The Journal of International Lingual, Social and Educational Sciences
Year: 2016, Volume: 2, Number: 2

Submitted: September 2016
Revised: November 2016
Accepted: December 2016

IDENTİFYİNG AND INTERPRETİNG SUBJECTİVE WEİGHTS FOR COGNİTİVE AND PERFORMANCE CHARACTERİSTİCS OF MATHEMATİCAL LEARNİNG DİSABİLİTY: AN APPLİCATİON OF A RELATİVE MEASUREMENT METHOD

Onder KOKLU[1], Elizabeth JAKUBOWSKI[2], Tayfun SERVİ[3], Jiajing HUANG[4]

Abstract

This study utilizes an evaluation model AHP (analytic hierarchy process) which prioritized the relative weights of three general subtypes of mathematical disability (MD), Semantic Memory, Procedural, and Visuospatial in order to analyze and explain underlying cognitive and performance features of FCAT (Florida Comprehensive Assessment Test) benchmarks and corresponding items for grades 6-8 in all mathematics categories. For this purpose, extensive review of the literature has been conducted on mathematical disability to determine subtypes of learning disability in mathematics. Afterwards, a multi-step AHP approach is adopted to obtain the relative weights of criteria (subtypes of learning disability) by linking the independent evaluations of four content area experts for each benchmark. The results indicate that semantic memory deficiency is the dominant subtype of mathematical learning disability on vast majority of benchmarks. Another important finding of this study is that effect of visuospatial deficiency increases from grade 6 to grade 8. In addition, effect of procedural deficiency does not show big variability among reporting categories, although it has the highest effect on number categories.

Key words: *AHP, FCAT, subtypes of mathematical disability*

INTRODUCTION

It is claimed by Saaty (2008) that there are two different kinds of topology in the measurement area. According to Saaty (2008), one of them is called metric topology and the other is order topology. The metric topology, as we all have been engaged in from early years of our education, is dealing with measuring certain element or objects on a scale with a determined unit. And it involves a uniformly applied origin in order to measure all objects with respect to the given properties. Metric properties, according to Saaty (2008) were the main paradigm until decision making was introduced as a scientific method. There were little known about systematic ways of ordering elements and decision making process. Therefore, the second kind of topology is mentioned as order topology which focuses on ordering objects by measuring the dominance of one element over others with respect to a common attribute. Thus Saaty (2008) expressed that order topology is not metric topology and cannot be derived metrically in terms of how close things are. Thus, it is clear that relative importance of objects or elements to put in order depends on our subjective judgments which are led by our values and preferences. Therefore judgments are the main characters and always come first before measurement takes place. Numerical measurements of the outcome are called priorities (Satty, 2008).

In such an order topology, MCDM (Multiple Criteria Decision Making) methods have been widely used in research studies in order to answer both theoretical and practical questions (Ginevičius, 2008). MCDM involves systematic ways of making informed decisions over the available alternatives that are characterized by multiple criteria (Hwang and Yoon, 1981). MCDM models are introduced and developed in order to satisfy the need to evaluate the level of dominance of one alternative over others in a discrete set of alternatives considering multiple criteria (Choo et al.,

[1] Adiyaman University, Faculty of Education, Adiyaman, TURKEY. okoklu@adiyaman.edu.tr
[2] Florida State University, College of education, Tallahassee, USA, ejakubow@fsu.edu
[3] Adiyaman University, Faculty of Business and Administration, Adiyaman, Turkey, tselvi@adiyaman.edu.tr
[4] Florida State University, College of education, Tallahassee, USA, jhuang@fsu.edu

1999). The overall preference values or respective dominance values of the alternatives are determined based on the weights of each criterion on which the alternatives are evaluated. Although review of the literature disclosed that many different weighting methods have been proposed for obtaining criteria weights there are four kinds of general methods mostly used in the research studies because of their simplicity and effectiveness to obtain weights: Delphi Method, Rank Order Centroid Method, Ratio Method, Pairwise Comparison Method

The Delphi Method is way of obtaining the opinions of a group of experts on an issue by conducting series of inquiring communications (Chan et al. 2001). This method is basically useful for achieving a consensus among experts, who are anonymous, by exchanging ideas opinions and viewpoints on factors (Chan et al. 2001). All experts give weights to each factor with their reasoning. In this way, other experts can evaluate the weights based on the reasons given and accept, modify, or reject those reasons and weights.

The Rank Order Centroid Method is a simple way of giving weight to a number of items ranked according to their importance. The rational for using this method, according to Chang (2004), is that decision-makers usually can rank items much more easily than give weight to them. By this method ranks are taken as inputs and converted to weights for each of the criteria. Following formula is suggested to be used for the conversion:

$$W_i = \frac{1}{M}\sum_{n=i}^{M}\frac{1}{n}$$

Where M is the number of items and W_i is the weight for i^{th} item. For example, if there are 4 items, the item ranked first will be weighted (1 + ½ + 1/3 + ¼) / 4 = 0.52, the second will be weighted (1/2 + 1/3 + ¼) / 4 = 0.27, the third (1/3 + ¼) / 4 = 0.15, and the last (1/4) / 4 = 0.06. (Chang 2004).

Another method of obtaining criteria weight is the Ratio method. This method is another simple way of calculating weights by ranking all the criteria regarding their importance. Then weights are assigned to each item based on its rank. The important rule is that the lowest ranked item should be given a weight of 10. Weights of other items should be multiples of 10. Finally, normalization procedure is applied for the raw weights.

Pairwise comparison method allows the decision-maker to compare each item with the rest of the criteria and give a preferential level to the item in each pairwise comparison (Chang 2004). According to Brown and Peterson (2009) paired comparisons provides rich information on respondents' relative judgments about a set of items. The method of paired comparisons has a long history, Brown and Peterson (2009) reported that a text on experimental psychology by Titchener in 1901 covered paired comparisons. Later, several studies of Thurstone in the late 1920's (Brown and Peterson, 2009) brought considerable attention to the method with his psychological scaling proposals. Brown and Peterson (2009) also mentioned that some indications of the method of paired comparisons are also found in psychometric textbooks (such as Guilford 1954; Nunnally 1976; Torgerson 1958). These Psychologists argue that it is easier and more accurate to express one's opinion on only two alternatives than simultaneously on all the alternatives. It also allows consistency cross checking between the different pair-wise comparisons (Ishizaka and Labib, 2011). The most widely used paired comparisons method in research studies is the Analytic Hierarchy Process (AHP).

