for this problem in a way that the problem also requires square-triangle-side-angle relations. The student needs to link relations between sides and angles of a square and those of a triangle to find the solution, thereby, connecting chunks of information stored in the memory. The student at the first level of expertise, or novice level, has difficulty in solving the third problem because the solution of the third level problem is more related to conceptual representation of facts than isolated representation of single facts. Note that this model of classification of levels of expertise can apply to a group of people who have similar experience in a knowledge domain, such as a group of elementary students, a group of high school students or a group of graduate students in mathematics. The levels of expertise also can be applied to a more extreme comparison of groups, such as high school students as novices, college students as developing experts and college professors as experts.

### Part III

#### Psychological and Neuropsychological Investigations of Creative and Analytical Abilities

The purpose of this part of the literature review is to provide psychological and neuropsychological supports to the idea of differing mathematical minds. The mode of support includes empirical research findings about analytical and creative abilities directly or indirectly related to mathematical ability. I draw some distinguishing lines between the two abilities from research findings in a spectrum of studies on human abilities.

Creativity is a rare trait, and it involves dealing with novelty. Analytical ability involves analysis, comparison, contrasting and evaluating, not necessarily novelty. The consensus among researchers is that creativity is a higher order “emergenetic” trait (Eysenck, 1995; Martindale, 1999a, 1999b). That is, creativity is a manifestation of the confluence of several genetic, cognitive, personality and environmental factors (Sternberg & Lubart, 1995). The confluent nature of creativity also accounts for distribution in that while any trait (e.g., intelligence) related to creativity may be distributed normally in the population, the distribution of genius is not. Eysenck (1995) articulated that all the necessary traits for creativity would occur simultaneously in anyone but geniuses. To Poincare (1958), while all creative people also are analytical, analytical people are not necessarily creative.

*Successful Intelligence: A Psychological Theory of Creative, Analytical and Practical Minds*

Perhaps the first process-oriented theory of intelligence that psychologically specified creativity and analytical ability separately is the theory of successful intelligence. Sternberg (1997) proposed the theory of successful intelligence based on process skills, which were conceptualized in his earlier theory of the triarchic mind (1988). According to Sternberg, intelligence has three aspects that underlie the theory of successful intelligence: creative, analytical and practical abilities. According to this theory, a common set of processes underlies all aspects of intelligence. These processes are thought universally important, but they may show developmental differences in different contexts. For example, Sternberg (2002) asserted that solutions to problems in

one culture might be intelligent while they are not in another culture. However, the intellectual processes used in solutions are the same. Defining a problem and translating strategies to solve problems, for example, exist in any culture.

Analytical ability involves the use of the components of intelligence in analyzing, comparing, contrasting, evaluating and judging relatively familiar problems. Thinking occurs at an abstract level. Sternberg (1977) analyzed analytical kinds of problems, such as syllogisms and analogies, into components to understand the information processing origins of individual differences in analytical ability. Encoding, inference and comparison were found to be important processes in the solving of analytical problems. Particularly interesting was the finding that good reasoners allocated more time in the encoding phases, which Sternberg interpreted in his finding that good reasoners also used their metacomponents efficiently (e.g., defining a problem, and planning and monitoring problem solution). On the other hand, creative ability is invoked when the information processing components of intelligence are applied to relatively novel problems or situations to create, design, imagine, suppose, explore, invent or discover; that is, novelty and production are the focus. For example, Sternberg (1982) investigated how people dealt with novel uses of some words, as opposed to their conventional uses, such as blue-green and grue-bleen. The information processing component requiring people to switch from conventional uses and novel uses or backward and forward was a good measure of the ability to cope with novelty. In addition, practical ability involves solving real life problems and is related to tacit knowledge. However, I conceptualize practical ability as

the knowledge aspect of expertise, since it is more related to knowledge and experience. Sternberg (2002) also stated that tacit knowledge is acquired by experience.

Sternberg and his colleagues have carried out a series of studies to investigate the validity of the theory of successful intelligence (a complete review in Sternberg, 2002; Sternberg, Castejon, Prieto, Hautakami, & Grigorenko, 2001; Sternberg, Grigorenko, Ferrari, & Clinkenbeard, 1999), using the Sternberg Triarchic Abilities Test (STAT). Their factorial research supported the triarchic theory of intelligence, revealing separate and uncorrelated analytical, creative and practical factors. STAT measured three kinds of ability in three domains: analytical, creative and practical abilities in verbal, figural, quantitative domains. Quantitative analytical ability was measured through a test of number series. Students had to figure out what came next in a series of numbers. I think this part of the test may involve some aspects of creative ability because students have to discover rules that form number series. Rule discovery is an important aspect of mathematical creativity according to mathematicians (Poincare, 1952b; Polya, 1954a). Creative ability was measured by a test of novel number operations. Students were presented with rules that defined what operations students had to use. For example, “flix” involves numerical manipulations that differ as a function of whether the first of two operands is greater than, equal to, or less than the second. From my point of view, this kind of problem or operation might be novel to students in the first encounter; however, once they start to solve problems of the same nature, problem solving would be easier and no longer novel, because students automatically would know what to do to

solve problems of the same nature. Further, quantitative practical ability was measured by scenarios, requiring students to solve real life problems using mathematical skills.

*Psychological Investigations into Creativity and Analytical Ability*

According to the neural network model of creativity, creative individuals have access to higher numbers of potentially useful mental associations (Martindale, 1981, 1995). However, a creative person must eliminate less promising alternatives prior to testing the potentially more promising ones. That is, the creative process is characterized by the generation and subsequent selection of hypotheses (Eysenck, 1993; Simonton, 1989). The theory behind this lies in the context of chance-configuration theory. Campbell (1960) argued that creativity is the result of a two-step process. The first step, referred to as blind variation, involves the generation of heterogeneous mental activity, which gives rise to a set of configurations or potential hypotheses. The second step involves subjecting those configurations to selection, thereby reducing unsuccessful configurations and arriving at the most viable hypothesis. What is important here is what helps eliminate unpromising ideas and what processes are used in the generation of heterogeneous hypotheses and the subsequent elimination of them. *Selection* seemingly plays a major role in the confirmation of the most useful hypotheses among many.

In an attempt to understand the relationships among creativity, inductive reasoning and selection, Vartanian, Martindale and Kwiatkowski (2003) tested participants using the Alternate Uses Test and Wason's 2-4-6 rule discovery test. The former is a divergent thinking test while the latter is a test of rule discovery by inductive reasoning. The results of the study revealed that performance on the 2-4-6 task was

related to potential creativity, as measured by fluency scores on the Alternate Uses Test. That is, the group that was successful in discovering the rule generated more confirmatory and disconfirmatory hypotheses than the unsuccessful group. Fluency accounted for a significant amount of variance in rule discovery in the step-wise regression analysis. The authors concluded that fluency, as in the ability to generate ideas, was associated with success on the 2–4–6 task because participants with higher fluency scores would generate more hypotheses, thereby increasing the probability of discovering the target rule.

Another line of research in the psychological and neuropsychological investigations about creativity is related to the influence of selective attention on creativity. Cognitive attention on given information to solve problems has been debated as one source of mathematical creativity (Poincare1957; Polya, 1954a). The controversy is that focused attention promotes the generation of strong associations while diffused attention supports the generation of remote associations (Finke, Ward & Smith, 1992; Martindale, 1995). That is to say, individuals who disperse their cognitive resources more easily might be more likely to generate more unusual associations than are those whose cognitive resources are more narrowly focused. Researchers in this line of study, however, suggest that those who are creative and routinely allocate their attention in a diffuse manner have more difficulty completing the target task than do those who are less creative and who maintain a more narrow focus. For example, Dykes and McGhie (1976) found that under certain conditions, highly creative individuals showed more shadowing errors on a dichotic listening task than did less creative individuals. Rawlings (1985)

reported that creative individuals showed more intrusion errors on the shadowing task than did less creative individuals; however, compared to the less creative participants, creative individuals showed a better memory for the secondary information. That is, creative individuals were better able to complete a task that required diffuse attention (i.e., the memory task) than were the less creative individuals, but the pattern was reversed for a focused attention task such as shadowing. In addition, Mendelsohn and Griswold (1966) found that, when solving anagrams, creative individuals were more likely to take advantage of incidentally presented hints than were those who scored lower on a measure of creativity. Mendelsohn's review of empirical research also showed a positive relationship between measures of attentional capacity and non-verbal and verbal indicators of creativity.

According to the present author, a diffuse attentional strategy would result in trivial outcomes if not supported by focused attention. That is, although creative individuals seem to have a propensity toward allocating attention broadly, when the situation demands, they must be able to focus their cognitive resources on certain parts of the situation. In fact, Martindale (1995), in his description of a connectionist model of creativity as discussed before, asserted that the ability to change cognitive states between defocused and focused attention is a crucial characteristic of creative thinking. For example, Dallob and Dominowski (1993) found that when participants' attention was experimentally drawn to certain aspects of insight problems (focused or selective attention) solution rates increased significantly.

An important question related to creativity, analytical ability and selective attention is whether the pattern of resource allocation or diffuse attention as a trait possessed by individuals different in creative and analytical ability. Here, creative thinking is thought of as the ability to generate associations or relations. Analytical ability is thought of as the ability to dissect, divide or analyze parts, aggregates or juxtapositions. Martindale (1995) discussed analytical and creative thinking by saying that a creative insight is not possible with deductive reasoning because the conclusion is implicit in the premises, and creative productions usually require remote associations. One assumption based on the notion of selective attentional differences between creative people and analytical people is that the acquisition of a large amount of information during problem solving might result in a trade-off of crucial processing capacity. That is, for some types of problems (e.g., deductive reasoning problems) gathering incidental information may not allow efficient and effective processing of more central information. Because analytic thinking involves an evaluation/dissection of the problem elements, sustained focus on the problem elements is required for solution—attention directed to peripheral items simply wastes cognitive resources (Dykes & McGhie, 1976). That is to say, analytical thinkers should not exhibit an inclination to diffuse attention. Therefore, the propensity to allocate attentional resources to aspects of the problem solving situation, which are not obviously central, should be uniquely characteristic of creative thought and not exhibited by analytical thinkers. To test this assumption and to investigate selective attentional differences between analytical thinkers and creative thinkers, Ansburg and Hill (2003) tested participants, using the Remote Associates Test (RAT) to measure

creativity, a variety of deductive problems to measure analytical thinking, and a variety of focal and control anagrams, along with some peripheral cues.

Those who tended to make unusual connections were more likely to allocate their attention in a diffuse manner than were those who were more analytical. Ansburg and Hill (2003) suggested that allocating attention broadly was not a strategy routinely employed by all good problem solvers. Instead, this cognitive trait was one that distinguished creative problem solving from other kinds of problem solving. The finding that creative thinkers used a different cognitive resource allocation strategy from analytical thinkers also is consistent with the view that creative problem solving is distinct from other kinds of problem solving. For example, Shaw and Conway (1990) found that during a word-detection task creative individuals were more sensitive to nonconscious information than were less creative individuals. That is, highly creative thinkers were more likely to produce unconsciously primed solutions than were the less creative thinkers. Schooler, Ohlsson, and Brooks (1993) found that when participants were forced to become aware of their problem-solving procedures through verbalization, insightful problem solving was inhibited. Shaw (1992) suggested that creative individuals gather information using attentional processes that might occur unconsciously. According to Shaw, this unmonitored stream of information activates prior knowledge and can account for the suddenness with which creative solutions come into consciousness.

*Neuropsychological Investigations about Creativity and Analytical Ability*

Neuropsychological researchers also have investigated the structure and the functions of the brain related to analytical and creative abilities. Kris (1952), for example, distinguished among primary process thinking, which is analogical, and free associative thinking, commonly associated with creativity, and secondary process thinking, which is logical and reality-oriented, commonly associated with analytical thinking.

Neuropsychological research implies that creative people are more flexible along the primary process–secondary process continuum and that creative insights are more likely to occur in a primary process mode of cognition (Martindale, 1989). Deductive reasoning, for example, represents a form of secondary process thinking (Martindale, 1995); whereas inductive thinking, characterized by the generation or discovery of original ideas, can be associated with primary process thinking.

*Creativity, analytical ability, and hemisphericity.* According to Martindale (1999a), the right hemisphere of the brain operates in a primary-process fashion while the left hemisphere operates in a secondary-process fashion. Researchers in this line of reasoning suggest that verbal, sequential, and analytical processing occur in the left hemisphere, while right-hemisphere capabilities involve parallel and holistic processing of nonverbal stimuli and creativity. In other words, the right frontal lobe is more involved in spontaneous production of non-verbal representations; whereas, the left lobe may exert control and secondary evaluative and verbal analysis. For example, Martindale, Hines, Mitchell and Covello (1984) found that creative people had significantly more right- than left-hemisphere activity as measured by an electroencephalography (EEG), as

opposed to low creative people on a creativity task, but not on a non-creativity task. In their study, student artists showed much greater right than left hemisphere activity during the drawing test than did the control group. They also included a reading task to measure asymmetry during a noncreative task. On this task, artists showed more asymmetry than did the control group, but in the opposite direction from that found during the drawing task. Left hemisphere activation was greater than right hemisphere activation, and this was more the case for the artists than for non-artists. Thus, Martindale and his colleagues concluded that creative people relied more on the right hemisphere than the left hemisphere. Also, Hoppe and Kyle (1990) examined patients with commissurotomy (split-brain) as well as normal subjects, and they concluded that creativity depended on whether the presentational symbolization and imagery in the right hemisphere were available to the left hemisphere via the corpus callosum.

Another method of study of hemisphericity that might be related to creativity is the use of right and left hand advantages and right and left visual field priming. For example, Poreh and Whitman (1991) compared undergraduate right-handed males on a variety of creativity tests such as the Torrance Test of Creative Thinking Verbal and Figural forms, the Remote Associates Test, and Verbal Closure based on participants' ear advantage (hemispheric dominance). They found that verbal convergent scores varied as a function of hemisphericity. Scores on this factor were found to be higher for all subjects with a right ear advantage (left hemisphere dominance). They concluded that verbal processing and verbal convergent scores were significantly related to

hemisphericity. Also important was that individuals with a left ear advantage generated a small number of ideas when processing nonverbal stimuli.

Another neuropsychological method to study thinking and brain functions is the measurement of regional cerebral blood flow (CBF). Carlsson, Wendt, and Risberg (2001) investigated the relationship between creativity and hemispheric asymmetry as measured by CBF. They found that the highly creative group used bilateral prefrontal regions when doing the Brick task (a creativity test), while the low creative group used functions predominantly on the left side. When the activation response during the fluency task was compared with that of the Brick task, the highly creative group showed increases in all three bilateral prefrontal areas. The low creative group showed more decreases and had an unchanged level only in the left anterior prefrontal region. Furthermore, the superior frontal regions seemed to play a special part in this investigation. Good performance on the Brick task was negatively correlated with high activity both in the left and in the right region.

*Creativity and cortical activation.* According to cortical activation and coherence theories, creative individuals are distinguished from noncreative ones by the distributed pattern of their cortical activation and the coherence among the regions of their brain (Martindale, 1999a). Cortical activation and coherence characteristics help creative people think of a creative idea because the spread of activation across a wide associative horizon makes it more likely that two distant nodes may be simultaneously activated and that a coherence between these nodes is established to form a novel composite concept. For example, Martindale and Hines (1975) measured the amount of EEG alpha wave

activity, an inverse measure of cortical arousal, while subjects took the Alternate Uses Test (AUT), the Remote Associates Test, and an intelligence test. High creative participants showed a differential amount of cortical activation across the three tasks, whereas low creative participants did not. Creativity, as measured by the Remote Associates Test, was connected with a tendency to exhibit differential amounts of alpha on different types of cognitive tasks, while creativity, as measured by the Alternate Uses test, was connected with a tendency to exhibit a high percentage of basal alpha on a variety of cognitive tasks. The finding does make psychological sense because the Alternate Uses Test is primarily a measure of ideational fluency or the ability to come up with a large number of ideas. The Remote Associates Test, on the other hand, requires both producing cognitive elements and putting these elements together to come up with a correct answer. Thus, both focusing and defocusing abilities would seem necessary for good performance on it. Creativity, as measured by either the RAT or the AUT, is associated with exhibition of a high percentage of basal alphas while taking the Alternate Uses Test. In contrast, creativity as measured by either test is not associated with greater alpha abundance during the basal or feedback conditions. This finding supports the hypothesis of an association between creativity and low cortical activation, specifically during tasks that call for or allow creativity.

*Creativity, analytical ability and neural coherence.* Electrical relatedness in some way reflects the functional relationship among brain areas (Sheppard & Boyer, 1990). Nunez (1995) argued that decreased overall coherence obtained among the regions of the brain when a cognitive task is performed could indicate that cognitive processing

involves a general shift from more global to local operation. On the other hand, Petsche (1996) suggested that increases in coherence might indicate closer cooperation of the brain areas; whereas, a coherence decrease shows that brain regions become functionally more separate. In both cases, the number of coherence changes centered on an electrode could be an indicator of the functional importance of this region for the task. Thatcher and Walker (1985) demonstrated a negative correlation between coherence increase and IQ. Notice that IQ mostly is a measure of analytical ability. Petsche (1997), who also correlated coherence measures with scores on a text composition task, found that most of the correlations obtained in males were positive and related to the left hemisphere. Note that the left hemisphere hosts analytical ability. In another study, Petsche (1996) compared coherence changes in people while they were performing creativity tasks in verbal, visual and musical domains. He further demonstrated that acts of creative thinking were characterized by a greater coherence increase between occipital and frontopolar electrode sites than in the solution of more closed problems.

Perhaps Jausovec's study (2000) is more related to the cooperation of brain regions with creative and analytical abilities. He investigated the differences in cognitive processes related to creativity and intelligence using EEG coherence and power measures. He compared intelligent, creative and average people in coherence and power measures (alpha wave) while participants were solving open and closed problems. Intelligent people were those who had high IQs, and creative people were those who had high scores on a creativity test, but not necessarily high IQs. He found that creative individuals displayed more inter- and intra-hemispheric cooperation among the far distant brain

regions. He concluded that generally increased coherence, together with more selective involvement of cortical zones, reflect the specificity of functions among the areas of the brain that, in turn, is related to creativity.

In conclusion, the research briefly discussed above implies that creativity and analytical ability might be very different, but they are not distinct process abilities because some processes underlie both abilities. At least, this distinction is my personal conviction based on the research findings. First, creativity requires cooperation among many regions of the brain to generate remote or original associations or products, and analytical thought occurs among nearer regions, particularly in the left frontal part of the brain which results in strong, connected associations that may be weak in originality. Second, creativity involves both focused and defocused attention on a problem situation while analytical thought requires only focused attention. In the context of creativity, an individual needs to generate diverse ideas. The production of diverse ideas is facilitated by defocused attention. Then, the individual selectively eliminates ideas that are unpromising for the problem solution. The selective elimination of ideas is facilitated by focused attention. In the context of analytical thought, an individual does not necessarily need to perform the first step, namely, production of many heterogeneous ideas or hypotheses. However, the individual has to analyze each piece to verify what has been found new or has been discovered in the first step. My conclusion is that an individual can have very strong analytical and creative abilities or only one of them, depending on the person's biological and experiential background.

## Part IV

### The Three-Mathematical Minds Model

The discussion in this chapter provides partial support for the idea of different mathematical minds. In this part, I offer the three-mathematical minds model to the study and assessment of mathematical ability. Before I discuss the model, I will present my conceptions of mathematical ability and of mathematical giftedness.

From the author's point of view, mathematical ability is the biopsychological potential to reproduce mathematical constructions or to produce new and useful constructions or to learn mathematical facts, rules, theorems, or laws that could be activated in mathematical problem solving. Mathematical giftedness is the mathematical competence demonstrated in the form of production, reproduction or problem solving in any branch of mathematics at a given time and is recognized as remarkable by members of mathematical communities (e.g., teachers or mathematicians). A mathematically gifted person is the one who demonstrates outstanding competence compared to his or her age, grade or experience peers to perform logical analysis of mathematical constructions, to produce new constructions or to solve mathematical problems that are recognized as being exceptional within the person's context.

Consider that underlying differences exist between mathematical ability and mathematical giftedness in the above definitions. The former reflects biopsychological potential while the latter is a form of demonstrated potential at a certain level of competence that is recognized as superior by members of mathematical communities. The second distinction is the emphasis placed upon productive, reproductive, and

problem solving components and upon different branches of mathematics, such as space or numbers. That is, an individual demonstrating remarkable competence in one or all components in any of the branches of mathematics may be thought the mathematically gifted.

The three-mathematical minds model, therefore, is aligned with the above definitions. This model is a tool to reconcile various views in that giftedness can emerge in many forms as postulated by Sternberg (2000), such as analysts, experts, creators, or in the interaction of any of them as seen in figure 2.1. Figure 2.2 shows how the three minds differ in certain aspects. They differ in three aspects, respectively. The first difference is the cognitive components, such as memory, intuition or logic, to carry out cognitive tasks. The second is cognitive tasks, such as domain facts, novelty or ambiguity. Finally, they differ in their end-products as a result of applying certain cognitive components in different tasks, such as knowledge production or reproduction. In this model, expert mathematical minds differ from the others in their knowledge representation, amount of knowledge, and in experience, but not necessarily in their cognitive ability and styles. Their knowledge is specialized, representing domain-specificity or task-specificity; therefore, their cognitive end-products are internalized knowledge. Unlike experts, creators and analytical thinkers manifest themselves more as a function of their thinking, such as the differences in their brains' neural activation, the way they approach a problem, their focused or diffused attention, their logic or intuition, and the way they deal with information, such as to search for novelty or to search for

ambiguity. Although experts, too, differ in their thinking compared to novices, this difference stems mostly from the nature of their knowledge.

I am certain that every mental work is achieved through many cognitive processes; but the one that remarkably differentiates creative people from the other types, analytical and knowledgeable, is the work of the unconscious, which I have not addressed yet. The unconscious may initiate a spectrum of ideas, which Poincare (1952b) compared to scattering atoms. According to Hadamard (1945), the unconscious might have levels that may vary in the proximity to the conscious. Some levels at which remote ideas congregate may be deeper while other levels at which near ideas come together may be superficial. The former characterizes a creative mind while the latter depicts an analytical kind. Further, the processes I have discussed hitherto can behave differently in different mathematical minds. For example, Hadamard said that the extent to which our minds work freely can be a plausible reason why we are more an intuitive type or a logical type, because scattered ideas can be brought into life depending on the extent to which our mind takes a direction of thought either narrower or wider. The former is typical of the analytical type. The latter is the distinctive mark of the creative type.

Figure 2.1 shows that it is plausible to think of mathematical giftedness in seven forms. That is, giftedness may be conceptualized based on the interactions of the three minds. For example, the interaction of expertise and analytical ability produces an *expert analyst*, who is competent both in domain knowledge and in analysis. The interaction of expertise and creativity makes a *creative expert* who is a good intuitive free thinker and has remarkable domain knowledge. By the same token, the interaction of analysis and

creativity gives birth to a *creative analyst*, who has both good, logical judgment and an *a priori* synthetic judgment. Finally, the interaction of all brings into being a “*master*,” who demonstrates remarkable analytical ability, domain knowledge, and creative productivity who, no doubt, is very rare.

In conclusion, I have proposed the three-mathematical minds model to the study and assessment of mathematical giftedness. Although they differ in many respects, they by no means are distinct constructs; rather, fundamental knowledge and skill components underlie all three. However, it is plausible to study, to assess, and to teach for all three types; otherwise, one type would be overlooked at the expense of the others (Sternberg, 2000). The rest of this study deals with the author’s empirical research to investigate the psychological validity of the three-mathematical minds model in the assessment of mathematical giftedness.

