Trends in Life Sciences
An International Peer-reviewed Journal
www.sciencejournal.in

THE APPLICATION OF COGNITIVE DIAGNOSTIC MODELS TO DETERMINE THE NATURE OF DIFFERENCES BETWEEN MALE AND FEMALE STUDENTS IN MATHEMATICS OF THE FIRST HIGH SCHOOL GRADE

Afshin Afzali[1*], Ali Delavar[1], Mohammad Reza Falsafinejad[1], Noor Ali Farrokhi and Ahmad Borjali[2]
[1]Department of Measurement and Assessment, Faculty of Psychology and Educational Sciences, Allameh Tabatabai University, Tehran, Iran.
[2]Department of Psychology, Faculty of Psychology and Educational Sciences, Allameh Tabatabai University, Tehran, Iran.
[*](Corresponding Author)

ABSTRACT
The causes and nature of differences between male and female students in learning mathematics has been interest of many scientists in education, but most of them have investigated mathematics as an overall constituent. In this study, the capabilities of CDM models are used for detailed examination of differences in the underlying math skills. Cognitive diagnostic measurement based on eight main attributes, consisting of 32 items, was run on a sample of 509 students, selected by a multi-stage cluster sampling among students of Tehran. Non-compensatory model of DINA showed differences in parameters of guess, slide, IDI, and RMSEA between the two groups. Male and female students had similar performance in the average dominance of math skills. However, female students were significantly more dominant than male students did in mathematical concepts (at alpha 0.01). In skills of Middle operations and the application of mathematics lessons, male students were significantly more dominant than female students did in real world problems (at alpha 0.05).

KEYWORDS: cognitive diagnostic models, DINA, mathematics of the first high school grade, male and female students

INTRODUCTION

Different structures, abilities and cognitive functions of boys and girls in different sciences, particularly mathematics, have been an interest of educational psychologists and education experts. Many researches have been done in this area, because support of such differences requires use of different teaching methods and models for both genders.

In his review, Zhou (2007) notes many studies on gender differences concerning mathematics. Most of these studies are based on data from standardized math tests such as the SAT-M[1] (Gallagher, 1990; Gallagher and Delisi, 1994; Hyde *et al*, 1990; Royer *et al*, 1999; Willingham, 1997). However, these studies are different in results; for example, some studies have evaluated the boys better than girls in mathematics (Maccoby *et al*, 1974; Halper *et al*, 2000). Some studies related gender differences with types of items, though confirming differences (Voyer *et al*, 1995). Hyde *et al* (1990) concluded that there was a very negligible, if existed, difference between the performance of boys and girls. Kaplan and Kaplan (2005) also reported a very weak relationship between gender and performance in mathematics. Zhou (2007) classified factors affecting the incidence of these differences as different patterns in solving mathematical problems, differences in cognitive skills, differences in psychological characteristics, and different educational experiences.

In addition to the above studies, studies based on the findings of the Trend in International Mathematics and Science Study (TIMMS) (1995, 1999 and 2003) conducted in different countries supported differences between boys and girls in mathematics (Kiamanesh and Pourasghar, 2009). In Iran, several studies examined the differences between boys and girls in mathematics, including Kiamanesh and by Pourasghar (2009). The results showed that gender indirectly influenced math progress through self-concept, while the total effect and direct effects on mathematics were not significant. Based on the findings of Thames, reviews conducted in Iran showed that average performance of girls for 20 open items was lower than that of boys in the 1999 study and higher than that of boys in 2003 for the same 20 items. In contrast, the average performance of girls for 69 multiple-choice items was lower than that of boys in the 1999 study and nearly equal to that of boys in 2003 for the same items (Kiamanesh, 2006). In most studies, performance in

[1] Scholastic Assessment Test - Mathematics

mathematics has been considered as a cognitive constituent and at most, they have reviewed differences based on the type or quality of the items. In recent years, with increased knowledge of and interest in cognitive diagnostic models (CDMs) in cognitive diagnostic fields, in the field of educational psychology and measurements, these models have been used to identify cognitive and functional differences in different groups based on different potentials in dominance on multiple cognitive skills required for successful performance in tests.