The Analytic Hierarchy Process (AHP) is a multi-criteria decision making (MCDM) process which is based upon pairwise comparison method. Weighting procedure in AHP is a ratio weighting method which is an algebraic and direct weighting method. Analytic Hierarchy Process (AHP) developed by Thomas Saaty. The oldest reference that we have found is a paper in the Journal of Mathematical Psychology (Saaty, 1977) which described the method. The majority of the applications of AHP are described more clearly in the book called "The Analytic Hierarchy Process" (Satty, 1980). This book provides a sketch of the major directions in methodological developments in this field. AHP has been widely used in research studies in the course of last decade (Haghighi, Divandari, & Keimasi, 2010; Seçme, Bayrakdaroglu, & Kahraman, 2009; T.S. Li & Huang, 2009; Yang, Chuang, & Huang, 2009; Sen & Çinar, 2010; Chamodrakas, Batis, & Martakos, 2010; Pan, 2009) in various areas.

Learning Disability Framework. Although there is limited literature about classification of mathematics learning disabilities, detailed review of literature (Geary, Hoard, Byrd-Craven & DeSoto, 2004; Geary, 2003; Jordan & Montani, 1997; Ackerman & Geary, 1993; Goldman, Hitch & McAuley, 1991; Goldman, Pellegrino, & Mertz, 1988) revealed that learning disabilities in mathematics are caused by three types of cognitive deficiencies: Deficiencies in Semantic memory, Procedural Deficiencies (primitive procedures used in computations),Visual and Spatial Deficiencies

The work of David C. Geary help to analyze and explain underlying cognitive and performance features of FCAT mathematics standards for grades 6-8. The model describing the subtypes of deficits which can result in mathematical learning disabilities was introduced by Geary (2003) as a result of several investigations (1993-2003) in which theoretical model was applied by using experimental methods in order to find out possible deficits in the ability to represent or process information.

Geary (2003) defined "Subtypes of Learning Disabilities in Mathematics" as *"…the integration of cognitive and behavioral genetic studies of individual differences in mathematical abilities provided clues as to possible sources of the problem solving characteristics of children with MD and resulted in a taxonomy of three general subtypes of Mathematical disability (MD)*" (p. 204). Table 1 represents the definitions of cognitive and performance characteristics of three general subtypes of mathematical disability (Geary, 2003).

Table 1.

Subtypes of Learning Disabilities in Mathematics

Subtype	Cognitive and Performance Features
Semantic Memory	• Difficulties retrieving mathematical facts. • What facts are retrieved, there is a high error rate. • Reaction Time (RT) for correct retrieval is unsystematic.
Procedural	• Relatively frequent use of developmentally immature procedures. • Frequent errors in the execution of procedures • Poor understanding of the concepts underlying procedural use. • Difficulties sequencing the multiple steps in complex procedures.
Visio-spatial	• Difficulties in spatially representing numerical and other forms mathematical information and relationships. • Frequent misinterpretation or misunderstanding of spatially represent information.

Statement of Purpose. Although previous studies have identified cognitive and performance characteristics of general subtypes of mathematical disability, the evaluation of the relative importance of these factors (subtypes) in a high-stake criterion-referenced assessment has not been empirically determined. Therefore, the main aim of this research study is to analyze and explain underlying cognitive and performance features of FCAT mathematics standards for grades 6-8 in terms of the model describing the subtypes of learning disabilities in mathematics. More specifically the aim of this study is to determine relative weights of the subtypes of learning disabilities in mathematics on each FCAT benchmarks for grades 6-8 in all mathematics categories tested.

METHOD

This study utilizes an evaluation model AHP (analytic hierarchy process) which prioritized the relative weights of three general subtypes of mathematical disability (MD), Semantic Memory, Procedural, and Visio-spatial in order to analyze and explain underlying cognitive and performance features of FCAT benchmarks for grades 6-8 in all mathematics categories tested (Table 2). The Florida Comprehensive Assessment Test (FCAT) is a criterion-referenced test that measures student achievement of the Next Generation Sunshine State Standards in reading, mathematics, science, and writing. All Florida schools are required to teach the Next Generation Sunshine State Standards, and the FCAT provides parents/guardians, teachers, policy makers, and the general public with an understanding of how well students are learning these standards. The best understanding of a student's academic achievement comes from looking at multiple pieces of evidence (including test scores) collected over time (FLDOE,)

Table 2.

FCAT mathematics categories and number of benchmarks for grades 6-8

Grade	Name of the Category	Number of Benchmarl
6	Number: Fraction, Ratios / Proportional Relationships and Statistics	9
	Expression and Equation	4
	Geometry and Measurement	3
7	Number: Base Ten	5
	Ratios / Proportional Relationship	6
	Geometry and Measurement	6
	Statistics and Probability	4
8	Number: Operation, Problems and Statistics	5
	Expression, Equation and Function	7
	Geometry and Measurement	5

A multi-step AHP approach is adopted to determine the relative weights of the criteria (three general subtypes of mathematical disability). The first of these steps involves the construction of a hierarchy that includes the relevant aspects of the criteria and incorporates the individual preferences (judgments) that reflect the relative importance of the alternatives through a pairwise comparisons matrix, proposed by Saaty (1980). In the second step, individual priorities (judgments) of alternatives are synthesized by means of an aggregation procedure (the weighted geometric means were extensively used), to obtain the *final priorities* of the alternatives. Finally, third step, provides computational procedures of obtaining weights for each element of the hierarchy and computational procedures of checking consistency of judgments. The eigenvalue method (Saaty, 1980) involving the column normalization method and row arithmetic mean method was employed to obtain weights which are measures of the priorities of the elements (criteria).