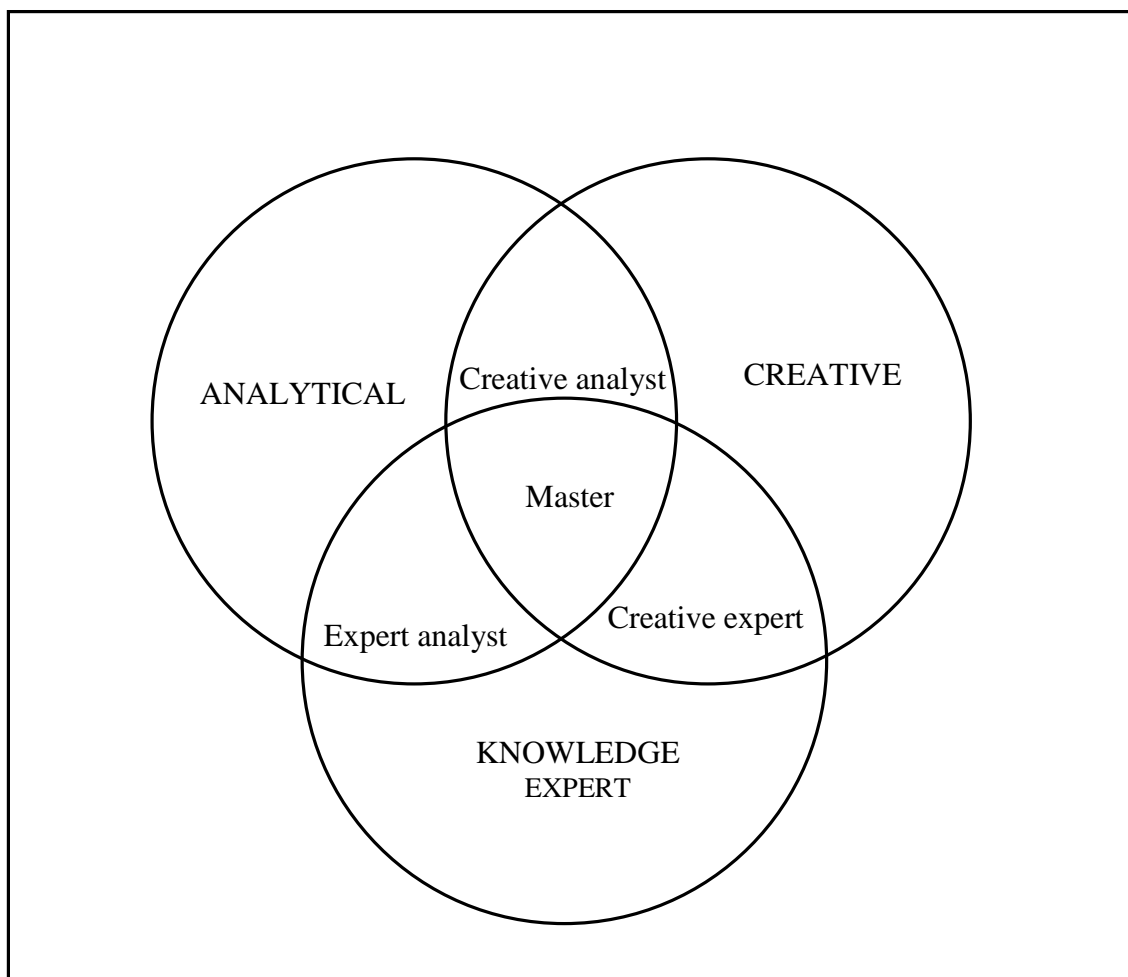


Figure 2.1. The three-mathematical minds model and seven forms of mathematical giftedness. This model is based on patterns of giftedness proposed by Sternberg (2000) and studies on expertise (Chi, Glaser, & Farr, 1988).

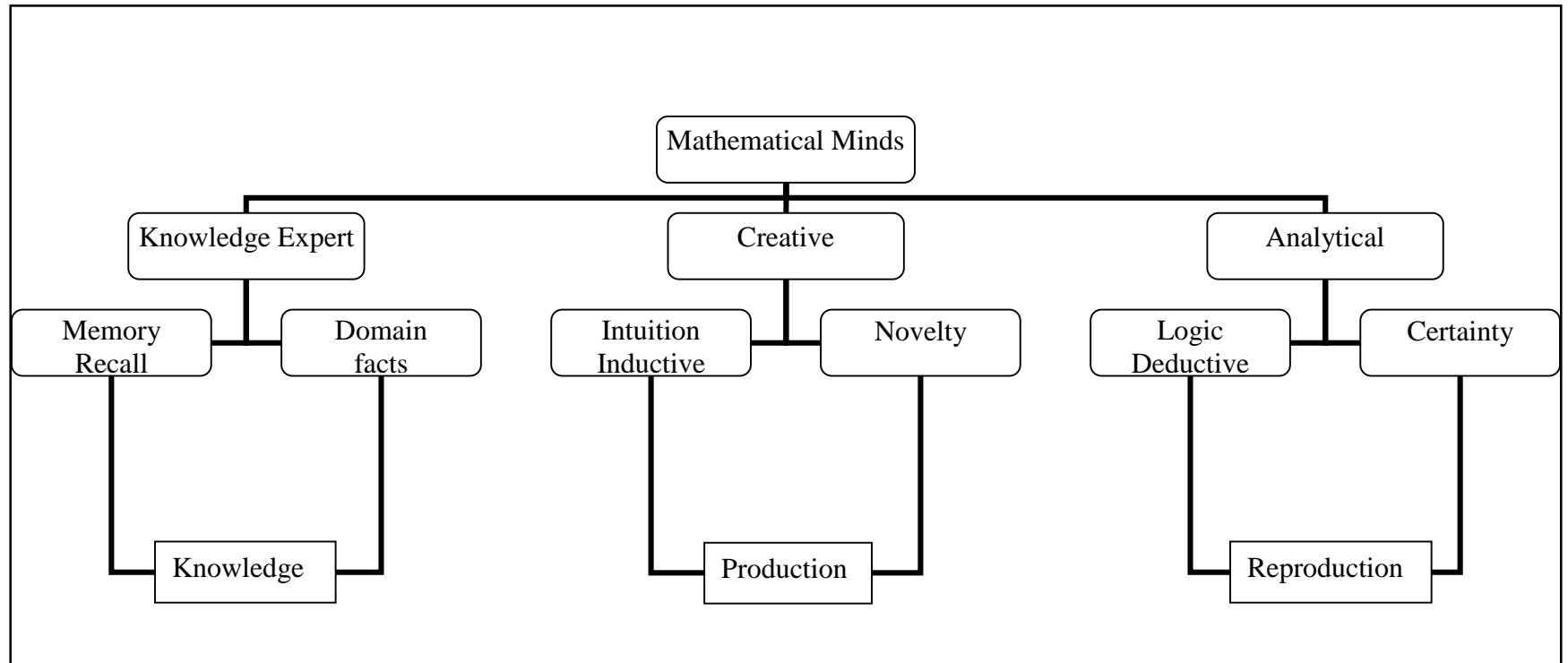


Figure 2.2. Major instruments of mathematical minds applied in cognitive tasks, and their end-products

## CHAPTER III

### METHOD

This chapter includes descriptions of participants involved in this study, procedures used to develop test items according to the three-mathematical minds model and the three-level cognitive complexity model. Finally, statistical data analyses used to answer the research questions are presented.

#### Participants

The total number of participants was 291. Participants included 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grade students from four different schools. Schools A and B were located in one city and school D was in another city. School C was located in a rural area. All schools were located in the southwestern part of the United States. The socioeconomic status of students in each school varied from lower to upper. Students of the schools A and D came mostly from middle class families. Students of the schools B and C came mostly from low-middle class families. The participants' ages ranged from 10.5 to 15.5. The mean age was 13.04. In addition, tables 3.1, 3.2, 3.3 and 3.4 show frequency and percentage distributions of the participants by school, grade, gender and race.

#### Procedures

A test of mathematical ability was developed through multiple steps and used to assess the participants' mathematical ability. The procedures used in the test development and administration of the test follow.

Table 3.1

Frequency of participants by school

School	Frequency	Percent
A	41	14.1
B	27	9.3
C	71	24.4
D	152	52.2
Total	291	100.0

Table 3.2

Frequency of participants by grade

Grade	Frequency	Percent
6	63	21.6
7	117	40.2
8	111	38.1
Total	291	100.0

Table 3.3

Frequency of participants by race

Race	Frequency	Percent
American Indian	4	1.4
Asian	10	3.4
Black	7	2.4
Mexican American	42	14.4
White	210	72.2
Other	18	6.2
Total	291	100.0

Table 3.4

Frequency of participants by gender

Gender	Frequency	Percent
Female	133	45.7
Male	158	54.3
Total	291	100.0

### *Instruments*

The test of mathematical ability was developed based on the three-mathematical minds model and the three-level cognitive complexity model. The test contained nine subtests for a total of 27 problems. Each subtest had three problems (see Appendix A for the entire test battery).

In addition to mathematics problems, the test booklet included two questionnaire items about students' liking of mathematics and their beliefs about the strength of their mathematical ability. They rated their liking and beliefs on a five-point scale, responding to the two questionnaire items below:

- A. How much do I like mathematics? (a) very much (b) much (c) some (d) a little (e) not at all
- B. How am I in mathematics? (a) excellent (b) good (c) ok (d) weak (e) very weak

Another questionnaire item was the teachers' rating of students' mathematical ability. The cover page of the test booklet included a space for the teacher to rate each student's ability according to the following scale. The teachers, who filled out the boxes after students completed the test, rated each as follows (Appendix C):

- 5) highly talented; 4) has high ability but not necessarily talented; 3) average; 2) weak ; 1) very weak

### *Development of the Test of the Three-Mathematical Minds ( $M^3$ )*

*Collaboration with experts.* A team of experts, with the leadership of the author developed mathematics problems according to the three-mathematical minds ( $M^3$ ) and

the three-level cognitive complexity model ( $C^3$ ), as seen in table 3.5. The content and complexity of items will be discussed later in this chapter. The team consisted of the following members: two mathematicians, one with a Ph.D in the science of mathematics and the other had a Ph.D in mathematics education; two middle and high school mathematics teachers; and the author, who specialized in the assessment of cognitive abilities, creativity and giftedness.

First, the author developed twenty-seven sample problems to measure the three aspects of analytical mathematical ability, creative mathematical ability and mathematical knowledge at three levels of cognitive complexity. Then, the sample problems were sent to the mathematicians for their review prior to the author's initial meetings with them. Afterward, the author met with each mathematician twice to review, to revise and to develop new problems. The first meeting resulted in modifying four problems, developing 20 new problems, keeping three original problems unchanged, and omitting 20 problems. The second meeting ended with modifying 22 problems, developing five new problems, keeping three original problems unchanged, and omitting two problems; thus, having a total of 30 problems.

In the second phase, the final 30 problems, of which 27 were main and 3 were additional problems, were sent to the teachers to review the content and difficulty level of each problem according to the level of 8<sup>th</sup> grade students' mathematical ability. The teachers used the following scale to rate problems: 1) very easy, 2) easy, 3) average, 4) difficult, and 5) very difficult. They rated the difficulty level of each problem by comparing it to other problems in the same subtest. For example, the problems in the

insight subtest were compared only to each other, not to other problems in the other subtests. The reason for such a rating procedure was the author's conviction that only problems of the same psychological nature could be compared in their psychological difficulty. The teachers' ratings of problems' difficulty later were used in the construct validity of the three levels of the cognitive complexity model discussed later.

Because of strong agreement between the author and the mathematicians on the difficulty level of the problems, I predicted high correlations between item cognitive complexity (ICC) levels of the 30 problems and the teachers' ratings of item difficulty for 8<sup>th</sup> grade students; however, correlational analysis indicated low and nonsignificant correlations, contradicting the initial agreements. The correlations were .14 between the ICC levels and the ratings of the first rater, and .25 between the ICC levels and the ratings of the second rater. On the other hand, the two teacher raters agreed at a high level. The correlation coefficient was .82 between the raters ( $p < .01$ ). The analysis of ICC and teacher ratings for each problem showed that problems in the induction and insight subcomponents were in conflict. The main reason for the conflict, perhaps, was the fact that these problems were not developed according to C<sup>3</sup> but the author and the mathematicians agreed on the difficulty level of each problem. Therefore, I set out a third phase, during which I revised 7 problems, replaced 3 problems, and omitted 3 additional problems. The teachers' ratings on the problems in the insight and induction subtests also were integrated in the final revision. Then, the mathematicians reviewed the final problems before the test was given to the participants. Table 3.5 shows the contents and the psychological sources of the ICC levels revised after the final phase.

*Item Content and the Use of the Three-Mathematical Minds Model in Item Development*

All problems included in the test were developed to measure some aspects of the three-mathematical minds as seen in table 3.5: knowledge component, analytical component, and creativity component. Each component had three subcomponents. Problems in each subcomponent hypothetically were developed to measure an aspect of one of the minds. The knowledge component consisted of problems with algebra, geometry and statistics subcomponents. The analytical component contained problems of linear, categorical and conditional syllogism subcomponents. The creativity component was made up of problems of induction, insight and selection subcomponents. The discussion of the theoretical and technical background for item development in each component, as well as in each subcomponent follows.

*Knowledge expert.* Studies on expertise provided the theoretical background in the development of knowledge problems. Three branches of mathematics (algebra, geometry and statistics) were used to develop three separate classes of problems. The theoretical purpose of these subtests was to measure factual, relational and schematic knowledge to distinguish between those who demonstrated the knowledge of a novice and those who demonstrated knowledge possessed by experts. *Algebra* problems required knowledge of substitution, transposing, and factoring, as well as that of solving an algebra word problem by translating it into an equation with two unknowns. *Geometry* problems required knowledge of area, perimeter, and of angles as well as knowledge of relationships between area and perimeter. *Statistics* problems required knowledge of rate, percent, interest and data tables.

Table 3.5

Item development according to the three-mathematical minds with the three-level cognitive complexity model

Construct		Cognitive Complexity		
<u>Minds</u>	<u>Subcomponents</u>	<u>Level I</u>	<u>Level II</u>	<u>Level III</u>
Knowledge Expert	Algebra	Factual	Relational knowledge	Conceptual-schematic knowledge
	Geometry	Factual	Relational knowledge	Conceptual-schematic knowledge
	Statistics	Factual	Relational knowledge	Conceptual-schematic knowledge
Analytical	Linear Reasoning	5 elements	5 elements + additions	6 elements + coefficients + divisors

Table 3.5 – continued

Construct		Cognitive Complexity		
<u>Minds</u>	<u>Subcomponents</u>	<u>Level I</u>	<u>Level II</u>	<u>Level III</u>
Analytical	Conditional Reasoning	One condition	Two conditions	Two conditions Double negation
	Categorical Reasoning	Two sets: one superset one subset	Three sets: dissections one superset one subset	Four sets: one intersection two dissections two supersets two subsets
Creative	Discovery – Induction	Free classification	one relation	one rule one relation
	Insight	Team agreement	Team agreement	Team agreement
	Selection	Selective encoding	Selective encoding + Selective combination	Selective encoding + Selective combination + Selective comparison

*Analytical ability.* Works of mathematicians Poincare and Polya and of psychologist Sternberg provided insights in the development of the analytical items. Analytical ability was measured by three deductive subcomponents: linear reasoning, conditional reasoning and categorical reasoning. The theoretical purpose of these subcomponents was to measure analytical mathematical ability to distinguish between novice analysts and expert analysts. Problems in this component were presented in the forms of numerical or geometrical notations or a mixture of both, such as  $A > B$ . In other words, these problems were transformations of typical deductive word problems, such as Mike is taller than Sally, who is shorter than Frank; or Mike is 5 years older than Bob and Bob is 3 years younger than Cathy; or Adam works three times as much as does Bob who works twice less than Cathy. The reasoning behind the use of mathematical notations instead of words in syllogistic problems was that a pilot study undertaken by the author (Sak, 2004) indicated that a problem with algebraic symbols had a better discrimination level in measuring analytical mathematical ability than did the same problem given in a word format.

In *linear syllogism*, two or more quantitative relations were given between each of two pairs of items, depending on the cognitive complexity of the item. One item of each pair overlapped with an item of another pair, such as  $A < B$  and  $B > C$ . The task of the problem solver was to figure out relationships between the nonoverlapping terms, and to verify a logical conclusion. In *conditional syllogism*, one condition was presented with five conclusions, of which only one option satisfied the condition. The following is an example of a conditional syllogism problem: If  $x < 0$ , then: (a)  $x^2 > 2x$ ; (b)  $x^2 < 3x$ ; (c)  $x^2$

$< 0$ ; (d)  $x^2 < x + x$ ; (e)  $x^3 > x^2$ . In *categorical syllogism*, participants had to figure out relationships between members of classes. Group memberships were presented in a table. The following is an example of a categorical word problem: In Ms. Flores' farm, half of the meat-eating animals also can eat plants. Half of the plant-eating animals also can eat meat. Sixteen animals only eat meat. How many animals cannot eat meat?

*Creative Ability.* This component, too, was framed on the works of mathematicians Poincare and Polya and psychologist Sternberg. Creative ability was measured by three subcomponents: Induction, selection and insight. The theoretical purpose of these subtests was to measure the creative mathematical ability to distinguish between novices and creative experts. The following is the review of the subcomponents.

*Induction* problems can be characterized either by mathematical rule discovery or by rule production. In rule discovery problems, the task of the individual was to discover a rule(s) that was constructed by the author between a series of numbers or commonalities between numbers and those between figures, such as side-corner-angle relations. The following was an example of this kind: what is the sum of internal angles of a shape that has 14 sides, given that the sum of internal angles of a triangle is 180, the sum of internal angles of a square is 360 and a pentagon's is 540 degrees. The task of the individual was to figure out the relationships between the number of sides and the sum of the internal angles of the shapes to find the rule. In this case, the rule is  $(n - 2)180$ . In the rule production problem, a set of numbers was presented. An individual had to develop or generate his or her own rule or identify commonalities between the numbers, such as grouping numbers from 1 to 20 based on a rule.

*Selection* problems can be characterized by finding out relevant or irrelevant information, by selective encoding, by selectively combining encoded information and by analogizing combinations to other constructions that are presented in the problem stem or among answer options, as related to the solution of the problem. The theoretical purpose of these problems was to measure selection in problem solving as articulated by Poincare (1952a) and Polya (1954a) and as theorized by Davidson and Sternberg (1984). Three types of selection problems were used. The first kind was *selective encoding*; an individual had to find out relevant or irrelevant information for the solution. The following is an example of this kind: In a classroom, twenty percent of the students passed mathematics; thirty percent passed English; and forty percent passed both. What information is NOT required to find the number of students in the classroom? (a) number of students who passed math; (b) number of students who passed English; (c) number of students who passed both; (d) percentage of students who failed both classes; (e) number of students who failed both classes. The second kind of problem required both *selective encoding* and *selective combination*; the individual first had to encode related elements in a problem and then find out which combination of given information was the correct combination for the solution. The third kind of problem required *selective encoding*, *selective combination*, and *selective comparison*; an individual had to perform selective encoding and selective combination and selectively compare his/her combination to a given construction.

*The insight* subcomponent contained problems of mathematical recreations. The theoretical background of these problems came from Gestalt psychology, according to

which an insight might occur through restructuring thinking and problems or through gestalts. Thus, the theoretical purpose of this subtest was to measure ability to restructure both thinking and problems. The task of an individual was to think in a flexible and unconventional mode or to scatter thinking over the problem to see the big picture. The following was an example of this kind: Write a fraction whose numerator is smaller than its denominator and whose result is larger than its numerator. An individual has to give up thinking only of positive numbers and start to think in negative numbers to come up with a correct answer for this problem, such as  $-8/2 = -4$ .

#### *Levels of Item Cognitive Complexities in Item Development*

Item cognitive complexity refers to psychological sources of item difficulty, such as levels and kinds of knowledge or cognitive processes a problem requires for the solution. The difficulty level of each problem used in this study was developed according to the three-level cognitive complexity model proposed by the author (see table 3.5). Consider the model below.

The three-level cognitive complexity model ( $C^3$ ) was developed psychologically according to a performance continuum on which novice and expert people can be categorized, based on their intellectual performance on some tasks that are related to the domain of mathematics. Figure 3.1 shows the continuum on a triangular shape. The triangle gradually gets more intense in color from the left side to the right side. The intensity of color is a connotation of intellectual complexity, less developed on the left and better developed on the right. In other words, the more an individual moves to the right, the more proficiently s/he performs on intellectual tasks. The continuum also gets

narrower from the left to the right representing the distribution of expertise in a domain. Novices, for example, are in the beginning of developing expertise, as compared to experts. That is to say, they have not been exposed enough to the domain, or could not benefit as much as could experts from the same amount of experience to master necessary skills and knowledge. Their performance is not recognized as superior or exceptional by members of the field in which the domain is located. Because the continuum reflects a developmental performance, people may be found at every point on the continuum. Note that the model was not developed based on a conception of giftedness that favors the measurement of pure capacities (experience free) in mathematics; rather, it was developed based on a conception of giftedness in which giftedness is demonstrated performance of resulting from the interaction of domain experience and developed mathematical skills.

Based on the continuum, therefore, expertise can be measured at different levels. Although I used three levels, novices, developing experts and experts, other levels might be found by dividing the continuum at different points that have psychological meaning. For example, another group of people, as masters, can be added at the highest level. However, the point of dissection should be about where giftedness is scattered as a form of expertise. A psychologically credible answer, probably, is to apply a developmental approach to distinguish among ability groups. That is, people, whose performance is recognized as being superior by members in a field (e.g., scientists, mathematicians or math teachers) compared to their experience peers, can be thought gifted. When this developmental approach is applied in the classification of school children, a good

comparison may be made by comparing students at the same grade level in the same school.

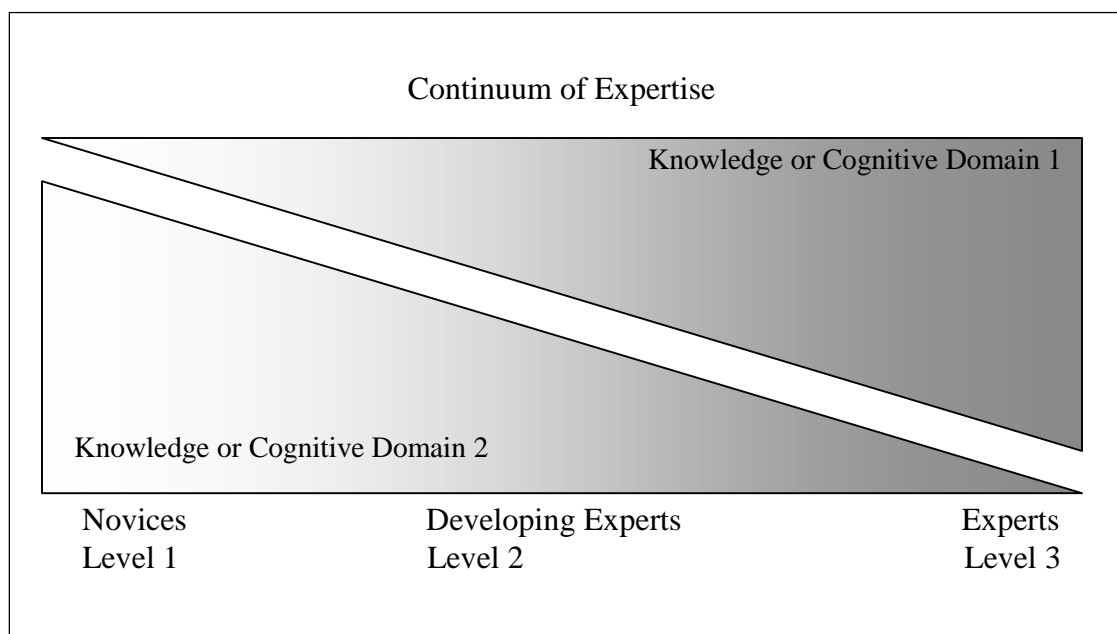


Figure 3.1. Distribution of expertise and knowledge and skills development in experts.

In this study, therefore, a specific ability or a psychological construct was measured at three levels only by three problems. In other words, each subtest had three problems developed to measure a particular psychological construct at a particular cognitive complexity level, such as novices at the first level and experts at the third level. Novices were expected to solve only first level problems while experts were predicted to solve problems at all levels.

The source of cognitive complexity was established on the level of cognitive demands for a particular cognitive ability or for some particular mathematical knowledge that a problem posed to the problem solver. Table 3.5 shows the three-level cognitive complexity model embedded in the three-mathematical minds model. The first level

problems constituted the simplest problems and the third level problems were the most complex ones. For example, a knowledge problem at the first level of cognitive complexity entailed knowing and recalling only one fact for the solution such as the area of a triangle. A problem at the second level required recalling two facts and relating these facts for the solution, such as finding the sum of the internal angles of a triangle given one internal angle and one external angle. A problem at the third level necessitated a conceptual knowledge for the solution, such as drawing two different shapes that had equal areas but unequal perimeters. Notice that level one and level two problems require novice knowledge, isolated facts, while a level three problem requires schematic knowledge; that is the relationship between the area and perimeter of a geometric shape.