In addition to estimate the total ability of an individual in a constituent, CDMs provide subjects with a profile which indicates their dominance on a series of attributes and basic skills predetermined for success in the test. Based on their results, different skills and attributes in different groups can be evaluated in order to achieve reliable detailed diagnostic information. Cognitive diagnosis modeling is a kind of training measure designed to measure knowledge structures and special skill process in students in order to provide information regarding their strengths and weaknesses. Its purpose is to provide valid educational information by which the teacher can effectively design next complementary trainings to respond learning requirements of a given student.

These models have been developed to determine dominance or non-dominance of subjects at multiple attributes based on a matrix of predetermined attributes. Unlike the item-response theory (IRT), which focuses on analysis of item level and the scores assigned to subjects, CDMs focus on informal and diagnostic feedback for subjects and teachers need to know which particular skills are to be improved.

In CDMs, attribute profiles of subjects include classified latent variables (discrete variables). These variables are the heart of scoring process based on which CDM classifies responses in classes of discrete latent variables (for example, a combination of zero and one values in which zero means non-dominance and one means dominance on an individual skill).

Various forms of CDMs have been discussed in literature. Generally, these models cover a variety of situations (e.g., types of constituents, response, and size) which are the interest of researchers in psychometrics, cognitive sciences and learning. Due to extension of these models, currently, researchers are trying to develop relationships between these models. These efforts have led to the growing popularity of these models, as well as a shift in attitudes of researchers for simpler and more feasible CDMs (Yang Sun *et al*, 2011).

DINA[2] model is the most forbearing CDMs which only requires two variable parameters. The model is the basis of other models used in cognitive diagnostic tests. DINA model is a non-compensatory model which assumes that a subject should have all the necessary skills to be able to respond an item correctly (Hanson *et al*, 2009). One subject who is dominant only to a number of attributes needed for an item can have an equal chance with another subject who is not dominant on any of the attributes.

For each item, responses are scored in two latent classes: one class representing the correct respond (score 1) includes subjects with all attributes needed to respond the item correctly; the other class representing the wrong respond to the item (score 0) includes subjects with at least one of the attributes required to respond items correctly (De la Torre, 2011). DINA model complexity is not influenced by the number of attributes by a test because, unlike other cognitive diagnostic non-compensatory methods (such as RUM), the model parameters are estimated for each item, rather than each attribute.

DINA model has two item parameters Slip (s_j) and Guess (g_j). Slip refers to the probability that a subject has all the attributes needed for an item, but fails in responding to that item correctly. Guess refers to the probability of a correct respond to the item in the absence of one or more attributes required for that item. However, these two parameters embrace other disturbing parameters. These disturbing factors alter the reasons why some subjects who do not dominate some required attributes can respond the items correctly as well as the reasons why subjects who dominate all the required attributes miss the correct respond. A general discussion of the DINA model that includes applications and related classification models is available in the works of de la Torre and Douglas (2008), Junker and Sijtsma (2001). According to above, purpose of the present study is to evaluate the difference in CDM of male and female students for

[2] deterministic input noisy and gate model

mathematics of the high school first grade based on DINA model and to compare the performance in cognitive diagnostic assessment and probability of dominance on skills and basic attributes in this filed.

MATERIALS AND METHODS

Methodology

In the study, CDM developed by Afzali (2014) to model mathematics of the first high school grade based on hierarchical attributes was used to examine and compare gender differences in mathematics. This model consists of eight basic attributes including basic math operations, apprehension of definitions, and dominance on concepts of definitions, application of definitions, primary math operations, advanced math operations, numerical calculations, and application of learnings in real-world problems. Based on the proposed model, these attributes have a hierarchical structure, so that dominance on lower-order attributes is required to succeed in a higher-order attribute. For example, a student who does not know basic math operations cannot have an adequate understanding of mathematical definitions, such as slope, trigonometric definitions, equations, and unity. The material designed to evaluate this model is a multiple-choice test consisting of 32 items measuring each attribute of the model by four multiple-choice items.