2.1. Scaling the relative importance of the criteria. In terms of construction of a hierarchy that includes the relevant aspects of the criteria, conventional structure of AHP (nine-point rating scale) was used (Table 3). Then four content area experts independently evaluated each benchmark based on the hierarchical structure described above. Evaluation of a benchmark was made considering both statements of each benchmark and sample items under each benchmark from a holistic perspective. Evaluators have more than 3 years of experience in the area of mathematics education. They were asked to comment on the meaningfulness, relevance, and clarity of the criteria. Therefore, the evaluating criteria have confirmed content validity.

Table 3.

AHP nine-point rating scale.

Importance Values	Value Descriptions
1	Both factors (Criteria) have equal importance.
3	1st factor is little more important than 2nd factor.
5	1st factor more important than 2nd factor.
7	1st factor much more important than 2nd factor.
9	1st factor incredibly more important than 2nd factor.

Individual preferences (judgments) that reflect the relative importance of the alternatives can be differ from each other. Pairwise-comparison matrices from 4 different evaluators are as follows:

$$E1_{3x3} = \begin{array}{cccc} & C_1(S) & C_2(P) & C_3(V) \\ C_1(S) & a_{1,1} & a_{1,2} & a_{1,3} \\ C_2(P) & a_{2,1} & a_{2,2} & a_{2,3} \\ C_3(V) & a_{3,1} & a_{3,2} & a_{3,3} \end{array} \qquad E2_{3x3} = \begin{array}{cccc} & C_1(S) & C_2(P) & C_3(V) \\ C_1(S) & b_{1,1} & b_{1,2} & b_{1,3} \\ C_2(P) & b_{2,1} & b_{2,2} & b_{2,3} \\ C_3(V) & b_{3,1} & b_{3,2} & b_{3,3} \end{array}$$

$$E3_{3x3} = \begin{array}{cccc} & C_1(S) & C_2(P) & C_3(V) \\ C_1(S) & c_{1,1} & c_{1,2} & c_{1,3} \\ C_2(P) & c_{2,1} & c_{2,2} & c_{2,3} \\ C_3(V) & c_{3,1} & c_{3,2} & c_{3,3} \end{array} \qquad E4_{3x3} = \begin{array}{cccc} & C_1(S) & C_2(P) & C_3(V) \\ C_1(S) & d_{1,1} & d_{1,2} & d_{1,3} \\ C_2(P) & d_{2,1} & d_{2,2} & a_{2,3} \\ C_3(V) & d_{3,1} & d_{3,2} & d_{3,3} \end{array}$$

If three elements (criteria) are presented by C_1, C_2, C_3, pairwise comparison of C_i, C_j are expressed by following four pairwise-comparison matrices for each evaluator:

$$E1_{3x3} = (a_{i,j}), (i, j = 1,2,3)$$

$$E2_{3x3} = (b_{i,j}), (i, j = 1,2,3)$$

$$E3_{3x3} = (c_{i,j}), (i, j = 1,2,3)$$

$$E4_{3x3} = (d_{i,j}), (i, j = 1,2,3)$$

Considering the AHP rating scale (table 2), values of $a_{i,j}$ $b_{i,j}$ $c_{i,j}$ $d_{i,j}$ are defined by following rules:

Rule 1. $a_{i,i} = 1,$ for all i values.

Rule 2. If $a_{i,j} = \alpha,$ then $a_{j,i} = 1/\alpha$, $\alpha \neq 0$

Rule 3. If C_i and C_j were determined as having equal importance, $a_{i,j} = 1,$ and $a_{j,i} = 1,$.

2.2. Determining the joint decision. There are two ways to analyze a joint decision problem in the classical literature on AHP (Ramanatham and Ganesh, 1994; Forman and Peniwati, 1998):

1. Aggregation of Individual Judgments where a new pair wise comparison matrix for the group is constructed aggregating the individual judgments by means of consensus, voting or statistical procedures such as, for instance, the weighted geometric mean. From this matrix, the priority vector is then calculated following any of the existing prioritization procedures.
2. Aggregation of Individual Priorities where the individual priorities are aggregated in order to obtain the priority of the group, with the usual aggregation procedure being the weighted geometric mean.

In this study, second method was used to obtain the priority of the group and to determine joint decision. Geometric mean of $a_{i,j}$ $b_{i,j}$ $c_{i,j}$ $d_{i,j}$ values from independently constructed pairwise comparisons matrices was used for final $f_{i,j}$ value in joint decision matrix.

$$f_{i,j} = \sqrt[4]{a_{i,j} \times b_{i,j} \times c_{i,j} \times d_{i,j}} \quad \text{and} \quad f_{j,i} = \sqrt[4]{\frac{1}{a_{i,j}} \times \frac{1}{b_{i,j}} \times \frac{1}{c_{i,j}} \times \frac{1}{d_{i,j}}}$$

Therefore,

$$f_{j,i} = \frac{1}{f_{i,j}}$$

2.3. Determining the weights and consistency. The eigenvalue method (Saaty, 1980) involving the column normalization method and row arithmetic mean method was employed to obtain weights which are measures of the priorities of the elements (criteria).

Let $f_{i,j}$ be the final value in joint decision matrix:

$$\begin{bmatrix} 1 & f_{1,2} & f_{1,3} \\ f_{2,1} & 1 & f_{2,3} \\ f_{3,1} & f_{3,2} & 1 \end{bmatrix} \xrightarrow{\text{Column Normalization}} \begin{bmatrix} w_{1,1} & w_{1,2} & w_{1,3} \\ w_{2,1} & w_{2,2} & w_{2,3} \\ w_{3,1} & w_{3,2} & w_{3,3} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$\sum_{i=1}^{3} f_{i,1} \quad \sum_{i=1}^{3} f_{i,2} \quad \sum_{i=1}^{3} f_{i,3} \qquad w_{i,j} = \frac{f_{i,j}}{\sum_{i=1}^{3} f_{i,j}} \qquad w_i = \frac{1}{n}\sum_{j=1}^{3} w_{i,j}$$

Saaty (1990) provides a consistency index to measure any inconsistency within the judgments in each pair-wise comparison matrix as well as for the entire hierarchy. The consistency of judgment matrices are tested using two equations shown below. If the calculated CR (Consistency Ratio) of a pair-wise comparison matrix is less than 0.1, the consistency of the pair-wise judgment can be thought of as being acceptable.