The difficulty levels of the problems in each subtest were established in an ascending order. In addition to the knowledge problems mentioned above, here are some sample linear syllogistic problems from the test battery in the form of algebraic notations. They are similar in structure and in purpose to measure linear syllogism, but they are different in their cognitive complexities (see appendix A for the other problems' characteristics):

1.  $A > B$ ;  $C > D$ ;  $D > E$ ;  $E > A$ . Which is the second largest?
2.  $a = b + 1$ ;  $c = d + 3$ ;  $a = f + 3$ ;  $b = d + 2$ . Which is the smallest?
3.  $A = 3B$ ;  $C = 2D$ ;  $F = G/2$ ;  $D = A/2$ ;  $2D = 3G$ . Which is the smallest?

Notice that the first problem is the easiest one because it requires a single pairwise comparison of five elements. This problem has no such quantifiers as coefficients, additions or divisors. It involves encoding of five elements and inferring relationships

between each pair and, then, comparing each pair to the other pairs. The second problem, on the other hand, has one additional element along with additions such as  $a = b + 1$ .

This problem is more difficult than the first one because it requires the employment of more encodings and inferences at a more abstract level because it also requires comparing and eliminating additions. Obviously, the addition of quantifiers, such as plus one or two, to the elements, puts more demands on processes of abstract thinking and on those of working memory. The third problem is more difficult than the second one because each element in the third problem has a coefficient or a divisor while the second problem only has additions. The third problem also has one additional element. Note that processing multiplication and division is more difficult than processing addition from the cognitive developmental perspectives.

#### *Item Format*

Two types of item formats were used. Most items were presented in a multiple-choice format consisting of a problem stem and five answer options. Only one option was correct in these types of problems. The second format was open problems – more than one method and solution or more than one method but only one solution was accepted as correct.

#### *Item Scoring*

One point was given for each correct answer in both multiple choice and open problems. No point reduction was taken for wrong answers.

### *Test Administration*

Mathematics teachers administered the test during students' regular mathematics classes in the beginning of the spring semester of the 2004-2005 school-year. The testing was done in one sitting, taking about 45 minutes. The teachers read standard instructions before the testing (Appendix C).

### *Data Analysis*

#### *Research Question 1*

This research question was about associations between different types of mathematical ability. Correlational analyses were used to examine associations between the hypothesized mathematical abilities. Exploratory factor analysis was used to examine the nature of mathematical ability.

#### *Research Question 2*

This research question was related to psychometric properties of the M<sup>3</sup> test battery. Kuder-Richardson reliability analysis was performed to answer the reliability question. Bivariate correlation coefficients were determined to analyze the convergent validity of the test battery when students' rating of their liking of mathematics, their rating of their own mathematical ability and teachers' rating of students' mathematical ability were used as converging variables. Point biserial correlations were used to examine item-total test and item-subtest relationships to provide information about the construct validity of the M<sup>3</sup>. Analysis of Variance (ANOVA) was used to examine performance differences of students at different grade levels.

### *Research Question 3*

This research question was developed to determine which M<sup>3</sup> items were good measures of mathematical knowledge, analytical mathematical ability, and creative mathematical ability. This research question was investigated to provide additional information about the construct validity of the M<sup>3</sup>. In this analysis, items constituted cases, just like individuals. Each item had a continuous discrimination index in each component. General Linear Modeling (GLM) Repeated Measures were used to analyze differences between discrimination indices derived from component scores, as well as from the composite score.

### *Research Question 4*

This research question was about associations between item cognitive complexity (ICC), item difficulty (ID) and item discrimination (D). The underlying assumption was that ICC is the major source of item difficulty, which accounts for variance in item discrimination. A Standard Regression Analysis was used to answer this research question and to test an associated hypothesis, as well as bivariate correlations to explore associations among the ICC, ID and D.

### *Research Question 5*

This research question was related to discrimination powers of items of different cognitive complexity. The underlying assumption was that gifted students (above 95% of the total participants), as identified by the composite score, outperform the rest of the participants in different degrees on items of different cognitive complexity. A nonparametric Chi-Square analysis was used to analyze differences in the proportions of

the groups passing at each item. A MANOVA was used performance differences of the ability groups on three levels of cognitive complexity.

### *Item Analysis*

Point biserial item-total test and item-subtest correlational analyses were used to analyze whether the M<sup>3</sup> items measured mathematical ability in the direction the entire test battery measured. In addition to point biserial correlations, the classical item discrimination model was used to explain item discrimination. I applied the classical mathematical model to different ability groups, with a focus on comparing the gifted group to nongifted groups to find out which items were good measures of mathematical knowledge, creative mathematical ability and analytical mathematical ability. Item discrimination looks at the proportion passing an item based on total test score. In this study, I estimated item discrimination indices based on the composite score. Also, the performance of the upper and the lower percentiles were compared on the factor scores to investigate further validity evidence at the item level. Table 3.6 shows how this comparison can be made. The capital letters represent different ability levels on a percentile scale while composite, expertise, analytical and creativity scores indicate types of abilities.

Table 3.6

The estimation of item discrimination indices based on different levels of different types of mathematical ability

Item	Item discrimination estimated based on											
	Composite			Expertise			Analytic			Creativity		
	<u>*A</u>	<u>**B</u>	<u>***C</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>A</u>	<u>B</u>	<u>C</u>	<u>A</u>	<u>B</u>	<u>C</u>
1												
2												
3												
...												

Note. \*the comparison of the upper 27% and the lower 27%; \*\*the comparison of the upper 85% and 50 to 84%; and \*\*\*the comparison of the upper 95% and 84%.

## CHAPTER IV

### RESULTS

This chapter includes findings for the research questions, which were investigated in this study. First, each research question will be reintroduced in the order in which it was introduced in the first chapter, and research findings will be presented.

#### Research Question 1

*How theoretically valid is the three-mathematical minds model ( $M^3$ )?*

*a. What is the underlying structure of mathematical ability?* As stated in the previous chapter, the  $M^3$  items were hypothetically grouped into nine categories; therefore, a subtest-level factor analysis was used to answer this research question. The nine subtests of the  $M^3$  were subjected to principal components analysis (PCA) to examine the theoretical validity of the three-mathematical minds model. Prior to performing PCA, the suitability of data for factor analysis was assessed. Inspection of the correlation matrix revealed the presence of coefficients of .3 and above. The Kaiser-Meyer-Olkin value for sampling adequacy was .81, exceeding the recommended value of .6 (Kaiser, 1974). The Barlett's Test of Sphericity reached statistical significance ( $p < .001$ ), which supported the factorability of the correlation matrix.

The initial, unrotated PCA of the nine subtests revealed the presence of three components with eigenvalues exceeding 1. As seen in table 4.1, Component 1 explained 31.31% of the total variance; Component 2 explained 12.32%; and Component 3 accounted for 11.40% of the total variance respectively. Also, table 4.1 shows initial factor loadings with .30 or above values for the unrotated PCA. Because only three

components had an eigenvalue above 1, these three components were investigated further, using Varimax rotation with Kaiser Normalization. Factors that had less than .40 absolute values were not reported in the rotated solution. The three factor solution explained 55.03% of the variance, with Component 1 contributing 29.30%, Component 2 contributing 13.58% and Component 3 contributing 12.16%. These components were labeled as follows: the knowledge-reasoning component, the creativity component and the analytical component. Table 4.3 shows the rotated factor solution for subtests and their loadings for each component. According to the three factor solution and subtest loadings, geometry, algebra, statistics, linear syllogism, conditional syllogism, and induction subtests were assigned in the first, the knowledge component; the selection subtest were assigned in the second, the creativity component; and categorical syllogism was assigned in the third, the analytical component. Because the insight component loaded equally on the first and second components, it was not assigned in a component by this author.

Although the factor analytic findings partially supported the three-mathematical minds model yielding three separate components, four subtests did not fit into the components for which they were developed. The induction subtest was developed to measure an aspect of the creative mind, but this subcomponent was found in the knowledge-reasoning component. Conditional and linear syllogism subtests were developed to measure the analytical mind, but they too were found in the knowledge-reasoning component. Finally, the insight subtest was developed one aspect of the

creative mind, but it loaded equally on the Knowledge component and the Creativity component.

Table 4.1

Unrotated component matrix

Subtest	Component		
	1	2	3
Geometry	.778		
Conditional Syllogism	.700		.310
Algebra	.612		
Statistics	.594		
Linear Syllogism	.593		-.309
Insight	.550		.417
Induction	.546		-.339
Categorical Syllogism		.794	
Selection		.568	.662
Total Variance Explained	%31.31	%12.32	%11.40

Table 4.2

Total variance explained by each component

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2.82	31.31	31.31
2	1.11	12.32	43.62
3	1.03	11.40	55.02
4	.82	9.1	64.11
5	.76	8.43	72.55
6	.73	8.15	80.70
7	.71	7.92	88.62
8	.56	6.20	94.81
9	.47	5.19	100.00

Table 4.3

Varimax rotation for the three factor solution for the M<sup>3</sup>

Subtest	Component		
	1	2	3
Geometry	.74		
Linear Syllogism	.65		
Algebra	.63		
Statistics	.63		
Induction	.58		
Conditional Syllogism	.58		
Selection		.85	
Insight	.43	.44	
Categorical Syllogism			.83
Total Variance Explained	%29.30	%13.58	%12.16

*b. How are the subcomponents of the  $M^3$  (hypothetically constructed as knowledge of algebra, of geometry and of statistics; as linear syllogism, conditional syllogism and categorical syllogism; and as induction, insight, and selection) associated?*

The relationships among the subcomponents of the  $M^3$  were investigated using Pearson Product-Moment correlation coefficient. Preliminary analyses were performed to check the assumptions of normality, linearity and homoscedasticity. Algebra, statistics and induction subtests distributed normally. Geometry, linear syllogism, conditional syllogism, categorical syllogism, insight and selection subtests were positively skewed. That is, scores were scattered around low performance (see Appendix A).

Table 4.4 shows a number of statistically significant and nonsignificant correlations among the subtests, with the highest associations among the knowledge subtests, linear syllogism, conditional syllogism, insight and induction subtests. Categorical syllogism and selection subtests had the lowest associations with the other subtests. In fact, the categorical subtest did not have any significant correlations with any other subtests, and the selection subtest had a statistically significant, but low correlation only with the induction subtest ( $r = .16$ ;  $p < .05$ ). The disassociations of the latter two subtests also were substantiated with factor analysis as reported in the previous research question.

*c. How are the components of the  $M^3$  (theorized as mathematical expertise, analytical mathematical ability, and creative mathematical ability) associated?*

Although factor analysis did not reveal three subtests in each component, the nine subtests were grouped in three components rationally, but not factor analytically, to

examine relationships among the three minds. As discussed in Chapter 3, algebra, geometry and statistics subtests were grouped in the Knowledge component; linear syllogism, conditional syllogism and categorical syllogism subtests were grouped in the analytical component; and selection, insight and induction subtests were grouped in the creativity component. As seen in table 4.4, statistically significant correlations existed among the components ranging from middle to middle-high correlations. The knowledge component had the highest correlations with the other components and the M<sup>3</sup> composite (.49, .44, .84,  $p < .01$ ). The creativity and analytical components had lower correlations with each other compared to the knowledge component.

Table 4.4

Bivariate correlations among all M<sup>3</sup> components

Component	Analytical	Creative	M <sup>3</sup> composite
Knowledge	.49**	.44**	.84**
Analytical		.41**	.82**
Creative			.72**

\*\* Correlation is significant at the 0.01 level (2-tailed).

Table 4.5

Bivariate correlations among all M<sup>3</sup> subcomponents

Variables	Geo.	Stat.	L. Syllo.	Co. Syllo.	Ca. Syllo.	Select.	Induct.	Insight	M <sup>3</sup> comp.
Algebra	.35**	.30**	.28**	.34**	.05	.03	.24**	.21**	.57**
Geometry		.36**	.42**	.47**	.03	.09	.34**	.36**	.73**
Statistics			.25**	.26**	.01	.03	.26**	.24**	.57**
L. Syllogism				.26**	.08	.02	.25**	.17**	.59**
Co. Syllogism					.03	.16**	.26**	.38**	.67**
Ca. Syllogism						.10	.13*	.00	.32**
Selection							.06	.07	.30**
Induction								.12*	.55**
Insight									.50**

Note. \*\* Correlation is significant at the 0.01 level (2-tailed). \* Correlation is significant at the 0.05 level (2-tailed).

## Research Question 2

*What are the psychometric properties of the M<sup>3</sup> test battery?*

*a. How reliable is the M<sup>3</sup>?* The purpose of this research question was to examine the consistency of scores in the M<sup>3</sup>. Kuder-Richardson reliability analysis was carried out to investigate the reliability of the M<sup>3</sup>. The analysis showed a .72 coefficient level, slightly exceeding the minimum desired level .70 for consistency of scores for psychological tests.

*b. What is the convergent validity of the M<sup>3</sup> when students' liking of mathematics, their rating of their own mathematical ability and teachers' rating of students' mathematical ability are used as converging variables?* This research question was investigated to examine if the M<sup>3</sup> correlated with other variables with which it theoretically should correlate. Partial correlation coefficients were computed, while grade was controlled in the equation, to investigate the relationships between teachers' rating of students' mathematical ability, students' rating of their own mathematical ability and their liking of mathematics and the M<sup>3</sup> subcomponents (see table 4.6 for bivariate correlations among student variables). Correlations ranged from low to high-medium, with the majority of correlations being statistically significant, as seen in table 4.7. This finding provided partial evidence for the convergent validity of the M<sup>3</sup>. Particularly important were the correlations between the M<sup>3</sup> composite and the other variables. These correlations, medium to high-medium were as follow: .45 between the M<sup>3</sup> composite and teachers' rating ( $p < .01$ ), .36 between the M<sup>3</sup> composite and students rating of their own ability ( $p < .01$ ) and .35 between the M<sup>3</sup> composite and students' rating of their liking of

mathematics. Teachers' rating of students' mathematical ability had higher correlations with the M<sup>3</sup> and its subcomponents than students' rating of their own mathematical ability and their liking of mathematics, as seen in table 4.7.

Table 4.6

Bivariate correlations among all student variables

Variable	Age	Teacher Rating	Student Liking	Student Rating
Grade	.84**	-.13	-.07	-.17**
Age		-.14	-.13*	-.21**
Teacher Rating			.58**	.50**
Student Liking				.55**

Note. \*\*  $p < .01$ , \*  $p < .05$  (2-tailed)

*c. Does the M<sup>3</sup> differentiate among students of various grade levels?* The purpose of this research question was to examine whether the M<sup>3</sup> showed developmental variance. Levene's test for homogeneity of variances in scores of the groups showed no violation ( $p = .077$ ;  $p$  is required to be greater than .05). A one-way between-groups analysis of variance (ANOVA) was conducted to inspect performance differences of 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grade students on the M<sup>3</sup> composite (see table 4.8 for mean performance for each grade level).

Table 4.7

Partial correlations between student variables and the M<sup>3</sup> components and subcomponents

Component	Grade	Age	Teacher Rating	Student Liking	Student Rating
Algebra	.02	-.03	.26**	.20**	.21**
Geometry	.16**	.11	.33**	.25**	.33**
Statistics	.06	-.02	.38**	.24**	.34**
Linear Syllogism	.14*	.07	.23**	.11	.08
Conditional Syllogism	.13*	.05	.31**	.24**	.27**
Categorical Syllogism	.03	.02	.09	.16**	.01
Selection	.05	.01	.13	.07	.07
Induction	.13*	.04	.29**	.19**	.22**
Insight	.16**	.10	.25**	.18**	.19**
Knowledge	.11	.03	.42**	.32**	.40**
Analytical	.15**	.07	.32**	.26**	.19**
Creative	.18**	.07	.35**	.24**	.26**
M <sup>3</sup> Composite	.18**	.07	.45**	.35**	.36**

Note. The effect of grade was removed. \*\* Correlation is significant at the 0.01 level (2-tailed). \* Correlation is significant at the 0.05 level (2-tailed).

The analysis indicated statistically significant differences among the grades ( $F[2, 288] = 7.5, p < .001$ , see table A.7). The effect size, calculated using eta squared, was .05, a medium effect. Post-Hoc tests using Tukey Honestly Significant Difference (HSD) revealed the following mean differences among the grades (table A.8). Eighth graders performed significantly higher than 7<sup>th</sup> and 6<sup>th</sup> graders ( $p < .01$  and  $p < .05$ , respectively). This finding provided developmental evidence for the construct validity of the M<sup>3</sup>. Although no significant performance difference existed between 7<sup>th</sup> and 6<sup>th</sup> graders on the M<sup>3</sup> composite, sixth graders performed slightly higher than 7<sup>th</sup> graders ( $p = .89$ ). This finding shows a contradiction for the developmental evidence obtained in the previous finding.

*d. What is the internal consistency of the M<sup>3</sup> for item-total score, item-subtest and subtest-total score correlations?* In this research question, the purpose was to analyze the degree to which test items and subtests were homogenous or heterogeneous. In other words, an attempt was made to determine whether the M<sup>3</sup> items measured a unified construct or multiple constructs. The degree of homogeneity or heterogeneity of a test has some relevance to its construct validity. Because the M<sup>3</sup> is a measure of multilateral aspects of mathematical ability, an investigation of item homogeneity and heterogeneity in the M<sup>3</sup> provides information about its construct validity. Point biserial correlations between the items and the M<sup>3</sup> composite and between the items and the subtests were computed. As seen in table 4.9, correlations between the items and the M<sup>3</sup> composite ranged from low to high, with all correlations being statistically significant except for item 26. It had a very low and statistically nonsignificant correlation with the M<sup>3</sup>

Table 4.8

Mean and standard deviation for student variables and the M<sup>3</sup> subcomponents by grade

Variable	Grade							
	6		7		8		Total	
	M	SD	M	SD	M	SD	M	SD
Teacher Rating	4.00	.95	3.45	1.08	3.50	1.41	3.59	1.23
Student Liking	3.06	1.57	2.87	1.30	2.78	1.43	2.88	1.41
Student Rating	3.89	.96	3.39	1.00	3.36	1.08	3.49	1.04
Algebra	1.60	.66	1.50	.71	1.61	.72	1.55	.70
Geometry	.67	.90	.72	.86	1.01	.90	.82	.90
Statistics	1.11	.86	1.08	.80	1.21	.79	1.13	.81
Linear Syllogism	.73	.87	.76	.72	1.01	.90	.85	.83
Conditional Syllogism	.81	.82	.89	.83	1.09	.99	.95	.89
Categorical Syllogism	.82	.92	.62	.79	.85	.94	.75	.88
Selection	.52	.72	.49	.62	.59	.67	.54	.66
Induction	.90	.79	.91	.66	1.11	.71	.99	.71
Insight	.32	.56	.27	.47	.53	.69	.38	.59
M <sup>3</sup> Composite	7.49	3.99	7.23	3.26	9.03	3.94	7.98	3.77

composite ( $r = .06$ , respectively). Though the correlations were significant, items 2, 6, and 15 correlated with the M<sup>3</sup> composite at low levels.

The second mode of analysis was point biserial correlation computed between the items and the subtests to examine the homogeneity and heterogeneity of the items. In other words, I assumed that an item was supposed to correlate highly with the subtest in which it was located to show homogeneity and to correlate at a low level with a subtest in which it was not located to show heterogeneity. As seen in table 4.9, correlations ranged from low negative to very high positive correlations. For example, the correlation coefficient was  $-.12$  between item 26 and the subtest selection in which it was not located ( $p < .05$ ). The correlation was  $.88$  between item 7 and the subtest insight in which it was located ( $p < .01$ ). What should be read from table 4.9 are correlations between the items and the subtests in which they are located. Item 7 had a high or very high correlation with the subtest in which they were located; whereas, item 26 had a very low to medium correlation with the subtest in which these items were not located. Only item 24 had a medium correlation with the statistics subtest in which it was located, and item 26 had a medium level correlation with the insight subtest in which it was located; however, the correlations still were statistically significant ( $p < .01$  for both items). As seen in the table, the pattern of correlations provided partial support for both homogeneity and heterogeneity of the M<sup>3</sup> because many items also had significant correlations with the other subtests in which they were not located; however, these correlations were low. In addition, Appendix A provides inter-item correlations showing associations among the items.

Table 4.9

Item-subtest and item-total test point biserial correlations

Item	Algebra	Geometry	Statistics	Linear	Cond.	Categor.	Selection	Induction	Insight	Total
				Syllo.	Syllo.	Syllo.				
1	.24**	.38**	.29**	.67**	.17**	.08	.01	.24**	.08	.46**
2	.47**	.08	.13*	.05	.06	.00	-.09	.11*	.04	.17**
3	.24**	.29**	.19**	.21**	.72**	-.06	.01	.16**	.29**	.44**
4	.20**	.76**	.29**	.30**	.35**	-.08	.03	.23**	.29**	.51**
5	.27**	.34**	.66**	.29**	.31**	.01	.07	.23**	.30**	.51**
6	-.03	.05	.02	-.02	.09	.04	.69**	.02	-.01	.16**
7	.15**	.29**	.22**	.13*	.31**	-.05	.08	.19*	.88**	.41**
8	.24**	.28**	.29**	.28**	.29**	.09	-.01	.69**	.12*	.47**
9	.14**	.28**	.16**	.16**	.61**	-.01	.10	.08	.21**	.37**
10	.17**	.19**	.12*	.63**	.21**	.00	.01	.09	.22**	.34**

Table 4.9 – *continued*

Item	Algebra	Geometry	Statistics	Linear	Cond.	Categor.	Selection	Induction	Insight	Total
				Syllo.	Syllo.	Syllo.				
11	.10	.18**	.11*	.10*	.23**	.13*	.08	.51**	.15**	.33**
12	.28**	.75**	.20**	.35**	.39**	.08	.06	.22**	.27**	.56**
13	.18**	.25**	.67**	.15**	.20**	-.04	-.03	.21**	.19**	.37**
14	.13*	.14**	.04	.09	-.02	.74**	.06	.11*	.05	.29**
15	-.03	-.06	-.05	.01	-.06	.79**	.03	.06	-.02	.15**
16	.00	-.01	.03	.09	.17**	.51**	.13*	.12*	-.06	.21**
17	.80**	.36**	.26**	.33**	.34**	.05	.10*	.27**	.23**	.56**
18	.03	.09	.05	.04	.11*	.11*	.54**	.05	.11*	.22**
19	.05	.01	-.03	.03	.08	.03	.51**	.04	.03	.14**
20	.10	.21**	.06	.56**	.11*	.08	.02	.13*	.03	.28**
21	.24**	.30**	.19**	.20**	.29**	.07	.08	.11*	.61**	.42**

Table 4.9 – *continued*

Item	Algebra	Geometry	Statistics	Linear	Cond.	Categor.	Selection	Induction	Insight	Total
				Syllo.	Syllo.	Syllo.				
22	.28**	.35**	.17**	.14*	.62**	.12*	.20**	.25**	.24**	.50**
23	.54**	.13*	.14**	.05	.16**	.04	-.03	.02	.09	.23**
24	.13*	.15**	.38**	.05	.07	.16**	.12*	.13*	-.02	.25**
25	.07	.15**	.04	.05	-.04	.03	.05	.58**	-.04	.18**
26	.02	.08	-.03	-.03	.09	.01	-.12*	-.06	.36**	.06
27	.26**	.56**	.27**	.21**	.24**	.11*	.13*	.30**	.18**	.49**

Note. \* Correlation is significant at the 0.05 level. \*\* Correlation is significant at the 0.01 level.

### Research Question 3

*Which M<sup>3</sup> items represent good measures of mathematical knowledge, analytical mathematical ability and creative mathematical ability?*

The purpose of this research question was to examine which M<sup>3</sup> items were good measures of mathematical knowledge, creative mathematical ability and analytical mathematical ability. This type of analysis provides evidence for item homogeneity and heterogeneity in construct validity. Item discrimination analysis was used to examine item characteristics. The index for item discrimination was computed based on a comparison of the performance of the upper 25<sup>th</sup> percentile group and that of the lower 25<sup>th</sup> percentile group on the M<sup>3</sup> composite, as well as on the components.