To evaluate the psychometric properties of the material, the classic theoretical techniques of test and IRT were used. Based on results of these techniques, the internal consistency of the test (Cronbach's alpha) was reported as 0.86. In addition, items were adequately recognizable and difficult to assess different levels of ability (theta) of the subjects. Thus, the lowest difficulty index based on IRT (beta coefficient) was -3.35 and the highest index of difficulty was +3.47. In addition, the lowest recognition index (coefficient alpha) was 0.47 and the highest index was 1.86. This research is based on the results of implementing cognitive diagnostic measurement on a sample of 509 people (including 231 female and 278 male students) sampled by multi-stage cluster sampling among first high-school grade students in Tehran.

To evaluate the fitness of the model, RMSEA index and IDI were used (Robitzsch *et al*, 2011). Guess and slip parameters, fitness indices of data, fitness indices of the model, and the possibility of dominance on eight skills are presented in comparative tables and graphs.

RESULTS

Descriptive findings:

Table 1: Descriptive findings

	Distribution	%	Mean	Variance
Girls	231	45.4	18.85	4.93
Boys	278	56.6	18.43	5.71
Total	509	100	18.62	5.37

Table 1 shows the number, mean and variance of subjects according to gender. Based on findings of the table, 231 out of 509 subjects (45.4%) were female and 278 (56.6%) were male. Mean and variance of female students were 18.85 and 4.93, respectively, and the mean and variance of boys were 18.43 and 5.71, respectively. In addition, the mean and variance of total students were 18.62 and 5.37, respectively.

Analytical Findings

At the following, results of cognitive diagnostic modeling based on the DINA model are presented by experimental data. Responses of 509 male and female students in cognitive diagnostic assessment of basic mathematics are analyzed based on matrix Q containing eight hierarchical attributes using the R software, CDM package. Above table shows the parameters, guess and slip, of the DINA model. Based on the information contained in this table, the lowest guess rate is related to item in girls and item 23 in boys. The highest guess rate is related to item 9 in girls and item 1 in boys. These coefficients indicate the probability of correct respond to the item for students who are not dominant on the skills needed to respond items. Moreover, the lowest slip rate is related to item 2 in girls and item 9 in boys. The highest slip rate is related to item 30 for both girls and boys.

Table 2: parameters of guess and slip for items

Item	Guess		Slip		Item	Guess		Slip	
	Girl	Boy	Girl	Boy		Girl	Boy	Girl	Boy
Item 1	0.805	0.791	0.048	0.092	**Item 18**	0.593	0.446	0.116	0.071
Item 2	0.783	0.744	0.069	0.087	**Item 19**	0.267	0.299	0	0.131
Item 3	0.127	0.299	0.28	0.192	**Item 20**	0.237	0.213	0.321	0.299
Item 4	0.424	0.197	0.166	0.13	**Item 21**	0.333	0.301	0.272	0.175
Item 5	0.489	0.286	0.101	0.164	**Item 22**	0.094	0.155	0.193	0.338
Item 6	0.202	0.304	0.164	0.26	**Item 23**	0	0.005	0	0.125
Item 7	0.37	0.6	0.121	0.205	**Item 24**	0.41	0.262	0.211	0.195
Item 8	0.212	0.28	0.504	0.314	**Item 25**	0.291	0.214	0.297	0.13
Item 9	0.87	0.712	0.03	0.042	**Item 26**	0.017	0.099	0.475	0.382
Item10	0.528	0.444	0.048	0.14	**Item 27**	0.473	0.366	0.036	0.09
Item11	0.431	0.263	0.079	0.054	**Item 28**	0.299	0.334	0.165	0.478
Item12	0.81	0.708	0	0.065	**Item 29**	0.181	0.173	0.475	0.656
Item 13	0.824	0.482	0.044	0.008	**Item 30**	0.127	0.123	0.644	0.739
Item 14	0.651	0.464	0.069	0.032	**Item 31**	0.182	0.246	0.451	0.248
Item 15	0.59	0.64	0.144	0.105	**Item 32**	0.318	0.235	0.064	0.103
Item 16	0.249	0.542	0.063	0.063	**Mean**	0.399	0.365	0.180	0.1912
Item 17	0.494	0.454	0.111	0.008					