The Consistency Index (CI) is formulated as follows:

$$CI = \frac{\lambda_{\max} - n}{n - 1}$$

where λmax is the maximum eigenvalue, and n is the dimension of matrix.

Accordingly, the "Consistency Ratio" (CR) can be computed with the use of following equation: $CR = \frac{CI}{RI}$

Where RI means “Random Consistency Index” (Table 4) and changes with n (number of criteria).

Table 4.

Random Consistency Index (RI)

Number Criteria	1	2	3	4	5	6	7	8	9	10
RI	0,00	0,00	0,58	0,90	1,12	1,24	1,32	1,41	1,45	1,49

RESULTS AND DISCUSSION

Major purpose of this study is to determine relative weights of the subtypes of learning disabilities in mathematics on each FCAT benchmarks for grades 6-8 in all reporting categories of mathematics. The relative weights of three general subtypes of mathematical disability (MD), Semantic Memory, Procedural, and Visio-spatial were obtained by the application of a multi-step AHP approach in order to analyze and explain underlying cognitive and performance features of FCAT benchmarks for grades 6-8 in all mathematics categories. First, AHP procedures were employed to obtain 54 pairwise comparison matrices. Then, we computed weights for 54 individual benchmarks in all the grade levels (Table 5).

We observed an average of 0.486 with a standard deviation of 0.108 under “Semantic Memory” subtype, and average of 0.298 with a standard deviation of 0.154 under “Procedural” subtype and average of 0.216 with a standard deviation of 0.129 under “Visual Spatial” subtype. After obtaining weights, which are measures of the priorities of the elements (criteria), for each benchmark, arithmetic mean method was employed to obtain mean weights for each reporting category. Mean weights and other descriptive information for each reporting category and grade were summarized in table 6.

Table 5.

Weights of All Benchmarks by Grade and Reporting Categories

Grade	Reporting Category	Benchmark	Weights S	P	V
6	Number: Fracti Ratios/Proportional Relationships, a Statistics	6A.1.1	0,63	0,20	0,17
		6A.1.3	0,24	0,61	0,15
		6A.2.1	0,39	0,50	0,11
		6A.2.2	0,51	0,35	0,14
		6A.5.1	0,48	0,42	0,11
		6A.5.2	0,46	0,29	0,25
		6A.5.3	0,66	0,24	0,10
		6S.6.1	0,64	0,22	0,14
		6S.6.2	0,59	0,31	0,10
	Expression and Equation	6A.3.1	0,48	0,40	0,12
		6A.3.2	0,35	0,41	0,25
		6A.3.5	0,64	0,26	0,10
		6A.3.6	0,47	0,25	0,29
	Geometry a Measurement	6G.4.1	0,64	0,20	0,16
		6G.4.2	0,64	0,10	0,26
		6G.4.3	0,54	0,20	0,26
7	Ratios/Proportional Relationship	7A.1.1	0,61	0,22	0,16
		7A.1.2	0,42	0,46	0,13
		7A.1.3	0,56	0,20	0,24
		7A.1.4	0,61	0,22	0,16
		7A.1.5	0,61	0,17	0,22
		7A.1.6	0,41	0,50	0,09
	Number: Base Ten	7A.3.1	0,45	0,36	0,18
		7A.3.2	0,30	0,60	0,09
		7A.3.3	0,37	0,52	0,11
		7A.3.4	0,42	0,45	0,12
		7A.5.1	0,56	0,35	0,09
	Geometry a Measurement	7G.2.1	0,61	0,13	0,26
		7G.2.2	0,41	0,14	0,44
		7G.4.1	0,39	0,13	0,48
		7G.4.2	0,31	0,09	0,59
		7G.4.3	0,49	0,09	0,42
		7G.4.4	0,48	0,43	0,09
	Statistics and Probability	7P.7.1	0,48	0,43	0,09
		7P.7.2	0,61	0,31	0,08
		7S.6.1	0,71	0,16	0,14
		7S.6.2	0,46	0,25	0,29
8	Expression, Equation, a Function	8A.1.1	0,44	0,31	0,25
		8A.1.2	0,50	0,11	0,39
		8A.1.3	0,46	0,13	0,42
		8A.1.5	0,46	0,13	0,41
		8A.1.6	0,45	0,10	0,45
		8A.4.1	0,31	0,59	0,11
		8A.4.2	0,27	0,46	0,27
	Number: Numb Operation, Problems, a Statistics	8S.3.1	0,49	0,10	0,40
		8S.3.2	0,56	0,31	0,12
		8A.6.1	0,43	0,47	0,10
		8A.6.2	0,38	0,47	0,14
		8A.6.4	0,35	0,55	0,10
	Geometry a Measurement	8G.2.1	0,48	0,36	0,16
		8G.2.2	0,55	0,10	0,35
		8G.2.3	0,52	0,11	0,37
		8G.2.4	0,47	0,21	0,32
		8G.5.1	0,53	0,38	0,10

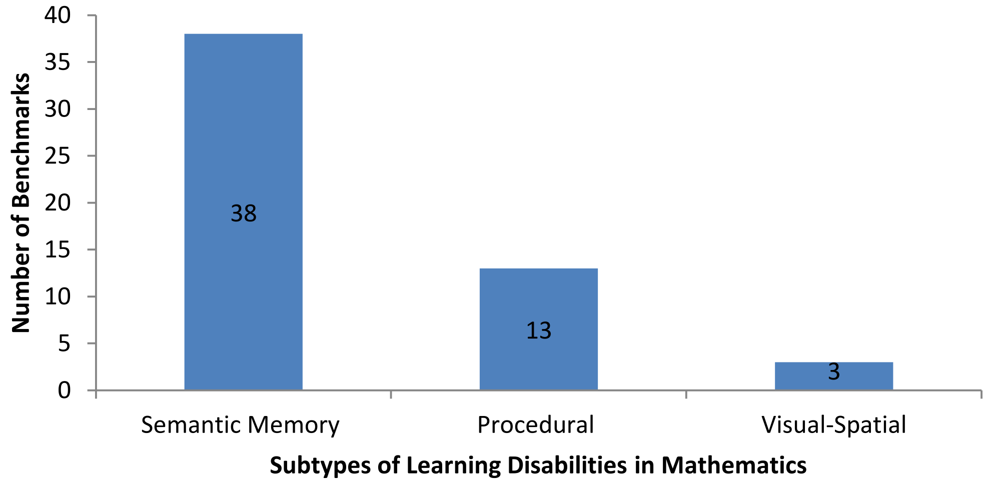
Table 6.