Table 4.10 shows discrimination indices for each item computed based on the M<sup>3</sup> composite and its components as well as mean indices and standard deviations for each component. Discrimination indices ranged from -.02 (negative discrimination) to .84 (very high positive discrimination). These indices mean that some items were homogenous, measuring similar constructs, whereas some items were heterogeneous, measuring separate constructs. The M<sup>3</sup> composite had a mean of .41 discrimination index and the knowledge component had a mean of .30 discrimination index. Both indices were moderate levels of discrimination for general mathematical ability, as measured by the composite. The analytical and creativity components had low levels of discriminations for general mathematical ability as measured by the composite (.29 and .24, respectively).

Table 4.10

Item difficulty, item cognitive complexity and item discriminations

Item	ID	ICC	D	D	D	D
			Composite	Knowledge	Analytical	Creative
1	.36	1	.62	.48	.59	.15
2	.93	1	.12	.13	.06	.02
3	.42	1	.60	.44	.57	.27
4	.41	1	.77	.79	.39	.32
5	.58	1	.80	.74	.40	.35
6	.27	1	.19	.05	.10	.41
7	.30	1	.58	.39	.25	.61
8	.67	1	.64	.48	.43	.45
9	.20	2	.44	.28	.40	.20
10	.23	2	.41	.22	.41	.16
11	.09	2	.27	.14	.17	.27
12	.30	2	.76	.68	.47	.30
13	.34	2	.48	.59	.18	.22
14	.31	1	.39	.14	.49	.12
15	.30	2	.14	-.01	.47	.08
16	.15	3	.21	-.02	.31	.09

Table 4.10 – *continued*

Item	ID	ICC	D	D	D	D
			Composite	Knowledge	Analytical	Creative
17	.51	2	.84	.82	.45	.37
18	.17	2	.26	.06	.09	.29
19	.10	3	.12	-.01	.08	.22
20	.26	3	.36	.15	.36	.09
21	.06	2	.30	.18	.15	.20
22	.33	3	.72	.42	.58	.39
23	.13	3	.19	.22	.05	.02
24	.15	3	.27	.25	.11	.13
25	.22	3	.20	.10	.01	.41
26	.02	3	.03	.02	.03	.03
27	.11	3	.40	.32	.24	.27
<i>Mean</i>	.29	2	.41	.30	.29	.24
<i>SD</i>	.20	.83	.24	.26	.19	.15

Note. Item difficulty range: .80-1.00 very easy, .60-.79 easy, .40-.59 moderately difficult, 20-39 very difficult and .00-.19 extremely difficult. Item discrimination range: .50 and above high, .30-.49 moderate, .15-.29 low, .00-.15 very low and negative.

A second mode of analysis was performed using the General Linear Modeling (GLM) Repeated Measures. Repeated Measures were used to analyze further if statistically significant differences existed between discrimination indices derived from component scores, as well as from the composite score. Four levels of within-subjects factors (composite, knowledge, analytical and creativity) were defined for use in the GLM Repeated Measures. Prior to the GLM, Mauchly's test of sphericity was performed to inspect the homogeneity of variance-covariance matrices to assure the validity of the F statistic for use in the univariate test. No violation existed ( $p < .001$ ). The GLM Repeated Measures indicated a statistically significant difference among the measures ( $F(3, 24) = 35.65, p < .001$ ; Wilks' Lambda = .18; partial eta squared = .82). The effect size of .82 is very large. This finding suggests that not all items were homogeneous, and some items were good or poor measures of multilateral aspects of mathematical ability. Findings related to the items will be discussed in the next chapter.

#### Research Question 4

*How psychologically valid is the three-level cognitive complexity model ( $C^3$ )?*

The purpose of this research question was to explore associations among item cognitive complexity (ICC), item difficulty (ID) and item discrimination (D). The underlying assumption was that ICC was the major source of item difficulty. Bivariate correlation was used to explore associations among ICC, ID and D. Standard Regression Analysis was used to analyze further this research question and to test an associated hypothesis.

a) *What relations, if any, exist among item cognitive complexity (ICC), item difficulty (ID) and item discrimination (D)?* The relationships among ICC, ID and D were investigated using the Pearson Product-Moment Correlation Coefficient. As seen in table 4.11, ICC had a high and statistically significant correlation with ID ( $r = .64$ ,  $p < .01$ ), which provided evidence for the validity of the three-level cognitive complexity model. Meanwhile, ICC had medium and statistically significant correlations with the discrimination indices except that the correlations between ICC and D computed based on the creativity and analytical components were not statistically significant.

Table 4.11

Bivariate correlations among ICC, ID and D

Variable	ICC	ID	D- Composite	D- Knowledge	D- Analytical
Item Difficulty (ID)	.64**				
D-Composite	.43*	.43*			
D-Knowledge	.39*	.46*	.93**		
D-Analytical	.37	.32	.72**	.49**	
D-Creativity	.33	.18	.55**	.45*	.15

Note. ICC was reversed. \*\* Correlation is significant at the 0.01 level (2-tailed). \* Correlation is significant at the 0.05 level (2-tailed).

Moreover, ID had high-medium and statistically significant correlations with the discrimination indices computed based on the composite and the knowledge component ( $r = .43, p < .05$ ;  $r = .46, p < .05$ ); while it did not have statistically significant correlations with the discrimination indices on the analytical and creativity components ( $r = .32, .18, p > .05$  for both correlations). High or medium correlations found in this study between ID and D did not provide support for the validity of the three-level cognitive complexity model. That is to say, a high correlation between item difficulty and item discrimination is not desirable.

*b. Hypothesis – ICC significantly predicts ID.* A Standard Regression Analysis was performed to test this hypothesis. ICC was the independent variable as predictor and ID was the dependent variable. As reported in the foregoing research question, ICC accounted for 41% of the variance in ID ( $R^2 = .408, p < .001$ ). As table 4.12 shows respective values of standardized and unstandardised Beta values, ICC had a statistically significant contribution to explain the variance in ID. Therefore, this hypothesis was accepted.

Table 4.12

Summary of Standard Regression Analysis for ICC predicting ID

Variable	Unstandardised	Standardised	Standard Error	Significance
	Beta	Beta		
ICC	15.49	.64	3.73	.000

### Research Question 5

*How do the three ability groups, gifted (above 95%), above average (85-94%) and average and below-average (below 85%) as identified by the composite score differ in their performance on the items at different levels of cognitive complexity?*

This research question was investigated to determine if each level of the cognitive complexity model (C<sup>3</sup>) discriminated between different ability groups. To analyze data, the items were put in three groups, level 1, level 2 and level 3, based on their cognitive complexity levels (as seen in table 4.13). Participants also were categorized in three groups according to their performance on the total score: gifted (above 95%), above average (85-94%) and average and below average (below 85%). Note that, grouping the participants based on their total scores on the M<sup>3</sup> might contribute to performance difference among the groups on each cognitive complexity level; however, an analysis of group differences on each item and each complexity level might provide additional information about discrimination and difficulty characteristics of items.

A MANOVA was used to investigate performance differences of the three ability groups on the three levels. Preliminary assumption testing was conducted to check normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity. Levene's test indicated inequality of error variance in the level 1 and level 3 groups ( $p < .01$  for both). Therefore, an alpha level of .01 was set for determining significance levels for variables. The analyses indicated a statistically significant difference among the ability groups on the combined dependent variables ( $[F(6, 572) = 68.42, p < .01; \text{Wilks' Lambda} = .34; \text{partial eta squared} = .42]$ ).

Then, Post-Hoc tests were conducted using the Tukey Honestly Significant Difference test (HSD) to investigate the following hypotheses (table 4.14):

*Null hypothesis – No significant difference exists between the performance of gifted students and that of above average students on items at the third level of cognitive complexity only.* Post-hoc comparisons indicated that gifted students scored significantly higher than above average students on the level three items ( $p < .01$ ; see table 4.13 for post-hoc comparisons and table 4.13 for group means). However, gifted students also scored significantly higher than above average students on the level-two items ( $p < .01$ ) but not on the level-one items. Therefore, this null hypothesis was rejected.

*Null hypothesis – No significant difference exists between the performance of above average students and that of average and below average students on items at the second and third level of the cognitive complexity.* Post-hoc comparisons indicated that above average students scored significantly higher than average and below-average students did on the level-one problems ( $p < .01$ ), as well as on the level-two and level-three problems ( $p < .01$  for both differences). Therefore, this null hypothesis was rejected.

*Null hypothesis – No significant difference exists among the performance of the three ability groups on items at the first level of the cognitive complexity.* Post-hoc comparisons indicated that gifted students scored significantly higher than average and below-average students on the level-one problems ( $p < .01$ ). Similarly, above-average students scored significantly higher than average and below-average students on level-one problems ( $p < .01$ ). Therefore, this null hypothesis was rejected. Table 4.14 shows

that only one nonsignificant difference among the groups existed: the one between the gifted and above average group on the first level items ( $p > .05$ ).

A Chi Square test was used to examine differences in the proportions of each ability group passing each item. Table 4.15 displays proportions of the ability groups passing each item and the Chi Square test results performed among group differences. According to Chi Square test, only items 2, 6, 15 and 19 did not differentiate significantly among the three ability groups. Items 2 and 6 are first level problems. Item 15 is a second level problem. Item 19 is a third level problem.

Table 4.13

Mean and standard deviation of performance of all ability groups on three levels of the C<sup>3</sup>

		Ability Group		
		Gifted	Above Average	Below Average
Level		N = 17	N = 42	N = 232
Level 1	Mean	7.23	6.42	3.63
	Std Deviation	.75	1.03	1.58
Level 2	Mean	5.64	3.71	1.67
	Std Deviation	1.27	1.38	1.17
Level 3	Mean	4.00	2.31	1.12
	Std Deviation	1.54	1.33	1.07

Table 4.14

Tukey HSD Multiple Comparisons

Complexity			Mean Group	Std.	
level	(I) Group	(J) Group	Difference (I-J)	Error	p
Level 1	Gifted ( $\geq 95\%$ )	Above Average	.81	.42	.140
		Average-Below	3.60	.37	.000
		Average ( $< 85\%$ )			
		Average			
Level 2	Above Average (85-94%)	Average-Below	2.79	.24	.000
		Average			
		Average	1.93	.34	.000
		Average-Below	3.97	.30	.000
Level 3	Above Average	Average	2.04	.20	.000
		Average			
		Average	1.69	.32	.000
		Average-Below	2.88	.28	.000
Level 3	Above Average	Average			
		Average-Below	1.19	.19	.000

Table 4.15

Chi Square test for group differences on each item and the proportion of the ability groups passing each item

Item	Ability Group			Chi Square Test	
	Gifted %	Average, Above      Below-		Pearson Chi- Square value	Asymp. Sig (2-sided)
		Average	Average		
		%	%		
1	94.1	59.5	28.0	41.19	.000
2	94.1	97.6	92.2	1.63	.442
3	76.5	78.6	32.8	29.51	.000
4	94.1	78.6	30.2	55.62	.000
5	100.0	90.5	48.7	38.64	.000
6	41.2	33.3	24.6	3.29	.192
7	64.7	71.4	20.3	54.30	.000
8	94.1	92.9	60.3	23.00	.000
9	58.8	35.7	14.7	26.35	.000
10	47.1	50.0	15.9	29.63	.000
11	35.3	21.4	5.3	25.68	.000

Table 4.15 – *continued*

12	100.0	64.3	19.0	76.27	.000
13	94.1	45.2	28.0	33.25	.000
14	64.7	40.5	26.3	13.27	.001
15	47.1	31.0	28.0	2.81	.256
16	41.2	19.0	12.1	11.37	.003
17	88.2	90.5	40.9	45.00	.000
18	52.9	16.7	14.7	16.32	.000
19	17.6	14.3	8.2	2.86	.240
20	58.8	33.3	20.0	12.70	.002
21	41.2	16.7	1.3	56.26	.000
22	94.1	61.9	22.8	55.70	.000
23	47.1	11.9	10.3	19.27	.000
24	35.3	23.8	11.6	10.22	.006
25	47.1	26.2	19.8	7.19	.027
26	0.0	7.1	1.3	6.41	.041
27	58.8	33.3	3.4	74.65	.000

### Summary of Findings

The exploratory factor analysis yielded three separate components partially supporting the three-mathematical minds model. The three-factor solution explained 55.03% of the variance. Some subtests did not fit in the components in which they were expected to cluster. The categorical syllogism and selection subtests did not correlate substantially with the other subtests. This indicates low association with other thinking skills, whereas the other subtests had statistically significant correlations with each other.

Regarding the M<sup>3</sup> reliability, the Kuder Richardson analysis showed that the M<sup>3</sup> test had a .72 coefficient level as a consistency of scores. The convergent validity analysis showed that the M<sup>3</sup> had medium to high-medium correlations with teachers' rating of students' mathematical ability and students' rating of their own ability and their liking of mathematics. Another mode of analysis involved developmental differences among the 6<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> graders. Eighth grade students performed significantly higher than the other two groups, while 6<sup>th</sup> graders scored slightly higher than the 7<sup>th</sup> graders. Most of the items in the M<sup>3</sup> had medium to high correlations with the components and the composite with the exception of one of the twenty-seven items.

Item discrimination indices ranged from -.02 (negative discrimination) to .84 (very high positive discrimination) depending on the component or the composite score on which item discrimination was computed. GLM Repeated Measures indicated a statistically significant difference among the four types of discrimination indices. Item difficulty analysis indicated that the difficulty level of the items ranged from .02 to .93, with a mean of .29, a difficult level.

The three-level cognitive complexity model had a strong association with item difficulty and accounted for 41% of variance in item difficulty. The analysis of discrimination power of the three-level cognitive complexity model revealed that gifted students scored significantly higher than above average students on the level two and level three problems, but not on the level one problems. Gifted students also scored significantly higher than average and below average students on the first, second and third level problems. Four of the twenty-seven problems (two first-level problems, one second-level and one third level) did not discriminate significantly among the three levels of mathematical ability. Research findings will be discussed in detail in the next chapter.

## CHAPTER V

### DISCUSSION AND CONCLUSION

The primary purpose of this study was to investigate the psychological validity of the three-mathematical minds model ( $M^3$ ). The secondary purpose was to examine the psychological validity of the three-level cognitive complexity model ( $C^3$ ). The author, along with mathematics experts, developed a test of mathematical ability according to the  $M^3$  and the  $C^3$ . The test was administered to 291 middle school students. Comparative and correlational analyses were conducted to analyze data. This chapter includes the discussion of research findings and implications, recommendations for future research and practice, and study limitations. Research findings related to the three-mathematical minds model will be discussed first. Then, research findings about the three-level cognitive complexity model will be discussed.

#### The Three-Mathematical Minds Model ( $C^3$ )

##### *Research Question 1*

*How theoretically valid is the three-mathematical minds model?* The factor analysis revealed three separate components out of nine subtests; that is, the three factor-solution (expert minds, analytical minds and creative minds) provides the best fit to explain mathematical ability. This finding supports the psychological validity of the three-mathematical minds model at the component level. However, the finding partially supports the  $M^3$  at the subcomponent level, in that factor analysis yielded that some subcomponents of creative and analytical minds clustered in the knowledge component, contradicting the author's theoretical position. The author's initial assumption was that

geometry, algebra and statistics subcomponents cluster in the knowledge component; the three syllogism subcomponents cluster in the analytical component; and insight, selection and induction subcomponents cluster in the creativity component. The discussion of componential and subcomponential research findings follows.

Table 4.3 shows loadings of each subcomponent on the component in which it clustered. According to the analysis, geometry, algebra, statistics, linear syllogism, conditional syllogism and induction cluster in the first component and explain almost 30% of the variance. This component can be labeled, with confidence, the knowledge-reasoning component or expert mind. An inspection of the subtests and problems in each subtest will support my label. In chapter three, I stated that the third level problems in the knowledge component require some reasoning because these problems entail conceptual understanding of subject matter. For example, the level three problem in the geometry subtest requires understanding perimeter and area relationships for the solution. Conceptual understanding means analyzing facts and relating them. However, the fact that these subcomponents cluster with the linear syllogism, conditional syllogism and induction subcomponents is unexpected to this author. Obviously, some overlap exists among the subcomponents.

What was unexpected was the clustering of induction, linear syllogism and conditional syllogism in the knowledge component. They have low to medium, but statistically significant, correlations with the other subcomponents of the knowledge component. An inspection of the problems in the induction subcomponent, which was supposed to cluster in the creativity component but clustered in the knowledge

component, indicates that the first level problem (categorizing numbers), and the second level problem (finding out the sum of the internal angles of a shape of 14 sides) can be solved by using mathematical knowledge, as well as by inductive processes. However, the third level problem (finding out the number of turns of a wheel) does not require mathematical knowledge. Needless to say, the first and second level problems contribute to the overlap between the knowledge component and the induction subcomponent from this author's point of view. Another reason for the overlap can be seen from the finding that the induction subcomponent has a higher correlation with the geometry subcomponent because the second level and the third level problems in the induction subtest also require spatial ability while the first level problem requires numerical ability.

The subcomponent conditional syllogism, which was expected to cluster in the analytical component, clustered in the knowledge component. The conditional syllogism has a similar nature of overlap as the induction subcomponent in that problems in the conditional syllogism were presented in algebraic expressions. Conditional syllogism problems also require the use of coefficients and factors. Therefore, the correlation between algebra and the conditional syllogism subtest is .34, a statistically significant finding (as seen in table 4.4). As a matter of fact, the conditional syllogism correlates with the geometry subcomponent at a higher level, .47. This author believes that the overlap between the conditional syllogism and the knowledge component is occurs because of the similar nature of the problems in the conditional syllogism and the algebra subcomponent. Perhaps some students, who were not good at algebra, did not even attempt to solve conditional problems because of surface similarities between algebra and

the conditional syllogism problems. The strong overlap between the conditional syllogism and the geometry subcomponent found in this study needs to be investigated further in future research related to the M<sup>3</sup>, in particular, and mathematics and mathematical ability, in general.

Unlike conditional syllogism and induction subcomponents, the linear syllogism subcomponent does not have problems similar to algebra problems; however, it, too, clustered in the knowledge component. Linear syllogism problems, in reality, require focused attention, as well as the comparing and contrasting aspects of analytical ability. Therefore, I predicted the linear syllogism subcomponent would correlate highly with the categorical syllogism, and cluster in the analytical component like the categorical syllogism, but it did not. What made the linear syllogism subcomponent cluster in the knowledge component needs further investigation.

Unlike the other subcomponents, which have multiple significant correlations with each other, the categorical syllogism subcomponent has very low correlations with the others. The categorical syllogism subcomponent was found to be a separate component in factor analysis explaining 12% of the variance. The categorical syllogism problems, like linear syllogism problems, require focused attention, contrasting and comparing. Therefore, this component was labeled as an analytical component by this author. The categorical subcomponent has a significant correlation, .13, with only the induction subcomponent. Categorical problems were presented in a table, and the test takers were supposed to figure out the number of individuals in intersecting and disjointed groups. Here, an association exists between the induction and the categorical

problems. However, the overlap is quite small, 1.6%. Keep in mind that a significant correlation might exist between two variables with a large sample size. Therefore, this correlation does not tell much about the relationship between the inductive processes and the categorical syllogism processes. Indeed, the finding might be just an artifact of the sample size. Furthermore, although the categorical syllogism subcomponent itself is a separate component in the factor analysis, the difficulty level of the categorical problems might have contributed to the distinction of this component. Categorical problems have a mean difficulty level of .25, a very difficult level (see table 4.10). Therefore, how much the difficulty level of categorical problems contributes to variance in the categorical subcomponent compared to the other components needs to be investigated further by this author to study mathematical ability both at the factorial level and at the item level.

Like the categorical syllogism subcomponent, the selection and insight subcomponents together appear as a separate component in factor analysis (as seen in table 4.3). However, the insight subcomponent contributes to both the first and second component equally. Although its loading is slightly higher on the second component (.44) than that on the first component (.43), it was not assigned in the creativity component. Therefore, the second component included only the selection subcomponent. The insight subcomponent has a medium level correlation with the geometry subcomponent because two of the insight problems require spatial ability. Therefore, the nature of the problems in the insight subcomponent might have contributed to the correlation between the insight subcomponent and the knowledge component. An inspection of the correlational matrix

presented in table 4.4 indicates that the insight subcomponent has weak associations with the other subcomponents.

The second component was labeled as the creativity component or creative mind by this author because, as stated in the literature review, selection problems requires an unusual mode of thinking, selective thinking; therefore, selection problems are different from problems in the other subtests. Recall that, as pointed out in the literature review (Davidson & Sternberg, 1984), many creative ideas come about through sorting out related and unrelated information and combining them selectively or through the use of analogies as in the third-level selective comparison problem in the selection subcomponent. Overall, the selection subcomponent seems to measure a different aspect of human ability. This difference can be seen more clearly from the point biserial item-subtest correlations in table 4.9. The problems in the selection subtest correlate significantly only with the total score of the selection subtest.

The factorial findings of this study differ in some aspects from those of prior factor analytic studies of mathematical ability. For example, Rogers' study (1918) showed correlations among math subtests, ranging from .02 to .59. Similarly, correlations among the subtests of the M<sup>3</sup> range from almost .00 to .47. Prior researchers such as Spearman (1927), Thurstone (1937) and Verdelin (1958) reported separate numerical and spatial-visual factors underlying mathematical ability. In this study, the author did not find any separate numerical or spatial factors; instead, algebra, geometry and statistics subtests underlying numerical and spatial ability, respectively, clustered in the knowledge-reasoning component. However, Verdelin also found a deductive factor

(analytical ability in this study) that was separate from other mathematical factors. Likewise, the categorical syllogism subcomponent is a separate component from the other components in this study. In addition, this study revealed a mathematical creativity component that has not been reported by other researchers with the exception of Sternberg (2002).

### *Research Question 2*

*What are the psychometric properties of the M<sup>3</sup> test battery?*

*Reliability findings.* The reliability coefficient may be interpreted in terms of the percentage of score variance attributable to different sources. For example, a reliability coefficient of .90 means that 90% of the variance in test scores is accounted for by true variance in ability measured, and 10% is explained by error variance. The mode of reliability analysis in this study was Kuder-Richardson reliability, a measure of interterm consistency. The analysis yielded a reliability coefficient of .72. This coefficient slightly exceeds the desired minimum level of the coefficient, .70 (Anastasi & Urbina, 1997).

Although the coefficients are above the minimum level, they are not very strong, in that over 25% of the variance in scores in the M<sup>3</sup> is attributable to error variance. The M<sup>3</sup> is not a measure of a single trait, but a measure of multilateral aspects of mathematical ability. That is, the items in the M<sup>3</sup> are heterogeneous; therefore, the interitem consistency of the M<sup>3</sup> should not be expected to be very high because the interitem consistency is influenced largely by the heterogeneity of the behavior domain sampled (Anastasi & Urbina, 1997). In other words, the more homogeneous the domain, the higher the interitem consistency. Recall that the M<sup>3</sup> has nine subcomponents and

three components measuring separate aspects of mathematical ability. Overall, the reliability findings provide partial evidence for the reliability of the  $M^3$  as a measure of unified mathematical ability and good evidence for the reliability of the  $M^3$  as a measure of multilateral aspects of mathematical ability.

### *The Construct Validity of the $M^3$*

In this section, I will discuss research findings related to the convergent validity of the  $M^3$  first. I will discuss research findings related to whether the  $M^3$  shows developmental evidence by differentiating among different grade levels and whether it shows internal consistency for item-total score, item-subtest and subtest total score correlations.

*Convergent validity of the  $M^3$ .* Campbell (1960b) pointed out that a psychological test should correlate with other variables to which it should be related theoretically to show construct validity, meaning that the test measures what it intends to measure. Convergent validity, however, is only one way to investigate the construct-related validity of a psychological test. Convergent validity can be investigated through the correlation of the same ability measured by different tests or through the correlation of similar abilities measured by the same or different tests. The correlation between scores on aptitude tests and grades in math courses is one example of convergent validity. In this study, the convergent validity of the  $M^3$  was investigated by correlating students'  $M^3$  scores with teachers' rating of students' mathematical ability, students' rating of their own mathematical ability and students' rating of their liking of mathematics. Note that the first two ratings are measures of students' mathematical ability, also measured by the

M<sup>3</sup>. The liking of mathematics is not a measure of mathematical ability, but it is associated with a student's mathematical performance.