Descriptive Information by Grades and Reporting Categories

Grade	Reporting Category	S		P		V	
		M	SD	M	SD	M	SD
6	Number: Fraction, Ratios / Proportion Relationships and Statistics	0.51		0.35		0.14	
	Expression and Equation	0.48		0.33		0.19	
	Geometry and Measurement	0.61		0.17		0.23	
	6th Grade General	0.52		0.31		0.17	
7	Number: Base Ten	0.42		0.46		0.12	
	Ratios / Proportional Relationship	0.54		0.30		0.17	
	Geometry and Measurement	0.45		0.17		0.38	
	Statistics and Probability	0.56		0.29		0.15	
	7th Grade General	0.49		0.30		0.21	
8	Number: Operation, Problems and Statistics	0.44		0.38		0.17	
	Expression, Equation and Function	0.41		0.26		0.33	
	Geometry and Measurement	0.51		0.23		0.26	
	8th Grade General	0.45		0.29		0.26	

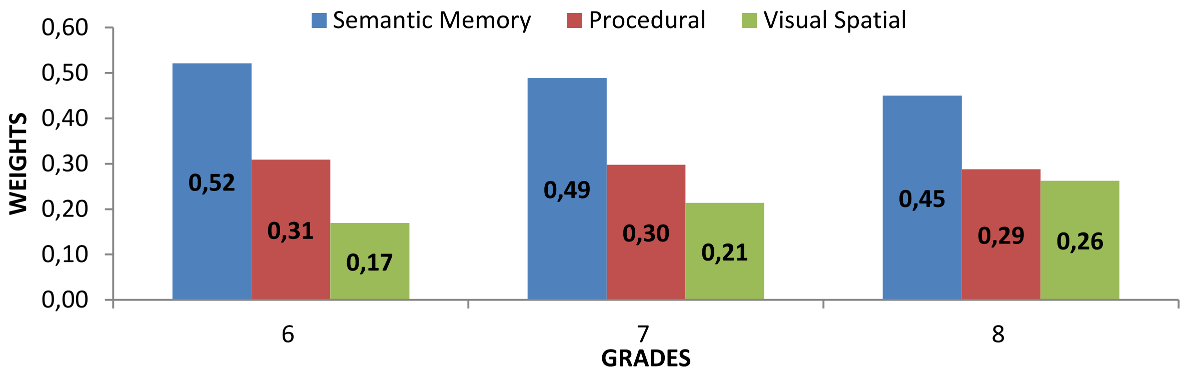
Semantic memory deficiency was identified as a dominant subtype of mathematical learning disability on vast majority of benchmarks considering the obtained relative weights of all benchmarks. Figure 1 presents the overall pattern of dominance by each of the three subtypes of learning disability. More specifically, results indicated that semantic memory deficiency has the major effect on 38 of the 54 benchmarks, while procedural deficiency has the major effect on 13 benchmarks and visual-spatial deficiency has the major effect on 3 benchmarks. Results also indicated that semantic memory deficiency has the major effect on 9 out of the 10 reporting categories, while procedural deficiency has the major effect on only1 reporting category, which is 7th grade "Number: Base Ten" category, and visual-spatial deficiency has no major effect on any of the reporting categories.

Figure 1. Number of Benchmarks Dominated By Subtypes of Learning Disability



Not surprisingly, semantic memory deficiency have also the greatest impact on all grade levels (Figure 2) comparing to Procedural and Visual-Spatial deficiencies. That means students who have this type of learning disability more likely to be expected to get low scores in FCAT because of having difficulties in retrieval of the mathematical facts or the high error rate in correct retrieval.

Figure 2. Learning disability subtype mean weights by grade levels



More detailed investigation on data revealed that deficiencies in "Semantic Memory" have the greatest impact on 6th grade 3rd reporting category which is "Geometry and Measurement" (0.61) and have the least impact on 8th grade 2nd reporting category which is "Expression, Equation and Function" (0.41). Examples for both categories were presented in figure 3 and figure 4. According to Geary (2003), from the general perspective, visuospatial deficiencies should affect the performance on some subjects such as geometry. However visuospatial deficiencies do not have this kind of expected effects if the domain requires retrieval of facts such as geometric formulas or theorems. Therefore, it is obvious that sample benchmark and corresponding item shown in figure 3 requires retrieving mathematical facts such as value of "Pi" and its common estimates and the formula for calculating the area of a circle. Although benchmark belongs to geometry and measurement reporting area, it doesn't very much require spatially representing numerical and other forms of mathematical information and relationships.

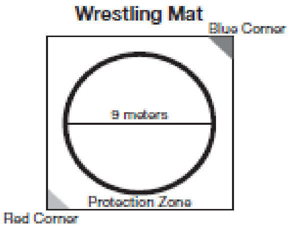
Figure 3. A sample benchmark and corresponding item in 6th grade 3rd reporting category

BENCHMARK MA.6.G.4.1

Reporting Category	Geometry and Measurement
Standard	**Supporting Idea** Geometry and Measurement
Benchmark	MA.6.G.4.1 **Understand the concept of Pi, know common estimates of Pi (3.14; $\frac{22}{7}$) and use these values to estimate and calculate the circumference and the area of circles.**

Sample Item 13 **MC**

In a regulation wrestling match, wrestlers compete within a circular zone 9 meters in diameter.



Which of the following is closest to the area of the circular zone?

A. 28.3 square meters
B. 56.5 square meters
★ **C.** 63.6 square meters
D. 254.3 square meters

Sample benchmark and corresponding item shown in figure 4 have among those which require the least amount of retrieving mathematical facts. Contrary, this item mostly requires sequencing the multiple steps in procedures in solving the given inequality and spatially representing numerical and other forms of mathematical information and relationships by selecting the correct representation of solution set on the number line.

Figure 4. A sample benchmark and corresponding item in 8th grade 2nd reporting category

BENCHMARK MA.8.A.4.2

Reporting Category	Expressions, Equations, and Functions
Standard	**Supporting Idea** Algebra
Benchmark	MA.8.A.4.2 **Solve and graph one- and two-step inequalities in one variable.**

Sample Item 77 **MC**

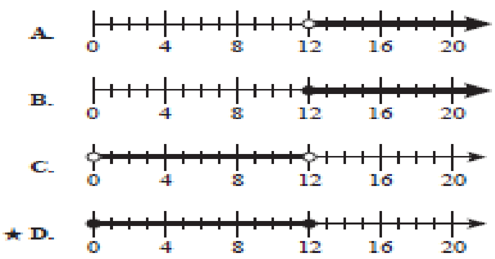
By United States law, any food labeled "reduced fat" must have at least 25% less fat per serving than the regular version of that food. The inequality below can be used to calculate the allowable fat content of a food labeled "reduced fat."

$$x \leq \frac{3}{4}y$$

where:

x = the number of fat grams per serving in the "reduced fat" food
y = the number of fat grams per serving in the regular-version food

One serving of regular crunchy peanut butter has 16 grams of fat. Which number line represents all possible numbers of fat grams that may be in one serving of "reduced fat" crunchy peanut butter while meeting the requirements of U.S. law?



Another important finding of this study is that effect of visuospatial deficiency increases from grade 6 to grade 8 (Figure 5). While the mean weight for Visual-Spatial deficiency for grade six was found to be .17, it is .21 in grade seven and .26 in grade eight. One notable difference is that benchmarks require more spatially representing numerical and other forms of mathematical information and relationships and also require more interpretation of spatially represented information. Detailed investigation on data revealed that visuospatial deficiency has the greatest impact on 7th grade 3rd reporting category which is “Geometry and Measurement” (0.38) and has the least impact on 7th grade 1st reporting category which is “Number: Base Ten” (0.12). Examples for these categories were presented in figure 6 and 7. This finding may be justified by salience of presence of subjects such as transformations, translations and symmetry in 7th grade curriculum. Items generated to assess these subjects require more interpretation and understanding of spatially represented mathematical information and visualization of context as seen in figure 6. Surprisingly, it is interesting that one of the highest weights of visuospatial deficiency among all reporting categories is in the 2nd reporting category of grade 8, which is “Expression, Equation and Function”. Upon further examination, it became evident that benchmarks in this reporting category require spatially representing numerical and other forms of mathematical information and relationships and also require interpretation of spatially represented information (figure 8).

Figure 5. Weights of Visuospatial Deficiency by Grade Level

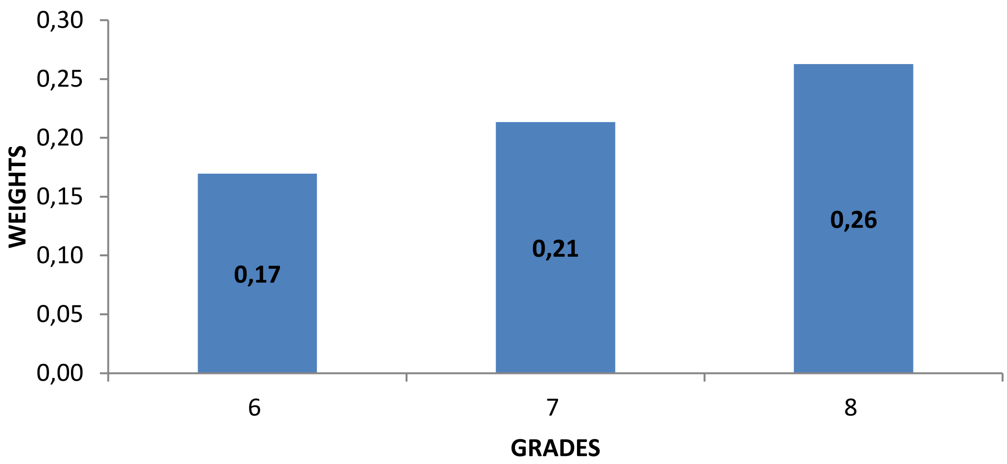


Figure 6. A sample benchmark and corresponding item in 7th grade 3rd reporting category

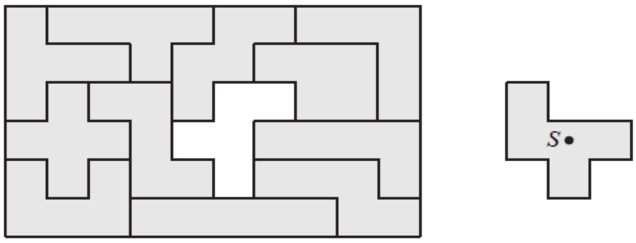
BENCHMARK MA.7.G.4.2

Reporting Category	Geometry and Measurement
Standard	**Supporting Idea** Geometry and Measurement
Benchmark	MA.7.G.4.2 **Predict the results of transformations, and draw transformed figures with and without the coordinate plane.**

Sample Item 46 **MC**

Tracy is playing a puzzle game on the computer. She has placed all the pieces in the puzzle except for one, as shown below.

Tracy's Puzzle



Tracy can complete the puzzle by performing two transformations on the remaining puzzle piece. Which two transformations should Tracy perform?