Table 4.7 summarizes partial correlations among the aforementioned variables. Recall that the overlapping effect of grade is statistically controlled in the correlational analysis. As read from the table, correlations between the M<sup>3</sup> and the ratings range from medium to high-medium, with all correlations being statistically significant at the .01 level. The correlation between teachers' rating and the M<sup>3</sup> composite score is the highest among the others ( $r = .45$ ), followed by students' rating of their own ability ( $r = .36$ ) and their liking of mathematics ( $r = .35$ ). What is interesting in the table is the pattern of correlations between the M<sup>3</sup> components and the ratings. Both teachers' ratings and students' ratings correlate with the knowledge component ( $r = .42$  &  $.40$ ) much higher than their ratings with the creativity ( $r = .35$  &  $.26$ ) and analytical component ( $r = .32$  &  $.19$ ). This author interprets these correlations to mean that teachers and students associate mathematical ability more with amount of mathematical knowledge and achievement in math classes than with creative and analytical ability. Perhaps, some creativity and analytical problems also were unusual to the students, as some students commented about some problems on the test booklet by saying, "weird," "different," "excellent," "never seen" or "impossible solution."

Overall, these findings provide evidence for the convergent validity of the M<sup>3</sup>. The author's future research agenda, however, should focus on associations between scores on the M<sup>3</sup> and scores on another test of mathematical ability for a more clear-cut picture of the convergent validity evidence. Also, further research is needed to determine

the relationship between scores on the M<sup>3</sup> and grades in mathematics classes or performance on an achievement test to provide criterion-related validity evidence for the M<sup>3</sup>.

*Developmental differences among students of various grade levels.* A major criterion used in the validation of a number of intelligence tests is age differentiation (Anastasi & Urbina, 1997). However, the use of age in the validation of aptitude tests, such as the M<sup>3</sup>, is not appropriate because they measure ability that is influenced largely by school learning. Therefore, the major criterion for aptitude tests should be grade differentiation, which is what the author used to check the M<sup>3</sup> against grade to determine whether the scores in the M<sup>3</sup> show a progressive increase with grade during middle school (6<sup>th</sup> through 8<sup>th</sup> grade).

As reported in the foregoing chapter, the findings provide partial validity evidence about whether the M<sup>3</sup> discriminates among different grade levels. The partial evidence means that 8<sup>th</sup> grade students scored significantly higher than 7<sup>th</sup> and 6<sup>th</sup> grade students; however, the 6<sup>th</sup> graders scored slightly higher than the 7<sup>th</sup> graders. Needless to say, the latter finding contradicts the former, even though the difference between the 6<sup>th</sup> graders and the 7<sup>th</sup> graders is not significant. The discussion of this contradiction follows.

An inspection of table 3.2 indicates that a significant difference exists among sample sizes. The size of the 6<sup>th</sup> grade is almost half that of the 7<sup>th</sup> grade. This difference might have contributed to the performance difference between the 6<sup>th</sup> and 7<sup>th</sup> graders, favoring the 6<sup>th</sup> graders. Another reason for the contradiction might come from the statewide achievement difference between 6<sup>th</sup> and 7<sup>th</sup> grade students. My personal

conversation with one of the teachers who administered the M<sup>3</sup> indicates that the 6<sup>th</sup> graders scored higher than the 7<sup>th</sup> graders on a state-approved achievement test (L. Chandler, personal communication, January 13, 2005). No reason has been found by state educators for this unexpected difference. Further, developmentally, some peaks and slumps might exist in students' performance during middle school. For example, Torrance (1968) reported that children demonstrated an early peak in divergent thinking followed by a slump around fourth grade, and a late increase. Likewise, Sak and Maker (2003) found stagnancy in fluency of students in mathematics at fourth grade. That is, children's cognitive development may show curvilinear trajectories with peaks and slumps in one or more facets of their ability. Interestingly, in the selection and insight subtests in this study, a slump exists around 7<sup>th</sup> grade and a peak around 8<sup>th</sup> grade. Briefly, the findings show partial, developmental evidence for the validity of the M<sup>3</sup>. Therefore, the author's future research agenda includes administering the M<sup>3</sup> to a different sample to see if the same or different results are obtained. These findings also suggest that researchers should check a newly-constructed ability or achievement test for grade differentiation.

*The internal consistency for item-total score, item-subtest and subtest-total score correlations.* Internal consistency correlations, whether based on items or subtests, show the homogeneity and heterogeneity of the test items used to measure the ability domain sampled by the test. Therefore, the degree of homogeneity or heterogeneity of a test has some relevance to its construct validity. In other words, the internal consistency correlations might provide evidence related to whether items in a test battery measure the

same construct, similar constructs or completely different constructs. Anastasi and Urbina (1997) stated that correlations among items, subtests and the total score are expected to be significant if the test battery is constructed to measure a single, unified construct. Because most psychological tests are developed to measure a unified construct, the degree of significance in correlations among items and between items and subtests of a test battery developed to measure separate abilities is unclear in the psychological literature. The author's assumption is that low correlations exist among items and between items and subtests that are developed to measure separate constructs even though they are parts of the same test battery. For example, the M<sup>3</sup> is designed to measure three aspects of mathematical ability; therefore, low correlations should be expected among items measuring different aspects of mathematical ability, such as the creative mind and the analytical mind. On the other hand, high correlations should be expected among items measuring the same aspect of mathematical ability, such as the analytical mind.

As seen in the point biserial correlational matrix in table 4.9, twenty-six of the twenty-seven items in the M<sup>3</sup> test battery correlate significantly with the total score. Indeed, most correlations are within the range of medium to high, providing evidence of strong internal consistency of item-total score relationships. In other words, the twenty-six items differentiate among the respondents in the same direction as does the entire test battery. However, item 26 have a low and nonsignificant correlation with the total score showing low discrimination. Therefore, this problem needs to be revised or removed from the test battery. Furthermore, bivariate correlations between the subtests and the M<sup>3</sup>

total score are significant, ranging from medium to high. The correlations between items and the total score and the correlations between the subtests and the total score provide evidence of internal consistency related to the construct validity of the M<sup>3</sup> (see table 4.5 for multiple comparisons).

The correlation pattern in table 4.9 also provides additional psychological evidence of the internal consistency of the M<sup>3</sup>. What is most important in this pattern are the correlations between the items and the subtests in which the items are located and the correlations between the items and the subtests in which the items are not located. As seen in the table, the items, such as 1, 10 and 20 located in the linear syllogism subtest, have high and significant correlations with the subtests in which they are located. In other words, each linear syllogism problem, for example, differentiates among individuals in the same direction, as does the entire subtest linear syllogism. This pattern repeats itself for the other subtests, as well. However, the only items that do not have high correlations with their associated subtests are items 24 and 26, though they have moderate and significant correlations. As pointed out in the foregoing discussion, these items may need further revision.

To this author, more interesting than the high correlations are low positive and low negative correlations between the items and the subtests in which these items are *not* located. In other words, these items and the subtests are not developed to measure the same constructs. The correlation between item 25, an induction problem, and the subtest conditional syllogism, for example, is -.04, meaning that they do not measure the same ability and they differentiate among different abilities. Low negative or low positive and

nonsignificant correlations exist between most items of the M<sup>3</sup> and the subtests that measure different constructs. These findings show the internal consistency of the M<sup>3</sup> at the item level and subtest level. The findings also provide validity evidence for the three-mathematical minds model, in that the items developed to measure one of the mathematical minds, such as the analytical mathematical mind, differentiates that mind from the other two minds and vice versa.

### *Research Question 3*

*Which M<sup>3</sup> items represent good measures of mathematical knowledge, analytical mathematical ability and creative mathematical ability?* The purpose of this research question was to attempt to find further evidence at the item level for the psychological validity of the M<sup>3</sup> and for the assessment of multilateral facets of mathematical ability. As stated before, an assessment model should be validated not only on the factorial level, but also on the item level. As Carroll (1996) maintained, one major problem of assessment practices is to determine the homogeneity or heterogeneity of items. The particular interest of this author was to find out which M<sup>3</sup> items were good measures of multilateral aspects of mathematical ability. This type of analysis provides evidence for item homogeneity and heterogeneity of discrimination powers of items to differentiate among individuals who may have high ability in different ability areas within the same domain. As pointed out in the first chapter, some items might show functional deviation, meaning that they measure abilities they theoretically are not constructed to measure. On the other hand, some items might be functionally fit; they only measure abilities they are constructed to measure.

As noted previously, the mode of analysis was a comparison of the performance of the upper 25% of the composite and the components with that of the lower 25%. Each item has four discrimination indices as a result of the use of four base scores: the composite, analytical, knowledge and creativity scores. As reported in the foregoing chapter, GLM repeated measures yield very significant difference among the four groups of discrimination indices. The effect size, computed by using Cohen's  $d$ , is .82, a very large effect. Therefore, discrimination indices for each item deserve further discussion.

As seen in table 4.10, each item has discrimination indices, most of which differ significantly from each other. Item one, for example, has a high discrimination for analytical minds, but has a low discrimination for creative minds. That is, item one is a good measure of analytical ability, but only the use of problems like item one in ability tests will miss many creative minds. Similarly, item 6 is a good measure of creative ability but it overlooks individuals who have high mathematical knowledge. Findings that are more radical are the discrimination indices of items 15 and 16. Although they differentiate analytical minds from nonanalytical minds, they discriminate against knowledge-expert minds. That is to say, knowledge-expert minds will be overlooked when only analytical types of problems are used in ability tests. Based on the discrimination indices presented in table 4.10, the author's conclusions about each item are as follow:

Problems 6, 7, 8, 18, 22 and 25 are good measures of creative ability. Problems 1, 2, 10, 14, 15, 16, 20, 23, 24, and 26 are not good measures of creative ability.

Problems 1, 3, 4, 5, 8, 9, 10, 12, 14, 15, 17, 20 and 22 are good measures of analytical ability. Problems 2, 6, 11, 18, 19, 21, 23, 24, 25 and 26 are not good measures of analytical ability.

Problems 1, 3, 4, 5, 7, 8, 12, 13, 17, 22 and 27 are good measures of mathematical knowledge. Problems 2, 6, 11, 14, 15, 16, 18, 19, 20, 21, 25 and 26 are not good measures of mathematical ability. In fact, problems 15, 16 and 19 discriminate against knowledge-expert minds.

Problem 2 is too easy and problem 26 is too difficult; therefore, they have low discrimination levels. They need further revision.

In addition, some items show functional fitness whereas others do not. Items that show functional fitness, meaning that they measure highly or moderately one type of mathematical mind or only what they are supposed to measure, are creativity problems 6, 11, 18, 19 and 25; analytical problems 9, 10, 14, 15, 16 and 20; and knowledge problems 13 and 23. The rest of the problems show functional deviations, meaning that they also differentiate more than one type of mathematical mind. Interestingly, most of these problems are knowledge problems.

Overall, some problems in the M<sup>3</sup> are homogeneous, differentiating only one type of ability whereas some are heterogeneous, differentiating more than one type of mathematical ability. Some problems have low discriminations not because they do not discriminate between high ability and low ability, but because they are just too difficult; even some of the individuals with the highest abilities could not solve them. Problems 11, 21 and 23 are good illustrations of these types of difficult and low discriminating

problems. As seen in table 4.15, only 35% to 47% of the top 5% of the participants passed items 11, 21 and 23. On the other hand, 1% to 10% of the below-average students passed these items. Therefore, discrimination indices of problems in an ability test should be computed based not only on a comparison of the upper 25% and the lower 25% but also on a comparison of the top 5% and above average individuals if the major purpose of the test is to differentiate between gifted and nongifted individuals. To this author's best knowledge, no prior research exists about this type of item discrimination computation. Finally, no prior researchers computed item discrimination based on different ability groups within the same domain (Linn, 1993). Based on the findings in this study, the author suggests that test items should be validated not only based on item-total score or subtest-total score correlations for their internal consistency, but also on item discrimination indices estimated based on factor scores. This type of item validation is essential particularly for ability tests that measure various aspects of human ability.

### The Three-Level Cognitive Complexity Model (C<sup>3</sup>)

#### *Research Question 4*

*How psychologically valid is the three-level cognitive complexity model?* The author investigated the psychological validity of the three level cognitive complexity model (C<sup>3</sup>) through correlational, regression and multivariate analyses and nonparametric analysis. The author's assumption was that item cognitive complexity, developed based upon the C<sup>3</sup>, was the major source of item difficulty. As pointed out in the first chapter, a good psychological test consists of problems that are in an ascending level of difficulty (Lohman, 2000; Sternberg, 2002); however, the difficulty level of problems should have

psychological sources. The psychological background of the C<sup>3</sup> was discussed extensively in chapter three. Sources of item cognitive complexity also were presented in chapter three. Here, I will discuss research findings related to the psychological validity of the C<sup>3</sup>.

As read from correlations in table 4.11, item cognitive complexity is associated significantly with item difficulty level ( $r = .64$ ). The association between the C<sup>3</sup> and item difficulty was investigated further by regression analysis to find out how much variance in item difficulty was explained by the C<sup>3</sup> or if the C<sup>3</sup> was the major source of item difficulty. As presented in table 4.12, the C<sup>3</sup> is the major source of item difficulty, explaining 41% of the variance. This finding provides support for the psychological validity of the C<sup>3</sup>, but the finding does not suffice as support for the effectiveness of the three levels in differentiating three ability groups: novices, developing experts, and experts. Research question five and associated hypotheses aimed at finding out if the three levels differentiate among the groups. The discussion of research question five follows.

#### *Research Question 5*

*How do the three ability groups, gifted (above 95%), above average (85-94%) and average and below-average (below 85%), as identified by the composite score, differ in their performance on items at different levels of cognitive complexity?* Three null hypotheses were tested to determine whether each level of the C<sup>3</sup> differentiated among novices, developing experts and experts, or average-below-average, above average and gifted students, respectively. The first null hypothesis is related to the discrimination

power of the third level problems between developing experts (average and below-average students) and experts (the top 5% of students): *no significant difference exists between the performance of gifted students and that of above average students on items at the third level of cognitive complexity only*. In other words, only third level problems are supposed to differentiate between developing experts and experts according to the C<sup>3</sup>; whereas, the first and second level problems do not differentiate. As post-hoc comparisons show in table 4.13, both level-two and level-three problems differentiate significantly between expert students and developing expert students; however, level-one problems do not differentiate between the two. These findings imply that level-two problems are very difficult for developing expert students, so most of these students were unable to solve these problems correctly. Therefore, level-two problems should be revised so that no significant performance difference exists between developing expert students and expert students on level-two problems. The findings also indicate that level-one and level-three problems function in the direction they were developed according to the three levels of the cognitive complexity model.

The second null hypothesis is related to the discrimination power of level-two and level-three problems between developing experts and novices: *No significant difference exists between the performance of above average students and that of average and below average students on items at the second and third level of the cognitive complexity*. In other words, only level-two and level-three problems are supposed to differentiate between developing expert students and novice students according to the C<sup>3</sup>. The findings presented in table 4.13 indicate that level-two and level-three problems

significantly differentiate between the two groups; however, developing expert students also perform significantly higher than novice students. As noted earlier, level-one problems do not differentiate between expert students and developing expert students. Therefore, level-one problems need further revision so that no significant performance difference exists between the two groups on level-one problems.

The third null hypothesis is related to the discrimination power of the level-one problems among the three groups: *no significant difference exists among the performance of the three ability groups on items at the first level of the cognitive complexity.* In other words, the first-level problems are not constructed to differentiate among the three groups according to the C<sup>3</sup>. The findings presented in table 4.13 illustrate that the first-level problems do not differentiate between expert students and developing expert students, but do differentiate between the novice students and expert students and between novice students and developing expert students. These findings mean that the first-level problems are difficult enough for novice students; thus, experts and developing experts outperform novices on the first-level problems as well as on the second and third level problems. Therefore, as recommended in the foregoing discussion, level-one problems need further revision so that they do not differentiate significantly among the three ability groups.

Overall, the three-level cognitive complexity model deserves further research, particularly as to whether each level differentiates among the three ability groups in the direction assumed. However, as discussed in research question four, the C<sup>3</sup> accounts for a significant degree of variance in item difficulty (41%), meaning that item difficulty

comes from psychological sources, such as demand for mathematical knowledge, analytical mathematical ability or creative mathematical ability, developed according to the C<sup>3</sup>. Although 41% of the variance in item difficulty is explained by the C<sup>3</sup>, what contributes to the rest of the variance is unknown. Probably, other factors, such as the difficulty of language, the length of verbal statements, the clarity of graphs or some external factors, contribute to item difficulty as well. Therefore, in future research related to the C<sup>3</sup>, the author may investigate the contribution of these factors to item difficulty as well as that of the C<sup>3</sup>. The author also suggests that every newly-constructed test should be validated through the analysis of sources of item difficulty. This type of analysis is missing in many ability tests. This type of analysis should be carried out and reported in test manuals so that test users associate the difficulty of a test with the characteristics of the students to be assessed.

### Study Limitations

A number of limitations exist in this study, most of which pertain to the sample. First, the number of the 6<sup>th</sup> grade students is much less than that of the 7<sup>th</sup> and 8<sup>th</sup> grade students. This difference might have contributed to the performance variance among the groups. Second, the sample is not an exact representation of the U.S. because it was drawn only from the southwest region of the country. Proportions of gender and ethnicity in this study, however, are close to those of the U.S. in general.

The test was given in the beginning of the spring semester; therefore, the participants had not completed their associated grade level. As a result, they had not mastered mathematical knowledge and had not developed skills taught at the end of their

respective grade level. The time of the school year in which the test was given might have contributed to item difficulty because problems were developed according to 8<sup>th</sup> grade students' level of mathematical ability as rated by mathematics teachers.

Other limitations are related to the methodology. First, the M<sup>3</sup> items were hypothetically assigned into their subtests; therefore, a subtest factor analysis was used instead of an item factor analysis. An item factor analysis is recommended for future research related to the M<sup>3</sup>. Second, although the conditional syllogism and linear syllogism subtests did not cluster in the analytical component, as reported by factor analysis, the author grouped them in the analytical component in the computation of multiple item discriminations based on component scores. Likewise, the induction and insight subtests were grouped in the creativity component even though the induction subtest clustered in the knowledge component and the insight subtest loaded equally on the analytical and creativity components according to factor analysis. The third limitation is related to ability grouping in research question five. The participants were categorized into three groups according to their M<sup>3</sup> total scores. Because each item contributes to the participants' total scores, some of the performance difference among the three ability groups on each level of the C<sup>3</sup> might come from the relative contribution of each item. Therefore, the author suggests that readers consider aforementioned limitations in their own interpretation of the research findings.

## APPENDIX A

Crosstabulation of Participants

Inter Item Biserial Correlations

Distribution of Correct Answers in the M<sup>3</sup>

Table A.1

Crosstabulation of participants by school, grade, gender and race

School	Grade	Gender	Race						Total
			Asian	Black	Hispanic	Indian	Other	White	
A	6	Female	0	0	0	0	0	1	1
		Male	0	0	0	0	0	5	5
		Total	0	0	0	0	0	6	6
	7	Female	0	1	1	0	0	1	3
		Male	0	0	2	1	0	12	15
		Total	0	1	3	1	0	13	18
	8	Female	0	0	0	0	0	9	9
		Male	0	0	3	1	0	4	8
		Total	0	0	3	1	0	13	17
B	6	Female	0	1	3	0	0	2	6
		Male	0	0	2	0	0	5	7
		Total	0	1	5	0	0	7	13
	7	Female	0	0	0	0	0	1	1
		Male	2	2	4	0	1	1	10
		Total	2	2	4	0	1	2	11

Table A.1 – continued

School	Grade	Gender	Asian	Black	Hispanic	Indian	Other	White	Total
C	8	Female	0	0	1	0	0	1	2
		Male	0	0	1	0	0	0	1
		Total	0	0	2	0	0	1	3
	7	Female	0	0	3	0	2	9	14
		Male	0	0	2	2	3	10	17
		Total	0	0	5	2	5	19	31
	8	Female	0	0	2	0	0	17	19
		Male	0	0	3	0	0	18	21
		Total	0	0	5	0	0	35	40
	6	Female	1	0	1	0	3	18	23
		Male	2	0	0	0	1	18	21
		Total	3	0	1	0	4	36	44
D	7	Female	4	0	5	0	0	20	29
		Male	0	1	2	0	3	22	28
		Total	4	1	7	0	3	42	57
	8	Female	1	0	5	0	1	19	26
		Male	0	2	2	0	4	17	25
		Total	1	2	7	0	5	36	51

[illegible]

Table A.2 – continued

Item	14	15	16	17	18	19	20	21	22	23	24	25	26	27
1	.11*	-.00	.05	.29**	.03	.05	.01	.05	.05	.01	.00	-.01	-.01	.24**
2	.06	-.06	-.00	.14*	-.06	.04	-.03	.01	.02	.02	.01	.08	-.06	.10
3	-.08	-.09	.08	.25**	.02	.01	.06	.17**	.12*	.05	.01	-.02	.07	.12*
4	.04	.16**	-.05	.24**	.05	.01	.15*	.27**	.17**	.04	.00	.09	.08	.18**
5	.02	-.04	.04	.23**	-.02	.02	.06	.18**	.23**	.14*	-.13*	.06	-.02	.17**
6	.04	.05	-.01	.04	-.05	.09	-.07	-.05	.09	-.02	-.04	-.01	-.09	.09
7	-.01	-.05	-.04	.17**	.08	.01	-.01	.22**	.15*	.02	-.11	-.01	.06	.15**
8	.05	.07	.07	.29**	-.03	.01	.10	.08	.24**	-.02	.01	-.03	-.05	.20**
9	-.07	.01	.06	.09	.07	.01	.02	.13*	.07	.17**	-.06	-.09	-.01	.10
10	-.00	.01	-.02	.19**	-.01	.02	.06	.18**	.06	.06	-.13*	.01	.04	.05
11	.10	.09	.10	.13*	.11	.02	.03	.22**	.21**	.06	-.05	.03	-.04	.08
12	.13*	.02	.00	.27**	.06	.01	.18**	.19**	.34**	.13*	-.06	.04	.06	.22**

13	.00	-.04	-.06	.21**	.04	.04	-.04	.19**	.14*	-.02	-.06	-.01	-.00	.19**
14		.39**	.04	.13*	.07	.01	.04	.12*	.11	.04	.05	.06	.06	.15*
15			.16**	.02	-.02	.02	.01	.03	-.02	.02	.00	-.04	.01	.04
16				.02	.20**	.06	.13*	-.02	.19**	.02	.06	.06	-.06	.04
17					.08	.06	.14*	.22**	.32**	.09	-.00	.05	.05	.24**
18						.08	.04	.16**	.13*	-.01	.07	.04	-.07	.10
19							.10	.07	.15*	-.02	-.03	.08	-.05	.03
20								.16**	.14*	.01	.09	.10	-.09	.09
21									.26**	.17**	-.06	-.06	.17**	.15*
22										.11	-.10	.01	.10	.25**
23											.13*	.02	.02	.13*
24												.02	-.02	.11
25													-.02	.23**
26														.03

Note. \* Correlation is significant at the 0.05 level (2-tailed). \*\* Correlation is significant at the 0.01 level (2-tailed).