★ **A.** a 90° clockwise rotation about point *S*, followed by a translation to the left
B. a 90° counterclockwise rotation about point *S*, followed by a translation to the left
C. a reflection across a vertical line, followed by a 90° clockwise rotation about point *S*
D. a reflection across a vertical line, followed by a 90° counterclockwise rotation about point *S*

Figure 7. A sample benchmark and corresponding item in 7th grade 1st reporting category

BENCHMARK MA.7.A.3.3

Reporting Category	Number: Base Ten
Standard	**Big Idea 3** Develop an understanding of operations on all rational numbers and solving linear equations.
Benchmark	MA.7.A.3.3 **Formulate and use different strategies to solve one-step and two-step linear equations, including equations with rational coefficients.**

Sample Item 41 **MC**

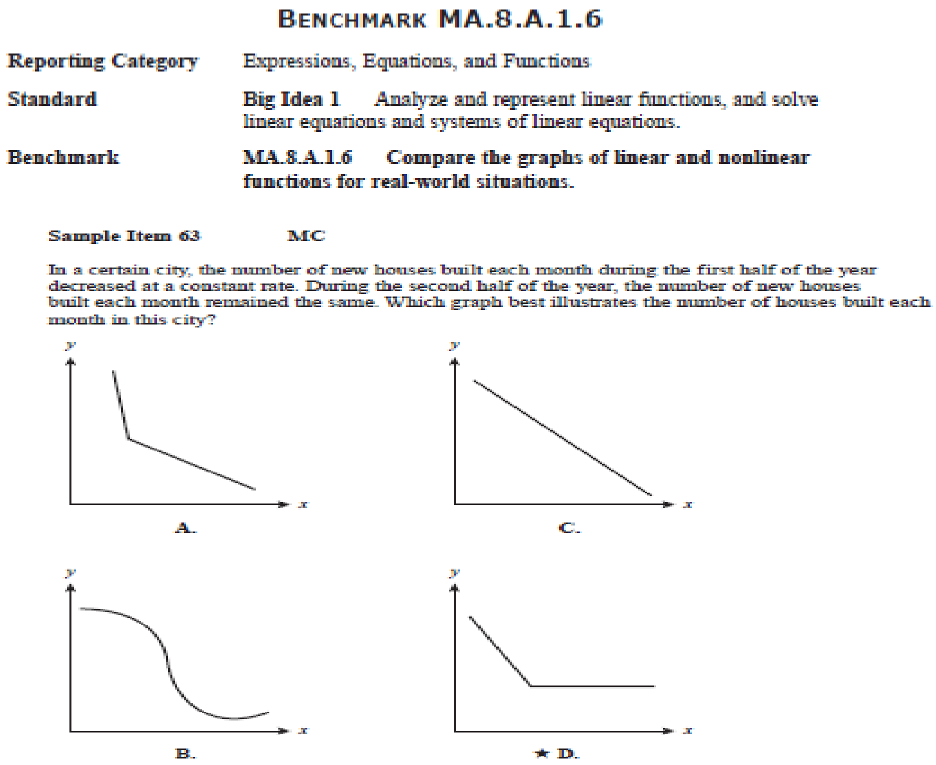
Which of the following steps would solve $\frac{2}{3}x - 4 = 10$?

A. Add 4 to both sides of the equation, then multiply both sides by $\frac{2}{3}$.
★ **B.** Add 4 to both sides of the equation, then multiply both sides by $\frac{3}{2}$.
C. Subtract 4 from both sides of the equation, then multiply both sides by $\frac{2}{3}$.
D. Subtract 4 from both sides of the equation, then multiply both sides by $\frac{3}{2}$.

Sample benchmark and corresponding item shown in figure 7 have among those which require the least amount of visual and spatial skills. Contrary, this item mostly requires sequencing the multiple steps in procedures in solving the given equation by selecting the correct steps and retrieving mathematical facts about operations on rational numbers.

On the other hand as mentioned before visuospatial deficiency has big impact on most of the benchmarks in 8th grade 2nd reporting category "Expressions, Equations, and Functions". As Geary reported this is one of the content areas in which semantic memory or procedural deficiencies are expected to be higher than the visuospatial deficiencies. However, when examined in detail, it is obvious to claim that most of the benchmarks and related items require spatially representing numerical and other forms of mathematical information and relationships and also require more interpretation of spatially represented information. For example, benchmark and the related item shown in figure 8 below requires students to determine appropriate visual representation of verbally provided mathematical information.

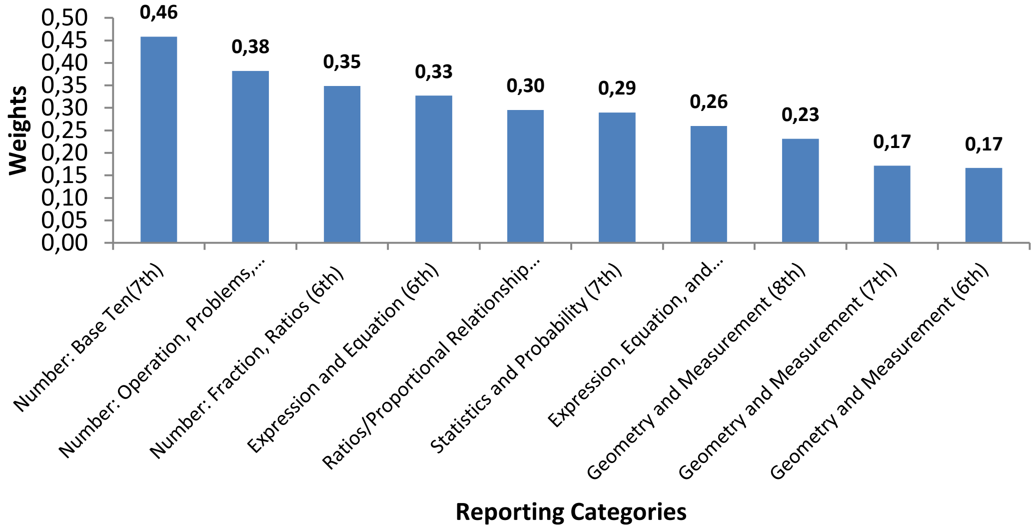
Figure 8. A sample benchmark and corresponding item in 8th grade 2nd reporting category



Comparing to other reporting categories procedural deficiency has its highest effect on "Number" categories in all grade levels. Although it only dominates the 7th grade "Number: Base Ten" category, it shows consistently its highest effects on all number categories as seen in the figure 9 below. As expected, benchmarks and related items in these categories require keeping track of multiple steps and executing complex procedures in problem solving. This finding may be justified by salience of presence of subjects such as operations and problem solving in number category. Items generated to assess these subjects require more execution of simple or complex procedures by sequencing the multiple steps.