Table A.3

Number of correct answers in the total test battery

Number of correct			
answers	Frequency	Percent	Cumulative Percent
1	8	2.7	2.7
2	4	1.4	4.1
3	23	7.9	12.0
4	17	5.8	17.9
5	21	7.2	25.1
6	28	9.6	34.7
7	39	13.4	48.1
8	39	13.4	61.5
9	30	10.3	71.8
10	23	7.9	79.7
11	12	4.1	83.8
12	11	3.8	87.6
13	7	2.4	90.0
14	12	4.1	94.2
15	5	1.7	95.9
16	3	1.0	96.9
17	3	1.0	97.9

Table A.3 – continued

Number of correct			
answers	Frequency	Percent	Cumulative Percent
18	4	1.4	99.3
20	1	.3	99.7
21	1	.3	100.0
Total	291	100.0	

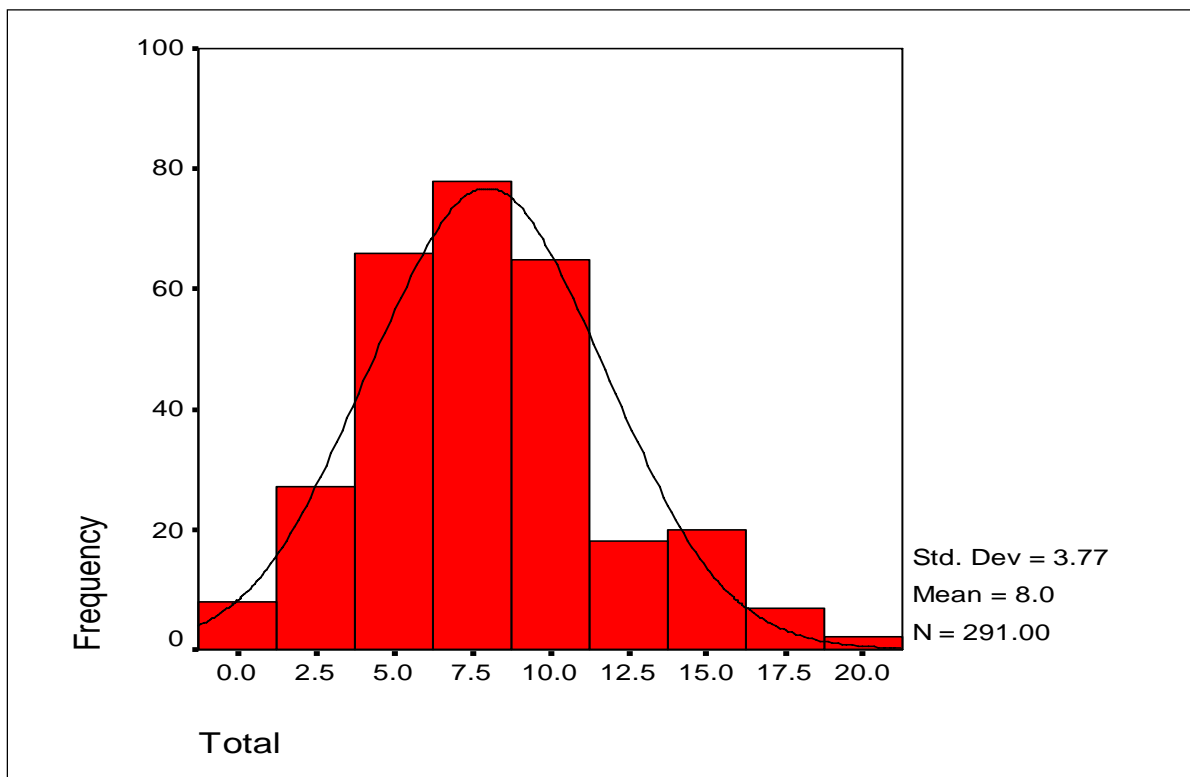


Figure A.1. Distribution of scores in the total test battery

Table A.4

Number of correct answers in the knowledge subtest

Number of Correct			
Answers	Frequency	Percent	Cumulative Percent
0	8	2.7	2.7
1	33	11.3	14.1
2	41	14.1	28.2
3	66	22.7	50.9
4	68	23.4	74.2
5	38	13.1	87.3
6	16	5.5	92.8
7	16	5.5	98.3
8	2	.7	99.0
9	3	1.0	100.0
Total	291	100.0	

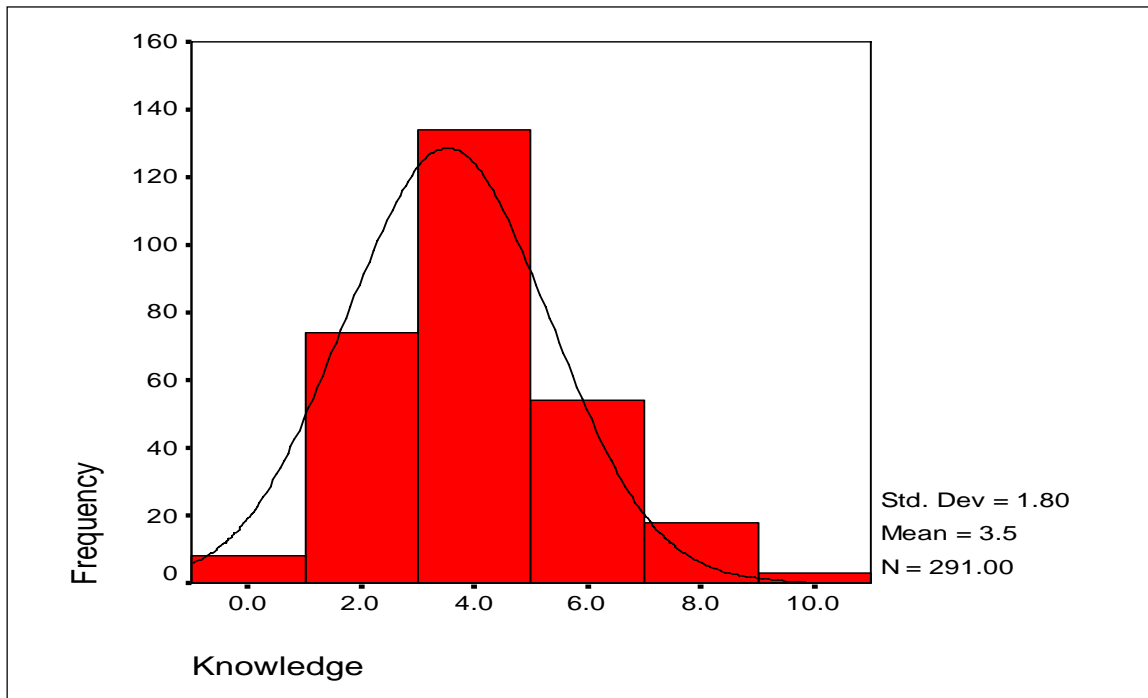


Figure A.2. Distribution of scores in the knowledge subtest.

Table A.5

Number of correct answers in the analytical subtest

Number of correct			
answers	Frequency	Percent	Cumulative Percent
0	25	8.6	8.6
1	62	21.3	29.9
2	69	23.7	53.6
3	63	21.6	75.3
4	34	11.7	86.9
5	21	7.2	94.2
6	10	3.4	97.6
7	5	1.7	99.3
8	2	.7	100.0
Total	291	100.0	

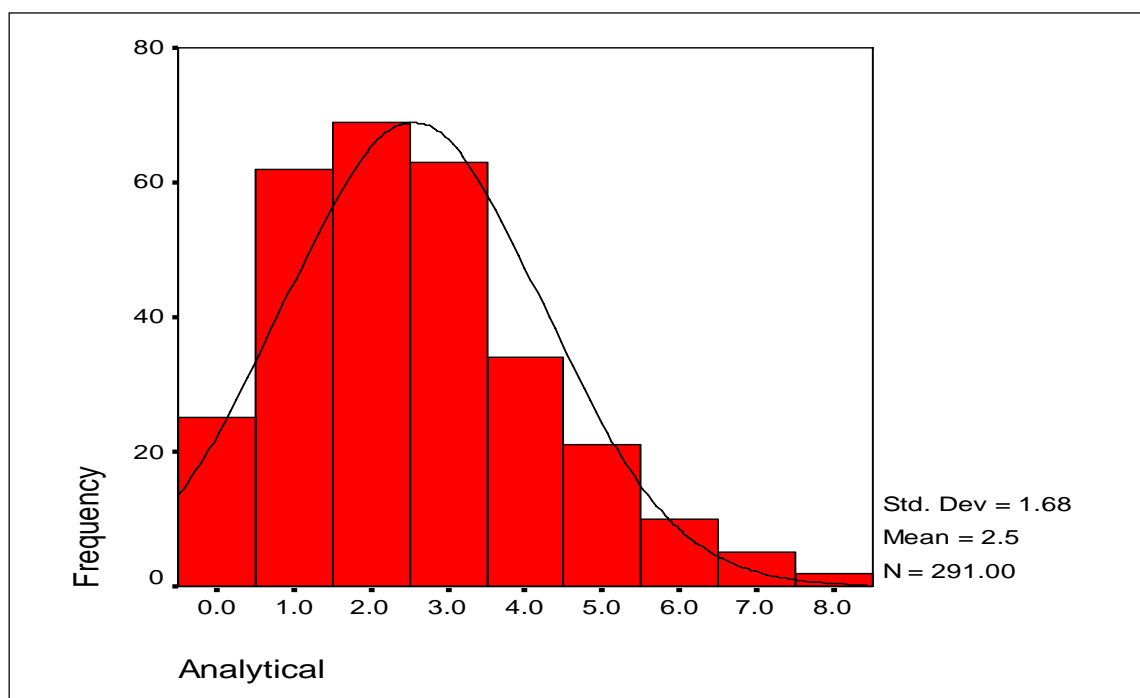


Figure A.3. Distribution of scores in the analytical subtest

Table A.6

Number of correct answers in the creativity subtest

---

N	Frequency	Percent	Cumulative Percent
0	37	12.7	12.7
1	70	24.1	36.8
2	109	37.5	74.2
3	44	15.1	89.3
4	23	7.9	97.3
5	6	2.1	99.3
6	2	.7	100.0
Total	291	100.0	

---

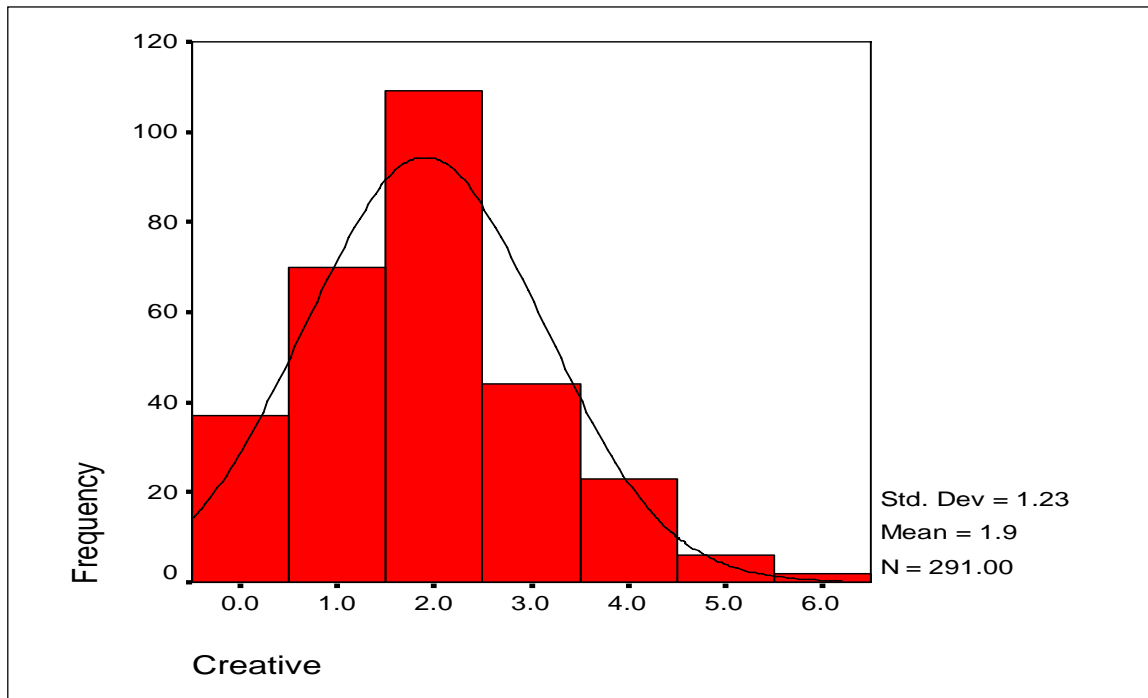


Figure A.4. Distribution of scores in the creativity subtest

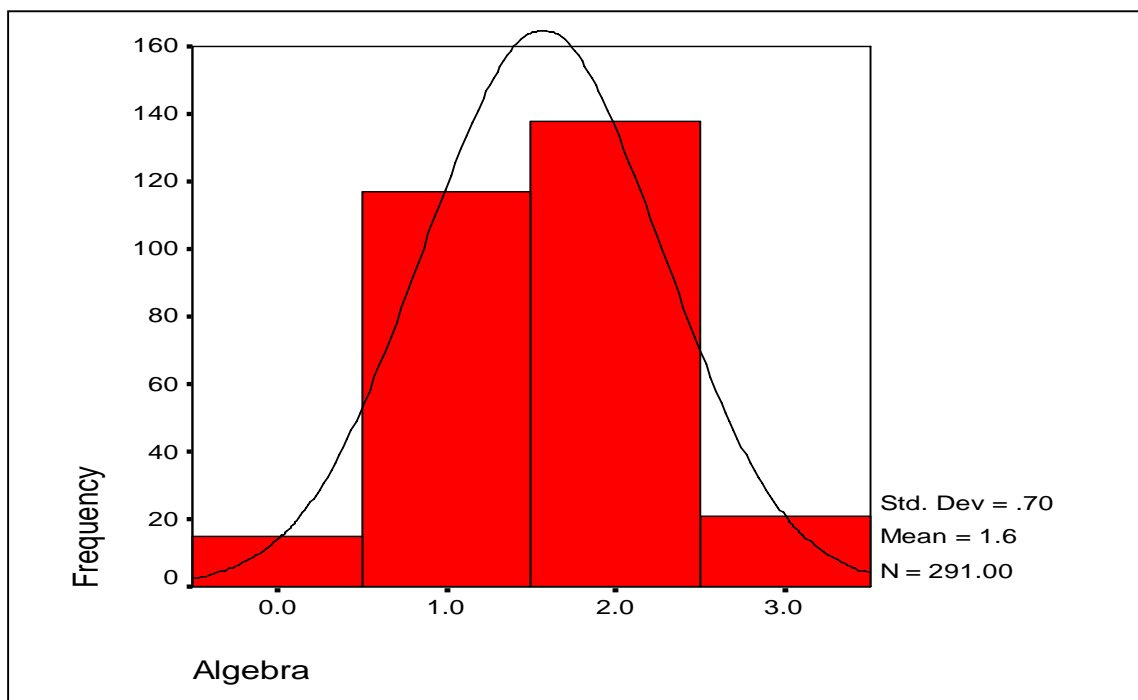


Figure A.5. Distribution of scores in the algebra subtest

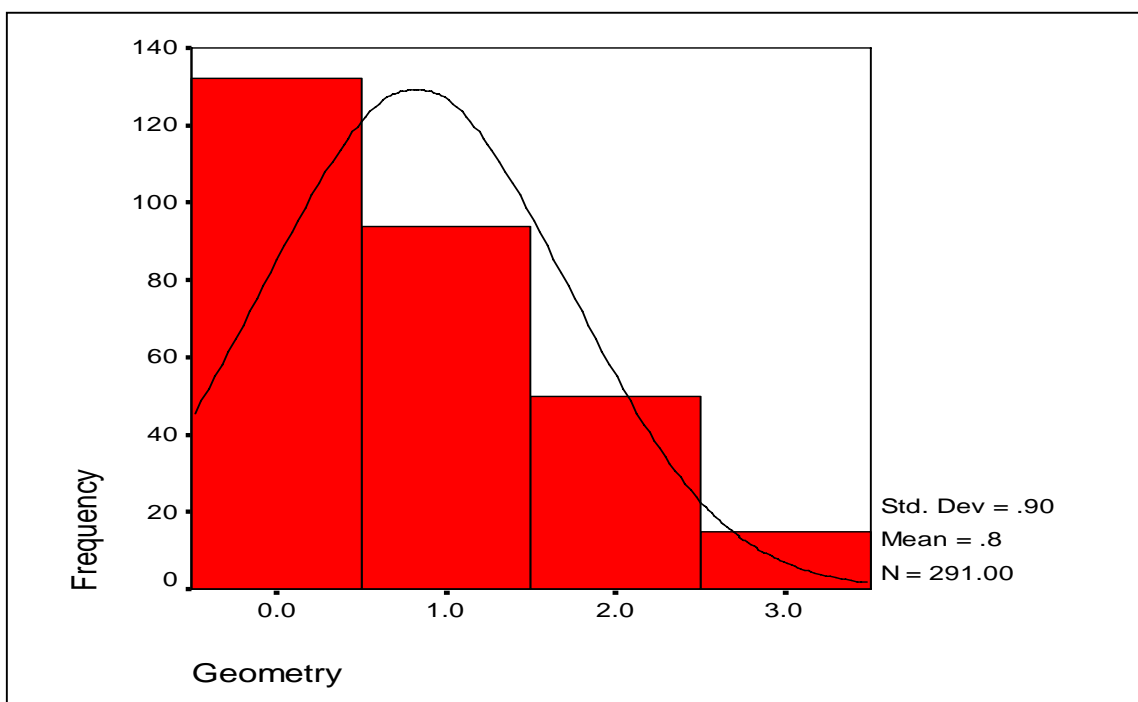


Figure A.6. Distribution of scores in the geometry subtest

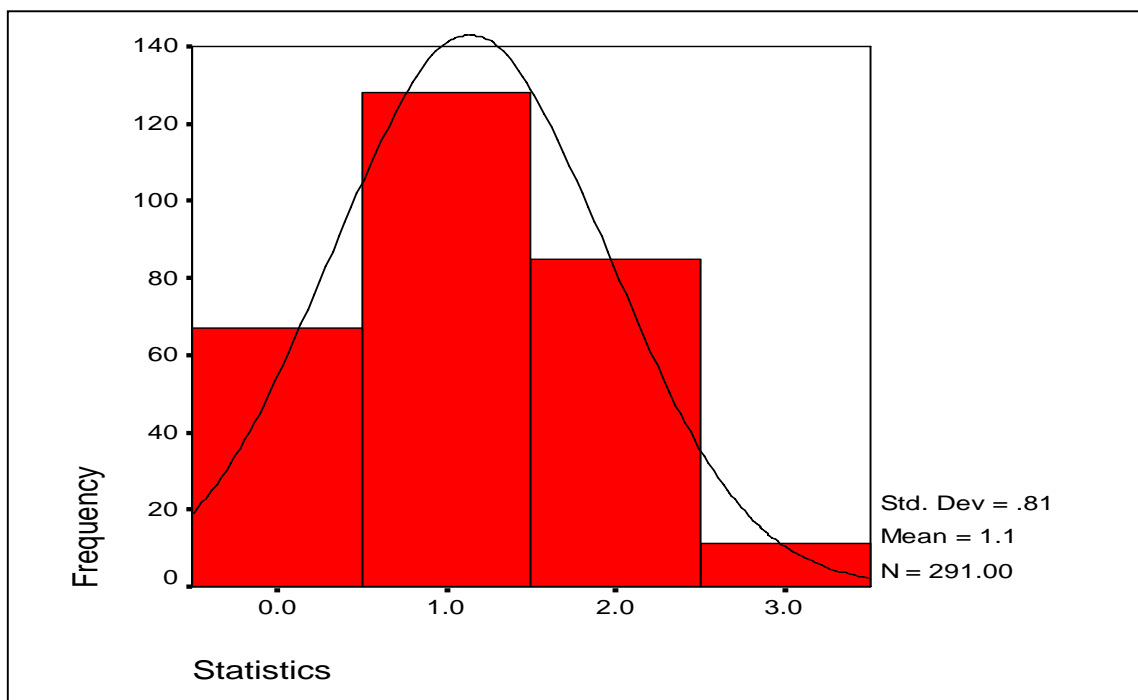


Figure A.7. Distribution of scores in the statistics subtest

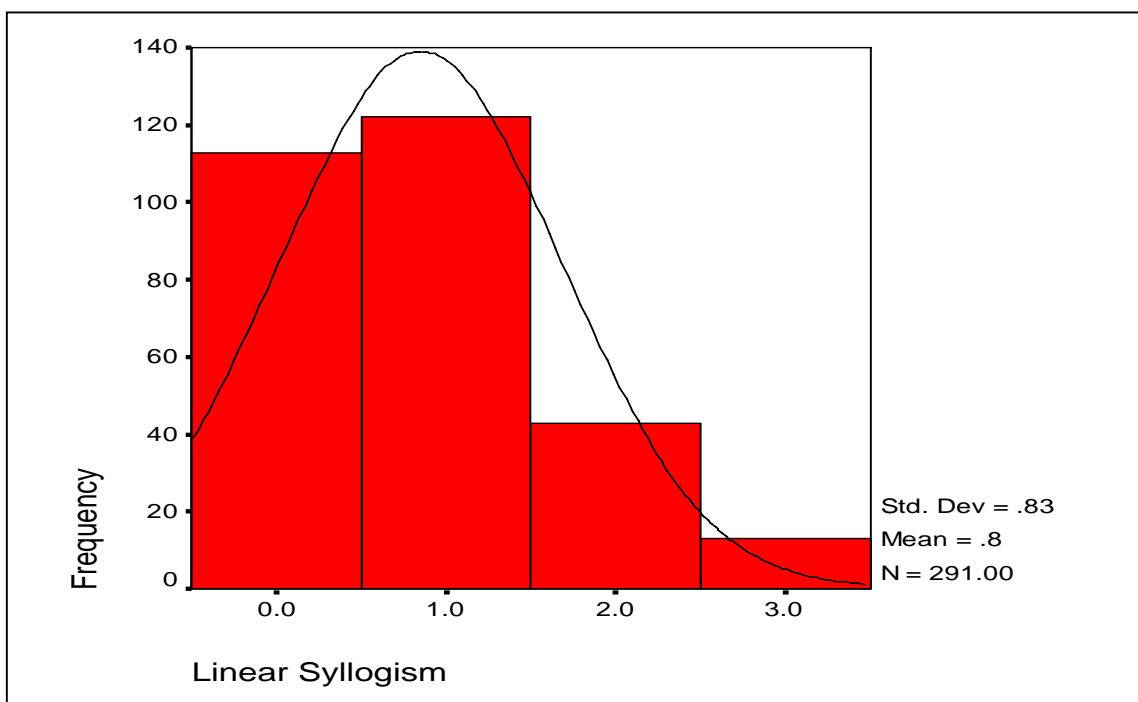


Figure A.8. Distribution of scores in the linear syllogism subtest

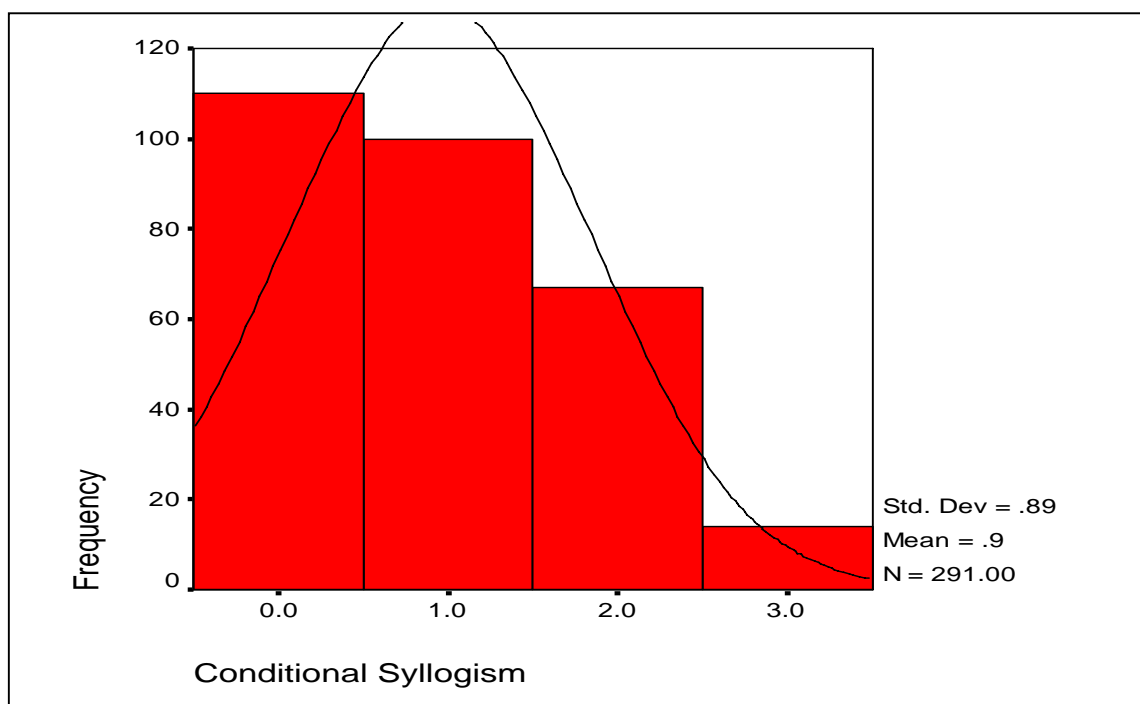


Figure A.9. Distribution of scores in the conditional syllogism subtest

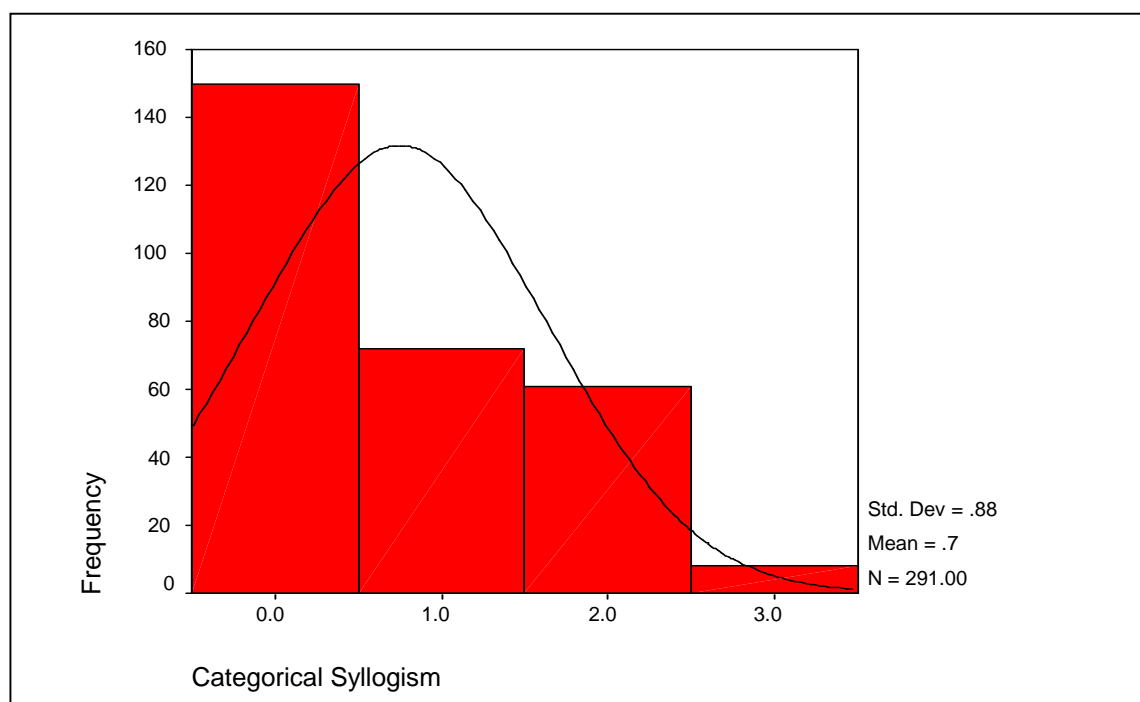


Figure A.10. Distribution of scores in the categorical syllogism subtest

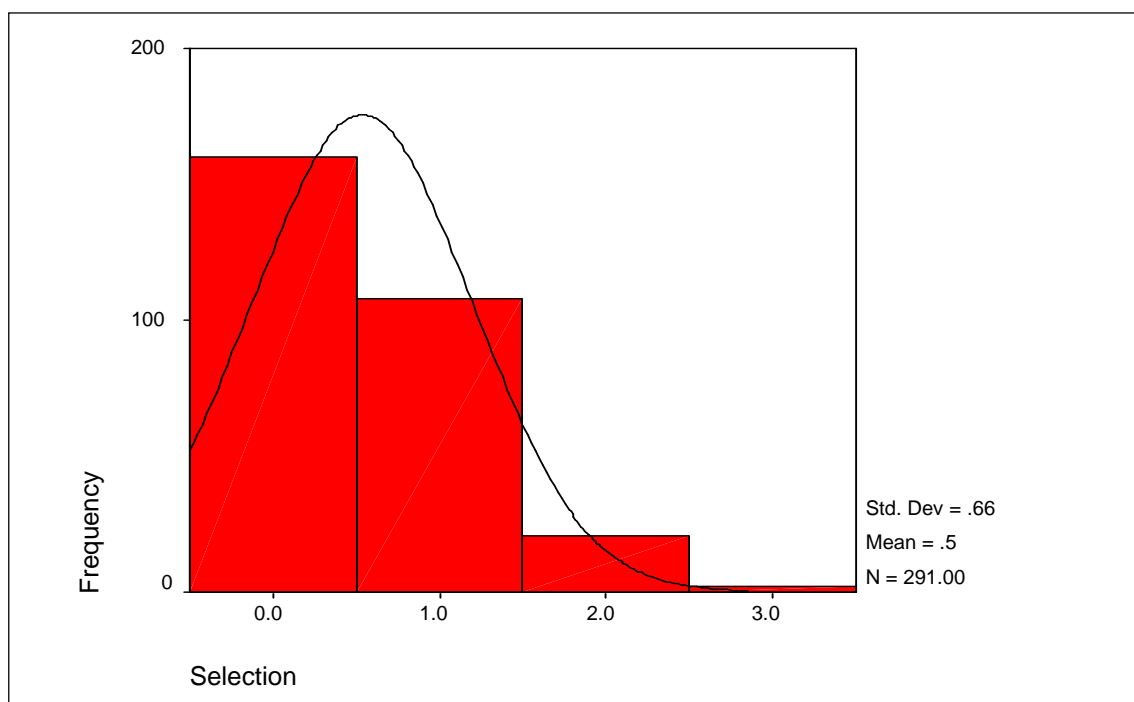


Figure A.11. Distribution of scores in the selection subtest

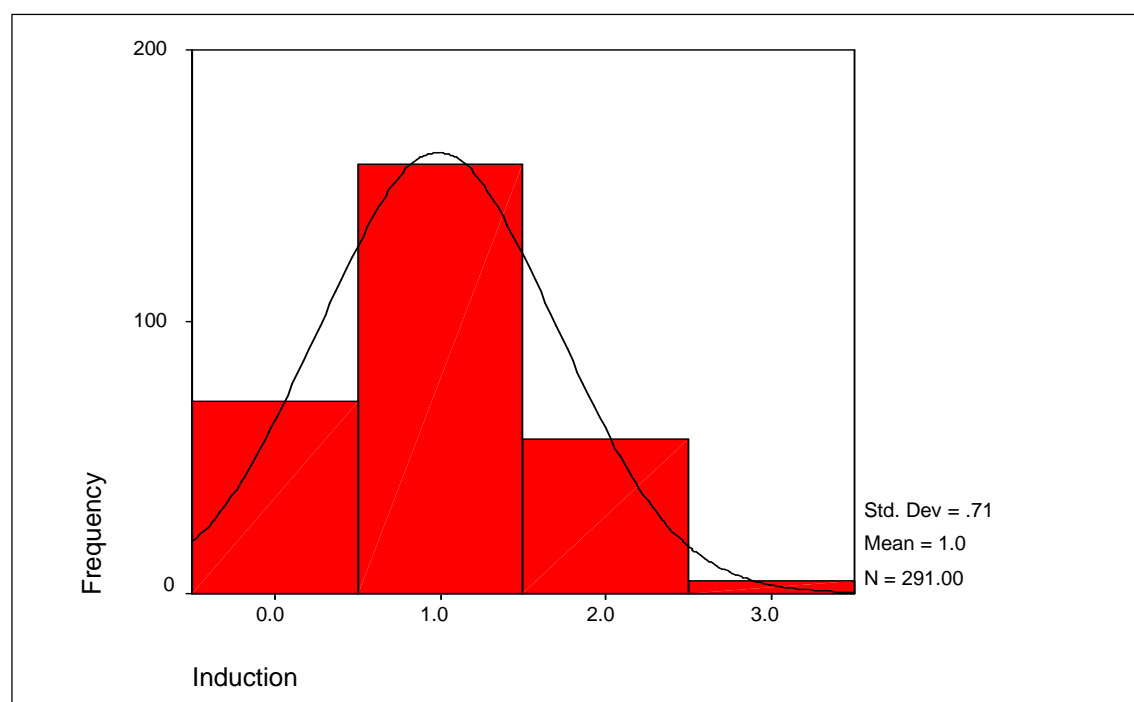


Figure A.12. Distribution of scores in the induction subtest

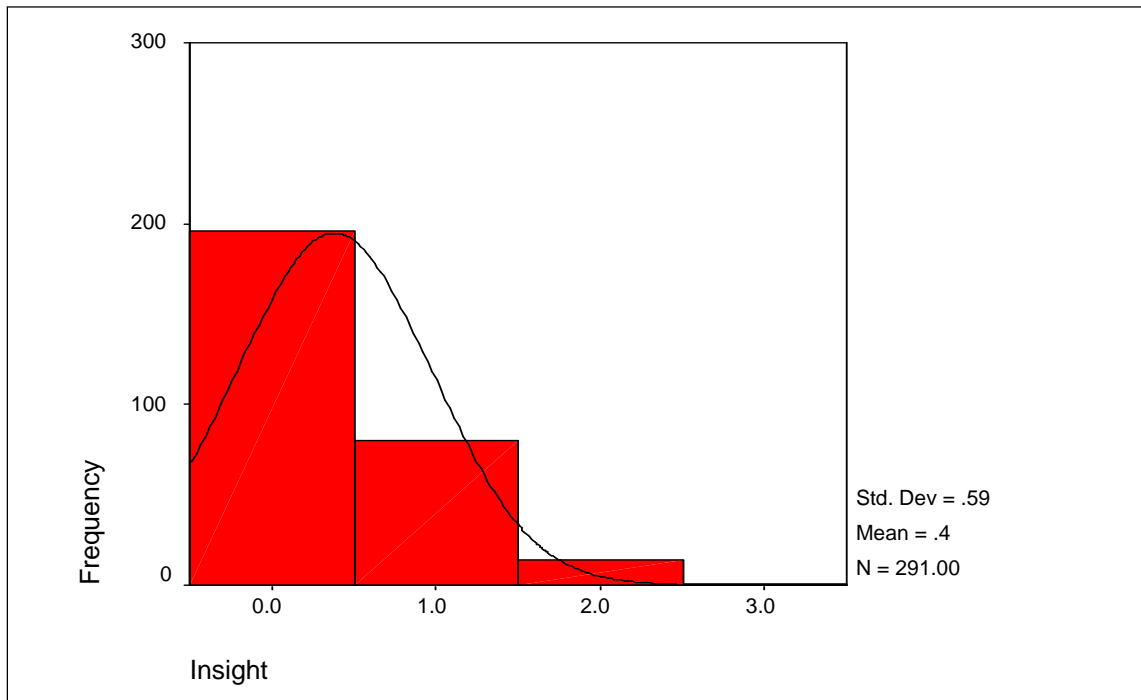


Figure A.13. Distribution of scores in the insight subtest

Table A.7

ANOVA for grade differences on the M<sup>3</sup> composite

Variable	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	204.46	2	102.23	7.50	.001
Within Groups	3924.37	288	13.62		
Total	4128.83	290			

Table A.8

Post-Hoc comparisons of mean differences among all grades on the M<sup>3</sup>

						95% Confidence Interval	
						Lower	Upper
Test	Grade	Grade	Mean Difference	Std. Error	Sig.	Bound	Bound
Tukey	6	7	.26	.57	.89	-1.09	1.62
		8	-1.54*	.58	.02	-2.91	-.17
	7	6	-.26	.57	.89	-1.62	1.09
		8	-1.80*	.48	.00	-2.95	-.65
		6	1.54*	.58	.02	.17	2.91
		8	1.80*	.48	.00	.65	2.95

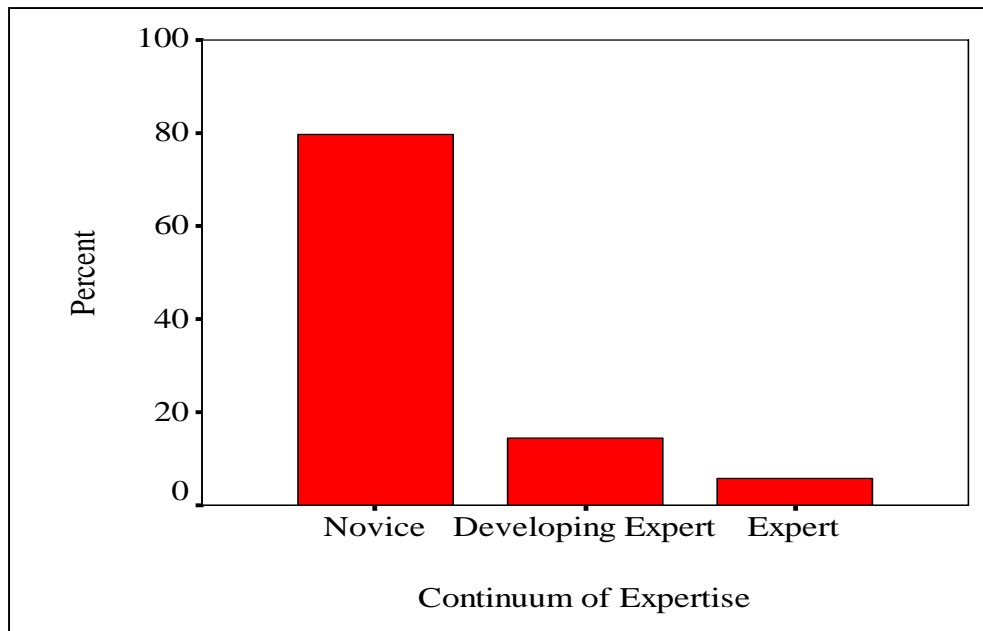


Figure A.14. Percentage of three ability groups on the M<sup>3</sup>

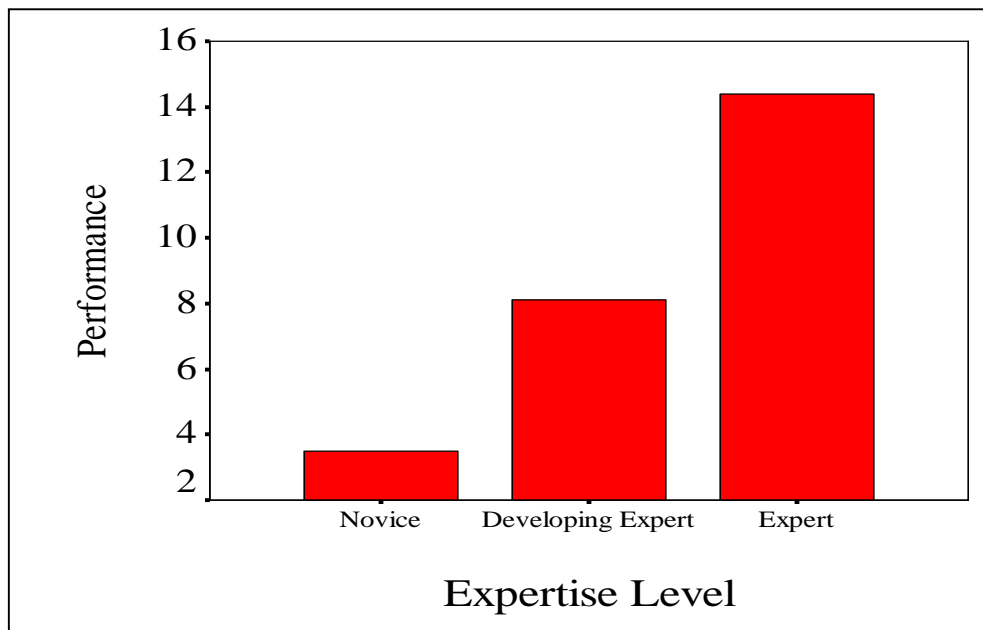


Figure A.15. Performance distribution of three ability groups on the M<sup>3</sup>

## APPENDIX B

The  $M^3$  Problems

Solutions of the Problems

Problem Characteristics

Cover Page of the  $M^3$

Table B.1

M<sup>3</sup> problems and their order in the test booklet

---

Item	
Number	Problems

---

1.	$A > B; C > D; D > E; E > A$ . Which is the <u>second</u> largest?  (a) A  (b) B  (c) C  (d) D  (e) E
2.	$x + 9 = 27$ . What is the value of $x$ ?  a) 36  b) 18  c) 9  d) 3  e) 27

---

---

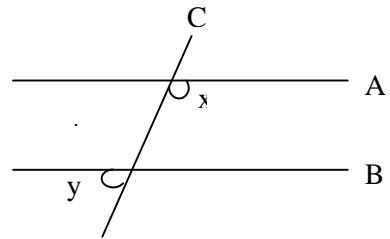
3. If  $x > 0$ , which one is definitely correct?

- a)  $2x > 1$
- b)  $x^2 > 1$
- c)  $x^2 > 2x$
- d)  $x^2 > x$
- e) none of the above

4. In the figure below, A and B lines are parallel, and the measure of angle y is 50.

What is the measure of angle x?

- a) 140
- b) 110
- c) 120
- d) 50
- e) 130



- 
5. Sam and Mike are carrying books for the school library. There are 64 books. How many books are left if Sam carries  $\frac{1}{4}$  of the books, and Mike carries  $\frac{1}{8}$  of the books?
- (a) 38
  - (b) 48
  - (c) 40
  - (d) 56
  - (e) 24
6. Thirty percent of students in a classroom play football; sixty percent play basketball, and ten students play both sports. What is NOT required to find the total number of students in the classroom?
- (a) number of students who play football
  - (b) proportion of students who play both football and basketball
  - (c) number of students who do not play basketball
  - (d) proportion of students who do not play football
  - (e) number of students who play basketball
-

- 
7. Suppose that you need to plant 5 trees in 2 rows, with 3 trees in each row in your backyard. Draw your answer below.

8. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 20, 25, 30

Make two different groups consisting only of the above numbers. Use each number only once. All numbers in each group must have some commonalities, or each group must be based on a rule. Write under each group commonalities you have found among the numbers or the rule on which you have based your grouping. Try to make the most original grouping! An example may be like this: some numbers are one digit and some are two digits, so one digit numbers can be placed in one box and two digits can be put in another box. Now, you should make two different groups, as said above.

Group 1

Commonalities or the rule:

---

---

---

---

Group 2

Commonalities or the rule:

---

---

---

---

---

9. If  $x < -1$  and  $y < x$ , which one is definitely correct?

a)  $y^2 > x^2$

b)  $y^2 < x^2$

c)  $x^2 < 2y$

d)  $3x > y^2$

e)  $2x < 2y$

10.  $a = b + 1$ ;  $c = d + 4$ ;  $a = f + 4$ ;  $b = d + 2$ . Which is the smallest?

(a)  $c$

(b)  $d$

(c)  $b$

(d)  $f$

(e)  $a$

---

---

11. What is the sum of the internal angles of the shape “x” in the table?

a) 1800

b) 2520

c) 1260

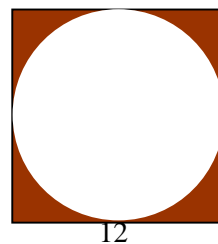
d) 5040

e) 2160

Shapes	Sum of internal angles	Number of sides
Triangle	180	3
Pentagon	540	5
Heptagon	900	7
X	?	14

12. What is the area of the shaded region in the figure below

if the center of the circle also is  
the center of the square whose  
one side is 12 inches?



a)  $144 - 36\pi$

b)  $144 - 12\pi$

c)  $144 - 6\pi$

d)  $48 - 6\pi$

e)  $42 - 12\pi$

---

---

Bottles	Ounces per bottle	Price per bottle	Extra discount At check out	Promotion
A	18	\$1.79	None	None
B	36	\$5.40	%55	None
C	6	\$1.00	None	Buy one get one free

13.

There are five different sizes of bottles of peanuts of the same quality as seen in the above table. The table shows the amount of peanuts each bottle has and its cost. Notice that some bottles have extra discounts or promotions.

Which is the best buy for peanuts?

- a) A
  - b) B
  - c) C
  - d) B and C
  - e) A and C
-

---

Answer questions **14, 15, and 16** according to the table below. The table shows the number or the proportion of students in school clubs. You must read the table from the left to the right, not from the top to the bottom. For example, the first row shows that there are 25 students in the chess club and all of them are in the technology club and none is in the poetry club. The second row shows that half of the technology students are in the chess club and five technology students are in the art club.

<b>Clubs</b>	Chess	Technology	Poetry	Art
Chess	25	All	None	?
Technology	Half	?	?	5
Poetry	None	Half	22	?
Art	?	All	All	?

14. What is the number of students in the art club?
- a) 22
  - b) 5
  - c) 6
  - d) 11
  - e) 17
-

15. How many chess students also are in the art club?

- a) 5
- b) 6
- c) 14
- d) 25
- e) none

16. How many students in the technology club are neither in the chess club nor in the poetry club?

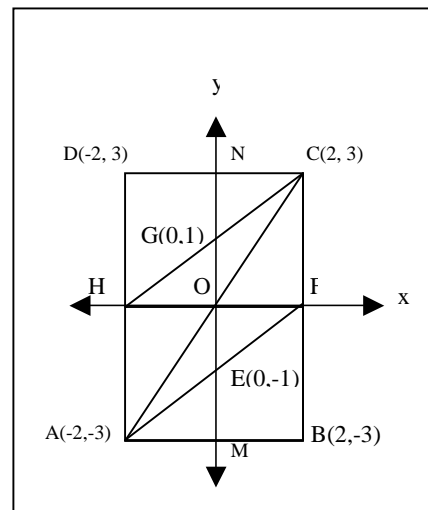
- a) 3
  - b) 11
  - c) 22
  - d) 14
  - e) none
-

17. If  $2x - 3y = 2$ ;  $y = 3z$ ; and  $z = 2$ , what is the value of  $x$ ?

- a) 10
- b) 20
- c) 6
- d) 9
- e) 8

18. Which combination of two triangles in the figure makes a square? (A square has four equal sides). The figure is not drawn on a scale.

- |        |         |
|--------|---------|
| 1. HDC | 8. NOC  |
| 2. HFC | 9. AOH  |
| 3. HOG | 10. AME |
| 4. EOF | 11. AFB |
| 5. CNG | 12. AHF |
| 6. COF | 13. ACB |
| 7. ACD |         |



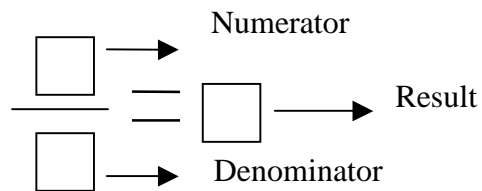
- a) 1 and 2
- b) 3 and 4
- c) 7 and 13
- d) 5 and 10
- e) 6 and 9

- 
19. Which combination of any two triangles in the above figure forms a rectangle similar to the 4 by 8 rectangle below?



- a) COF and AOH
  - b) ABC and DCA
  - c) AMO and HOA
  - d) FOC and HOA
  - e) EOF and HOG
20.  $A = 3B$ ;  $C = 2D$ ;  $F = G/2$ ;  $D = A/2$ ;  $2D = 3G$ . Which is the smallest?
- (a) G
  - (b) B
  - (c) F
  - (d) D
  - (e) C
-

21. Consider that the numerator of a fraction is smaller than its denominator and the result of the fraction is larger than its numerator. Write such a fraction by filling the boxes below.