Not surprisingly, effect of procedural deficiency does not show big variability among reporting categories, although it has the highest effect on number categories. Since most of the items in FCAT involve in procedural applications in some degree, it is expected that procedural deficiency has almost uniform effect on all reporting categories.

Figure 9. Weights of Procedural Deficiency by Reporting Categories



CONCLUSION

The complexity of mathematical learning disability has made it difficult to identify and explain underlying cognitive features that impact individual performance. Geary (2003) suggests that mathematical learning disability, by its nature, is complex and difficult to identify and explain. This complexity can create the perception that it is difficult to analyze and explain underlying cognitive and performance features of an assessment test that measures student achievement. However, detailed review of literature in this study revealed a model describing which can result in mathematical learning disabilities. This model provided opportunity to analyze and explain underlying cognitive and performance features of a high-stake assessment test (FCAT). In this study, a subjective weighting method was utilized in order to determine criteria (subtypes of deficits) weights by combining individual criteria scores (rankings) supplied by different decision makers based on the certain procedure. A multi-step AHP approach is adopted to determine the relative weights of the criteria (three general subtypes of mathematical disability).

Results of the current study disclosed that semantic memory deficiency is the dominant subtype of mathematical learning disability on vast majority of benchmarks considering the obtained relative weights of all benchmarks. That means most of the items in FCAT assessment require correct retrieval of the mathematical facts which may include recalling definitions, recognizing mathematical entities and objects (e.g., shapes, numbers, expressions, and quantities), retrieving information from graphs, tables, or other sources, reading simple scales and choosing appropriate units of measurement. Another notable finding of this study is that effect of visuospatial deficiency increases from grade 6 to grade 8. In other words, moving from grade 6 to grade 8 benchmarks require more spatially representing numerical and other forms of mathematical information and relationships and also require more interpretation of spatially represented information.

REFERENCES

Ackerman, P.T., & Dykman, R.A. (1995). Reading-disabled students with and without comorbid arithmetic disability. Developmental Neuropsychology, 11, 351–371.

Brown T.C., Peterson G.L. (2009). An enquiry into the method of paired comparison. U.S. Department of Agriculture, General Technical Reports. RMRS-GTR-216WWW, Fort Collins, CO, USA

Chan, A.P.C, Ho, D. C.K., and Tam, C. M. (2001). Design and build project success factors: Multivariate analysis. Journal of construction engineering and management, 127 (2), 93-100.

Chamodrakas, I., Batis, D., & Martakos, D. (2010). Supplier selection in electronic marketplaces using satisficing and fuzzy AHP. Expert Systems with Applications, 37, 490-498.

Choo, E.U., Schoner, B., Wedley, W.C. (1999). Interpretation of criteria weights in multi-criteria decision making. Computers and Industrial Engineering, 37: 527-541.

Forman, E. & Peniwati, K. (1998). Aggregating individual judgments and priorities with the Analytic Hierarchy Process. European Journal of Operational Research, p.108,165−169.

Geary, D.C. (2004). Mathematics and learning disabilities. Journal of Learning Disabilities, 37, 4–15.

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

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

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

Ginevičius, R. (2008). Normalization of quantities of various dimensions, Journal of business economics and management, 9(1): 79-86.

Goldman , S. R., Pellegrino , J. W., & Mertz , D. L. (1988). Extended practice of basic addition facts: Strategy changes in learning disabled students. Cognition and Instruction, 5, 223-265.

Haghighi, M., Divandari, A., & Keimasi, M. (2010). The impact of 3D e-readiness on e-banking development in Iran: A fuzzy AHP analysis. Expert Systems with Applications, 37, 4084-4093.

Hitch, G. J. & McAuley, E. (1991). Working memory in children with specific arithmetical learning difficulties. British Journal of Psychology, 82: 375–386.

Hwang, C.L., & Yoon, K. (1981). Multiple Attribute Decision Making: Methods and Applications. Berlin/Heidelberg/New-York: Springer Verlag.

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

Li, T.S., & Huang, H. (2009). Applying TRIZ and Fuzzy AHP to develop innovative design for automated manufacturing systems. Expert Systems with Applications, 36, 8302-8312.

Pan, N. (2009). Selecting an appropriate excavation construction method based on qualitative assessments. Expert Systems with Applications, 36, 5481-5490.

Ramanathan, R. & Ganesh, L.S.(1994). Group Preference Aggregation Methods employed in AHP: An Evaluation and Intrinsic Process for Deriving Members' Weightages,European Journal of Operational Research, 79, p. 249−265.

Saaty, T.L.(2008). Decision making with the analytic hierarchy process. International Journal of Services Sciences. Volume 1, Number 1/2008 Pages83-98.

Saaty, T.L. (1977). A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15 (3), 234-281.

Saaty. T.L., 1980. The Analytic Hierarchy Process, McGraw-Hill, New York, NY.

Saaty, T.L. (1990). How to make a decision: The analytic hierarchy process. European Journal of Operational Research, 48 1.

Seçme, N. Y., Bayrakdaroglu, A., & Kahraman, C. (2009). Fuzzy performance evaluation in

Turkish Banking Sector using Analytic Hierarchy Process and TOPSIS. Expert Systems with Applications, 36, 11699-11709.

Sen, C. G., & Çinar, G. (2010). Evaluation and pre-allocation of operators with multiple skills: A combined fuzzy AHP and max-min approach. Expert Systems with Applications, 37, 2043-2053.

Triantaphyllou, E. (2000). Multi-Criteria Decision Making Methods: a Comparative Study. Kluwer Academic Publishers, Boston.

Yang, C.L., Chuang, S.P., & Huang, R.H. (2009). Manufacturing evaluation system based on AHP/ANP approach for wafer fabricating industry. Expert Systems with Applications, 36, 11369-11377.