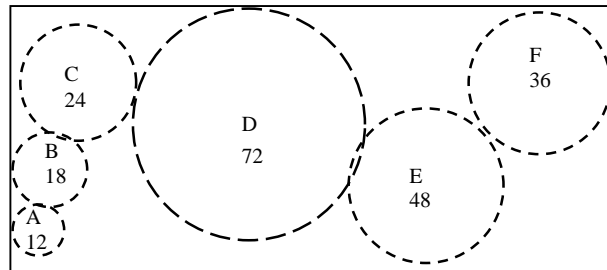


22. If  $x^2 + y = 19$  and  $y \neq 3$ , which one is definitely correct?
- a)  $x \neq 4$
  - b)  $x \neq 3$
  - c)  $x \neq 2$
  - d)  $x \neq 1$
  - e)  $x \neq 0$

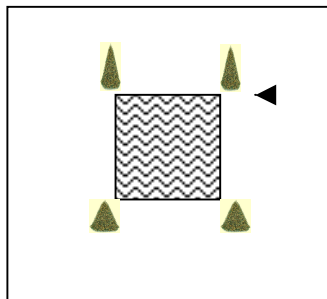
23. Mary worked 5 hours and Josh worked 6 hours in the first day, and together they made \$71. The second day Mary worked 3 hours and Josh worked 2 hours, and together they made \$33. How much was Josh paid per hour?
- a) 7
  - b) 6
  - c) 5
  - d) 8
  - e) 6.5
24. Sally spent 62.5% her money to buy a car and deposited  $\frac{1}{8}$  of her money in her saving account at a rate of 6% interest. One year later, she had \$742 in her account. How much did she pay for the car?
- a) \$742
  - b) \$3710
  - c) \$3500
  - d) \$5600
  - e) 2100
-

25. The figure shows a chain of six gear-wheels. The number on each one shows how many gears that particular gear-wheel has. How many times does F turn when A moves clockwise 18 full turns?

- a) 3
- b) 4
- c) 6
- d) 12
- e) 18



26. Mr. Sanchez has a square farm. A tree stands 5 feet from and 90 degree diagonal to each corner as seen in the figure. How can he double his farm in area, which still is square, without cutting or removing the trees or without owning the land upon which the trees stand? Draw your solution on the figure.



- 
27. Draw three different shapes that have equal perimeters but different areas.

Prove your answer.

---

Table B.2

Solutions of the M<sup>3</sup> problems

Item	
Number	Solutions
1.	Option D
Random values:	
$A = 20$ and $B = 19$ ( $A > B$ ), $E = 21$ ( $E > A$ ),	
$D = 22$ ( $D > E$ ), $C = 23$ ( $C > D$ ); therefore,	
$C > D$ , $D > E$ , $E > A$ , and $A > B$	
2.	Option B
$x = 27 - 9$ ; $x = 18$	
3.	Option E
If $x$ is a number between 0 and 1, then all the other options are wrong.	
4.	Option E
$X$ is an alternative angle to $y$ . The sum of them must be equal to 180.	

- 
5. Option C

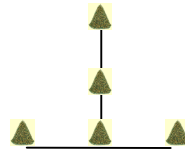
$$64 - (64/4 + 64/8) = 40$$

6. Option D

Proportion of students who do not play football is not required because the number of football players or that of those who do not play football is not given.

7. Many solutions

An intersection between two rows is required, such as a shape of T, L, X, V or a cross:



8. Many solutions

For example, 2, 4, 8, and 10 are divisible by 2 and 3, 9, and 15 are divisible by 3.

9. Option A

When two negative numbers are smaller than -1, the square of the smaller one always is larger than the square of the larger number.

---

---

10. Option D

Random values:

$$a = 10, b = 9, c = 11, d = 7, f = 6$$

11. Option E

$(n - 2) 180$  is the rule to find the sum of internal angles of any shapes. This can be induced from the number of sides and the sum of internal angles of the shapes given in the problem stem. The second solution is adding 180 degrees with each additional side after the triangle.

12. Option A

The square's area is 144 ( $12 \times 12$ )

The circle's area is  $36\pi$  because its radius equals to one half of the square's one side ( $12/2$ )

13. Option B

Bottle B has the lowest price per ounce. This can be found after a 55% discount is taken away from the total price and the remaining is divided by the amount per ounce.

---

---

14. Option B

All art students are in the technology class and only 5 of the technology students are in the art class.

15. Option E

All art students are in the poetry class and none of the poetry students is in the chess class; therefore, there is no intersection between art class and chess class.

16. Option D

The number of students in the intersections and dissections among chess, technology and poetry classes should be taken into account for a correct solution. The following is one way to find the correct solution:

The total number of technology students is 50 because half of the technology class is in the chess class, which has 25 students. Half of the poetry class, which is 11, is in the technology class. No intersection between poetry and chess classes exists; therefore, the sum of 25 and 11 gives the total number of technology students who are either in chess or poetry class. Finally,  $50 - 36 = 14$  is the correct answer.

---

---

17. Option A

$$2x - 3y = 2; y = 3z; \text{ and } z = 2$$

$$2x - 18 = 2$$

$$2x = 20; \text{ therefore, } x = 10$$

18. Option D

Two triangles with equal bases and heights should be found.

A 2 x 2 square is an example.

Others are not square.

19. Option E

A rectangle with sides in  $\frac{1}{2}$  rate should be composed from triangles in the graph; that is, selective encoding and combination should be carried out to construct an analogous side relation in the rate of  $\frac{1}{2}$  between two rectangles.

20. Option C

Random values:

$$A = 15, B = 5, C = 15, D = 7.5, F = 2.5, G = 5$$

---

---

21. Many solutions

Negative numbers or decimals have to be used for a correct solution, such as

$$-6/2 = -3$$

22. Option A

If  $y$  is not equal to 3, then  $x^2$  cannot be 16; therefore,  $x$  cannot be 4.

23. Option B

$$5\text{Mary} + 6\text{Josh} = \$71; \underline{3\text{Mary} + 2\text{Josh} = \$33}$$

$$-3(3\text{Mary} + 2\text{Josh}) = \$33; \underline{5\text{Mary} + 6\text{Josh} = \$71}$$

$$-4\text{Mary} = -\$28; \text{Mary} = \$7 \text{ per hour}$$

$$\text{Josh} = \$6 \text{ per hour (substitute Mary in the first or second equation)}$$

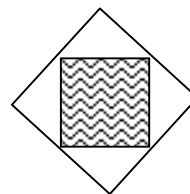
24. Option C

25. Option C

Relation: Wheels increasing by multiples of 6 in gears. Rule: speed rate of  $1/3$

---

- 
26. One correct solution



27. Many solutions

For example, a square, a rectangle and a circle with 2 inches in perimeters will have areas from the largest to the smallest in the order of circle, square and rectangle. That is, the area always becomes different in units no matter how the shape is changed even if the perimeter remains the same. The following is an example:

Square's perimeter ( $2 \times 2$ ): 8 inches

Square's area: 4 square inches

Rectangle's perimeter ( $1 \times 3$ ): 8

Rectangle's area: 3

Circle's perimeter ( $\pi 2.54$ ): 8, respectively

Circle's area: 5.06, respectively

---

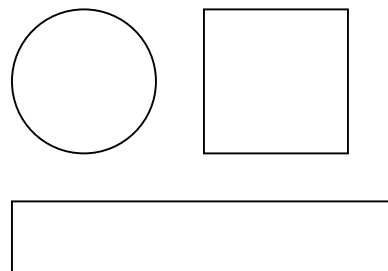


Table B.3

Problem complexity levels and cognitive components and subcomponents measured by each problem

Item			Complexity
number	Mind measured	Subcomponent measured	level
1	Analytical	Linear syllogism	1
2	Knowledge expert	Algebra	1
3	Analytical	Conditional syllogism	1
4	Knowledge expert	Geometry	1
5	Knowledge expert	Statistics	1
6	Creative	Selection	1
7	Creative	Insight	1
8	Creative	Induction-rule production	1
9	Analytical	Conditional syllogism	2
10	Analytical	Linear syllogism	2
11	Creative	Inductive rule discovery	2
12	Knowledge expert	Geometry	2
13	Knowledge expert	Statistics	2
14	Analytical	Categorical syllogism	1
15	Analytical	Categorical syllogism	2
16	Analytical	Categorical syllogism	3

Table B.3 – *continued*

Item	Complexity		
Number	Mind measured	Subcomponent measured	level
17	Knowledge expert	Algebra	2
18	Creative	Selection	2
19	Creative	Selection	3
20	Analytical	Linear syllogism	3
21	Creative	Insight	2
22	Analytical	Conditional syllogism	3
23	Knowledge expert	Algebra	3
24	Knowledge expert	Statistics	3
25	Creative	Induction-rule discovery	3
26	Creative	Insight	3
27	Knowledge expert	Geometry	3

For Teacher Use Only:		
<p>The mathematics test you are about to take has 2 questions about your perceptions of mathematics and 27 mathematics problems you need to answer. You have <b>45 minutes</b> to complete the test.</p>		
<b>First Name:</b>	<b>Grade:</b>	<b>Gender:</b>
<b>Last Name:</b>	<b>Date of Birth:</b>	<b>Race:</b> African American____ American Indian____ Asian____ Hispanic____ White____ Other_____

Figure B.1. Cover page of the test booklet

## APPENDIX C

## Student Questionnaire and Teacher Instruction

## Student Questionnaire

C. How much do I like mathematics?

- a. very much
- b. much
- c. some
- d. a little
- e. not at all

D. How am I in mathematics?

- a. excellent
- b. good
- c. ok
- d. weak
- e. very weak

### Teacher Instructions and Questionnaire

#### Before the Test:

Please, read the following instructions before students start the test:

You will solve some mathematics problems today. The time for the test is 45 minutes. Results of this test will not affect your grade in any ways. The researcher is interested in how you solve different mathematics problems. Please, try your best!

#### After the Test:

After all students finish the test, rate each student's mathematical ability according to the following five-point rating scale. Write the rating in the box designated for teacher use on the cover page of each test booklet:

- 5) Highly talented
- 4) Has high ability but not necessarily talented
- 3) Average
- 2) Weak
- 1) Very weak

Thanks for your participation in this research.

## APPENDIX D

## Epilogue

## Epilogue

Every snowball begins its journey as a snow flake so does every spring begin with only a flower. In the beginning, Dr. Maker and I thought I should develop a new assessment of mathematical ability to identify mathematically talented students although we knew a lot of tests of mathematical ability in existed. Coming from a psychological tradition, I started reading current studies on conceptions of mathematical ability and its assessment from a psychological vantage point. Frankly speaking, I vehemently followed ideas of many contemporary psychologists; however, they provided little illumination about how a mathematician thinks and solves problems of the mathematical kind. I learned little about the content and structure of mathematics from a psychological point of view. However, I learned a lot about how to develop test items psychologically in order to assess mathematical ability objectively.

Then, I began extensive reading about mathematics and mathematicians. I discovered Poincare, Polya, Riemann, Russell and many others. Later, I realized that I had gone into the history of mathematics. I could not stop myself. I went to the library to check out one book, but I left the library with many books. I would not have finished this dissertation in time should I have not wittingly stopped delving into the writings of those who were both mathematicians and philosophers of mathematics.

I was not a mathematician, nor have I become a mathematician during the writing of this dissertation. However, I learned a lot about the philosophy and psychology of mathematical ability. I studied the content of mathematics like a high school student and got very interested in theorems and the discovery of theorems in mathematics rather than

just the content of mathematics. I inevitably admitted what Hardy once said, "...to many readers who never have been and never will be mathematicians,... there is more in mathematics than they thought" (Hardy, 1940, p. 77). "Hah!" said I, after facing the phenomenon.

## REFERENCES

- Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Upper Saddle River, NJ: Prentice Hall.
- Ansburg, P. I., & Hill, K. (2003). Creative and analytic thinkers differ in their use of attentional resources. *Personality and Individual Differences*, 34, 1141-1152.
- Biggs, J., & Collins, K. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. Washington D.C: National Academy Press.
- Cameron, A. E. (1928). A comparative study of the mathematical ability of boys and girls in secondary schools.
- Campbell, D. T. (1960a). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, 67, 380-400.
- Campbell, D. T. (1960b). Recommendations for APA test standards regarding construct, trait, and discriminant validity. *American Psychologist*, 15, 546-553.
- Carlsson, I., Wendt, P.E., & Risberg, J. (2001). On the neurobiology of creativity. Differences in frontal activity between high and low creative subjects. *Neuropsychologia*, 38, 873-885.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Carroll, J. B. (1996). Mathematical Abilities: Some results from factor analysis. In R. J. Sternberg & T. Ben Zeev. (Eds.). *The nature of mathematical thinking*. (pp. 1-26). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 1, 33-81.
- Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.

- Chi, M. T. H., Glaser, R., & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Connell, M. W., Sheridan, K., & Gardner, H. (2003). On abilities and domains. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The psychology of abilities, competencies, and expertise* (pp. 126-155). New York: Cambridge University Press.
- Csikszentmihalyi, M. (1997). *Creativity: Flow and the psychology of discovery and invention*. New York: Harper Perennial.
- Davidson, J. E. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight*. Cambridge, MA: The Bradford Book.
- Davidson, J. E. & Sternberg, R. J. (1984). The role of insight in intellectual giftedness. *Gifted Child Quarterly*, 28, 58-64.
- Davidson, J. E. & Sternberg, R. J. (1986). What is insight? *Educational Horizons*, 64, 177-179.
- DeGroot, A. D. (1965). *Thought and choice in chess*. The Hague, the Netherlands: Mouton.
- Dollinger, S. J., Levin, E. L., & Robinson, A. E. (1991). The word association Implications tests. *Journal of Personality Assessment*, 57, 368-380.
- Dykes, M., & McGhie, A. (1976). A comparative study of attentional strategies of Schizophrenic and highly creative normal subjects. *British Journal of Psychiatry*, 128, 50-56.
- Embretson, S. E., & McCollam, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Ericsson, K.A. (2003). The search for general abilities and basic capacities: Theoretical implications from the modifiability and complexity of mechanisms mediating expert performance. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The psychology of abilities, competencies, and expertise* (pp. 93-125). New York: Cambridge University Press.
- Eysenck, H. J. (1993). Creativity and personality: Suggestions for a theory. *Psychological Inquiry*, 4, 147-178.
- Eysenck, H. J. (1995). *Genius: The natural history of creativity*. Cambridge: Cambridge University Press.

- Ferrara, R. A., Brown, A. L., & Campione, J. C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. *Child Development*, 57, 1087-1099.
- Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: theory, research and applications*. Cambridge, MA: MIT Press.
- Gardner, H. (1983). *Frames of mind*. New York: Basic Books.
- Gardner, H. (1999). *Intelligence reframed*. New York: Basic Books.
- Gray, E., Pinto, M., Pitta, D., & Tall, D. (1999). Knowledge construction and divergent thinking in elementary and advanced mathematics. *Educational Studies in Mathematics*, 38, 111-133.
- Griffin, S. (2000). The development of math competence in the preschool and early school years: Cognitive foundations and instructional strategies. In J. M. Royer (Ed.), *Mathematical cognition* (pp. 1-32). Greenwich, CT: Information Age Publishing.
- Grigorenko, E. L. (2003). Expertise and mental disabilities. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The psychology of abilities, competencies, and expertise* (pp. 156-185). New York: Cambridge University Press.
- Hadamard, J. (1945). *The psychology of invention in the mathematical field*. New York: Dover Publications, Inc.
- Hardy, G. H. (1940). *A mathematician's apology*. London: Cambridge University press.
- Heller, K. A., Monks, F. J., Sternberg, R. J., & Subotnik, R. F. (2000). *International handbook of giftedness and talent*. New York: Elsevier.
- Hinsley, D. A., Hayes, J. R., & Simon, H. A. (1977). From words to equations: Meaning and representation in algebra word problems. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 89-106). Hillsdale, NJ: Erlbaum.
- Hoppe, K. D., & Kyle, N. L. (1990). Dual brain, creativity and health. *Creativity Research Journal*, 3, 150-157.
- Jausovec, N. (2000). Differences in cognitive processes between gifted, intelligent, creative, and average individuals while solving complex problems: An EEG study. *Intelligence*, 3, 213-237.
- Klauer, K. C., Meiser, T., & Naumer, B. (2000). Training propositional reasoning. *The Quarterly Journal of Experimental Psychology*, 53, 868-895.

- Kris E. (1952). *Psychoanalytic explorations in art*. New York: International Universities Press.
- Linn, R. L. (1993). The use of differential item functioning statistics: A discussion of current practice and future implications. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning*. (pp. 349-64). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence*. (pp. 285 -340). New York: Cambridge University Press.
- Martindale, C. (1981). *Cognition and consciousness*. Homewood, IL: Dorsey.
- Martindale, C. (1989). Personality, situation, and creativity. In J. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 211–228). New York: Plenum.
- Martindale, C. (1995). Creativity and connectionism. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 249–268). Cambridge, MA: MIT Press.
- Martindale, C. (1999a). Biological bases of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 137–152). New York: Cambridge University Press.
- Martindale, C. (1999b). Genetics. In M. A. Runco & S. R. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 767–771). San Diego, CA: Academic Press.
- Martindale, C., & Hines, D., (1975). Creativity and cortical activation during creative, intellectual and EEG feedback tasks. *Biological Psychology*, 3, 91-100.
- Martindale, C., Hines, D., Mitchell, L., & Covello, E. (1984). EEG alpha asymmetry and creativity. *Personality and Individual Differences*, 5, 77–86.
- Mayer, R. E. (1991). *Thinking, problem solving, cognition*. 2<sup>nd</sup> ed. New York: Freeman.
- Mayer, R. E. (2003). Mathematical problem solving. In J. M. Royer (Ed.). *Mathematical cognition*. (pp.69-92). Greenwich, Connecticut: Information Age Publishing Inc.
- McDaniel, E., & Lawrence, C. (1990). *Levels of cognitive complexity: An approach to the measurement of thinking*. New York: Springer-Verlag.
- Mendelsohn, G. A., & Griswold, B. B. (1966). Assessed creative potential, vocabulary level, and sex as predictors of the use of incidental cues in verbal problem solving. *Journal of Personality and Social Psychology*, 4, 423–431.

- Miyazaki, M. (2000). Levels of proof in lower secondary school mathematics. *Educational Studies in Mathematics*, 41, 47-68.
- Ni, Y. (1998). Cognitive structure, content knowledge, and classificatory reasoning. *The Journal of Genetic Psychology*, 159, 280-296.
- Novick, L. R. (1992). The role of expertise in solving arithmetic and algebra word problems by analogy. In J. I. D. Campbell (Ed.). *The nature and origins of mathematical skills* (pp.155-188). Amsterdam, the Netherlands: North-Holland.
- Nunez, P. L. (1995). Mind, brain, and electroencephalography. In P. L. Nunez (Ed). *Neocortical dynamics and human EEG rhythms*. New York: Oxford Univiversity Press.
- Pegg, J. (2003). Assessment in mathematics: A developmental approach. In J. M. Royer, (Ed.). *Mathematical cognition*. (pp. 227-259). Greenwich, CT: Information Age Publishing.
- Petsche, H. (1997). EEG coherence and mental activity. In F. Angeleri., S. Butler, S. Giaquinto, & J. Majakowski (Eds). *Analysis of the electrical activity of the brain*, (pp. 141–168). Wiley: Chichester.
- Petsche, H. (1996). Approaches to verbal, visual, and musical creativity by EEG coherence analysis. *International Journal of Psychophysiology*, 24, 145–159.
- Poincare, H. (1952a). Science and hypothesis. In S. J. Gould (series editor, 2001), *Henri Poincare*. New York: The Modern Library.
- Poincare, H. (1952b). Science and method. In S. J. Gould (series editor, 2001), *Henri Poincare*. New York: The Modern Library.
- Poincare, H. (1958). The value of science. In S. J. Gould (series editor, 2001), *Henri Poincare*. New York: The Modern Library.
- Polya, D. (1954a). *Induction and analogy in mathematics*. Princeton, NJ: Princeton University Press.
- Polya, D. (1954b). *Patterns of plausible inference*. Princeton, NJ: Princeton University Press.
- Polya, D. (1957). *How to solve it*. 2<sup>nd</sup> ed. NJ: Princeton University Press.
- Polya, D. (1962). *Mathematical discovery*. New York: John Wiley & Sons, Inc.

- Poreh, A. M., & Whitman, R. D. (1991). Creative cognitive processes and hemispheric specialization. *The Journal of Creative Behavior*, 25, 2, 169-179.
- Ostad, S. A. (1998). Developmental differences in solving simple arithmetic word problems and simple number-fact problems: A comparison of mathematically normal and mathematically disabled children. *Mathematical Cognition*, 4, 1-19.
- Rawlings, D. (1985). Psychoticism, creativity, and dichotic shadowing. *Personality and Individual Differences*, 6, 737-742.
- Rogers, A. L. (1918). *Experimental tests of mathematical ability and their prognostic value*. New York: Teachers College.
- Sak, U. (2004). *A cognitive expertise approach for assessing mathematical talent to identify mathematically gifted students*. Unpublished manuscript. University of Arizona, Tucson.
- Sak, U., & Maker, C. J. (2003, November). *Children's creative cognitive development in mathematics*. Paper presented at the 50<sup>th</sup> Annual Conference of the National Association for Gifted Children (NAGC), Indianapolis, IN.
- Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge representation in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 484-494.
- Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: when language overshadows insight. *Journal of Experimental Psychology: General*, 122, 166-183.
- Shaw, G. A. (1992). Hyperactivity and creativity: The tacit dimension. *Bulletin of the Psychonomic Society*, 30, 157-160.
- Shaw, G. A., & Conway, M. (1990). Individual differences in non-conscious processing: The role of creativity. *Personality and Individual Differences*, 11, 407-418.
- Shaw, J. B. (1918). *Lectures on the philosophy of mathematics*. London: The Open Court Publishing Company.
- Sheppard, W.D., & Boyer, R.W., (1990). Parietal EEG coherences a predictor of semantic priming effects. *Brain and Language*, 39, pp. 57-68.
- Silver, E. A. (1981). Recall of mathematical problem information: Solving related problems. *Journal for Research in Mathematics Education*, 12, 54-64.

- Simonton, D. K. (1989). Chance-configuration theory of scientific creativity. In B. Gholson, W. R. Shadish, R. A. Neimeyer, & A.C. Houts (Eds.), *Psychology of science* (pp. 170–213). Cambridge: Cambridge University Press.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan.
- Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical thinking*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Sternberg, R. J. (1982). Natural, unnatural and supernatural concepts. *Cognitive Psychology*, 14, 451-488.
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Penguin Books.
- Sternberg, R. J. (1996). What is mathematical thinking? In R. J. Sternberg and T. Ben-Zeew (Eds.). *The nature of mathematical thinking*. (pp. 303-319). New Jersey: Lawrence Erlbaum Associates, Publishers.
- Sternberg, R. J. (1997). *Successful intelligence*. New York, NY: Plume.
- Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher*, 27, 11-20.
- Sternberg, R. J. (2000). Patterns of Giftedness: A Triarchic Analysis. *Roeper Review*, 22, 231-235.
- Sternberg, R. J. (2002). Beyond g. In R. J. Sternberg & E. L. Grigorenko (Eds.), *The general factor of intelligence*. (pp. 447-479). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Sternberg, R. J., Castejon, J., Prieto, M. D., Hautakami, J., & Grigorenko, E. L. (2001). Confirmatory factor analysis of the Sternberg Triarchic Abilities Test (multiple-choice items) in three international samples. An empirical test of the thriarchic theory of intelligence. *European Journal of Psychological Assessment*, 17, 1-16.
- Sternberg, R. J., & Davidson, J. E. (1986). *Conceptions of giftedness*. New York: Cambridge University Press.
- Sternberg, R. J. & Grigorenko, E. L. (2003). *The psychology of abilities, competencies, and expertise*. New York: Cambridge University press.

- Sternberg, R. J., Grigorenko, E. L., Ferrari, M., & Clinkenbeard, P. (1999). A triarchic analysis of an aptitude-treatment interaction. *European Journal of Psychological Assessment*, 15, 1-11.
- Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd*. New York: The Free Press.
- Thatcher, R.W., & Walker, R.A. (1985). EEG coherence and intelligence in children. *Electroencephalography and Clinical Neurophysiology*, 61, 161.
- Torrance, E. P. (1968). A longitudinal examination of the fourth-grade slump in creativity. *Gifted Child Quarterly*, 12, 195-199.
- Thurstone, L. L. (1938). *Primary mental abilities*. Psychometric monographs. Chicago.
- Thurstone, L. L. (1950). *Multiple factor analysis*. Chicago: University of Chicago Press.
- Vartanian, O., Martindale, C., & Kwiatkowski, J. (2003). Creativity and inductive reasoning: The relationship between divergent thinking and performance on Wason's 2-4-6 task. *The Quarterly Journal of Experimental Psychology*, 56, 641-656.
- Weisberg, R. W. (1999). Creativity and knowledge: A challenge to theories. In R. J. Sternberg (Ed.), *Handbook of creativity*. (pp. 226-250). Cambridge, UK: Cambridge University Press.
- Werdelin, I. (1958). *The mathematical ability: Experimental and factorial studies*. Sweden: Lund.