Table 3: IDI and RMSEA of items

Item	IDI		RMSEA		Item	IDI		RMSEA	
	Girl	boy	Girl	boy		Girl	boy	Girl	boy
Item 1	0.147	0.117	0.187	0.17	**Item 18**	0.291	0.483	0.218	0.178
Item 2	0.148	0.169	0.167	0.16	**Item 19**	0.733	0.571	0.102	0.13
Item 3	0.593	0.509	0.178	0.156	**Item 20**	0.443	0.488	0.273	0.248
Item 4	0.41	0.673	0.178	0.096	**Item 21**	0.395	0.524	0.315	0.19
Item 5	0.41	0.551	0.197	0.147	**Item 22**	0.713	0.507	0.103	0.203
Item 6	0.634	0.436	0.163	0.148	**Item 23**	1	0.87	0	0.065
Item 7	0.509	0.194	0.161	0.18	**Item 24**	0.379	0.543	0.313	0.22
Item 8	0.284	0.406	0.277	0.202	**Item 25**	0.412	0.656	0.234	0.151
Item 9	0.10	0.246	0.181	0.162	**Item 26**	0.508	0.519	0.234	0.183
Item 10	0.424	0.416	0.177	0.176	**Item 27**	0.491	0.544	0.172	0.153
Item 11	0.491	0.682	0.158	0.124	**Item 28**	0.536	0.187	0.197	0.2
Item 12	0.19	0.228	0.129	0.177	**Item 29**	0.344	0.171	0.23	0.152
Item 13	0.131	0.509	0.12	0.149	**Item 30**	0.23	0.138	0.193	0.148
Item 14	0.28	0.504	0.189	0.135	**Item 31**	0.367	0.506	0.227	0.134
Item 15	0.266	0.256	0.245	0.185	**Item 32**	0.618	0.662	0.175	0.088
Item 16	0.688	0.395	0.095	0.155	**Mean**	0.420625	0.4436875	0.1847	0.16022
Item 17	0.395	0.538	0.221	0.162					

This coefficient indicates the probability of wrong respond to the item for students who are dominant on the skills needed to respond items. Lower guess and slip parameters indicate a better fitness between diagnostic assessment designs, experimental data of the cognitive data (Ravand *et al*, 2013). The results listed in the table can reveal the relative similarity of the parameters for both genders. Moreover, female and male subjects are relatively in similar conditions for average guess and slip parameters. Average guess parameter is slightly higher in girls than in boys (0.026) and slip parameter is slightly higher in boys than in girls (0.011). Table 3 presents Item detection indices (IDI) and RMSEA. IDI is inversely related to guess and slip parameters. RMSEA index shows the fitness of the item to cognitive model. This index is based on chi-square distribution; the smaller RMSEA, the higher fitness. According to results contained in the table, item 23 has the best IDI for both genders and the lowest RMSEA index and thus the highest fitness to model for both genders. Item 8 for girls and 20 for boys have the highest value of RMSEA and the least fitness to the model.

Table 4: probability of dominance on eight attributes in the case sample

Trait	Title	Probability		z	Level
		Girl	Boy		
Attitude 1	Basic math operations	0.5596	0.6211	-1.40622	0.08
Attitude 2	Apprehension of definitions	0.5489	0.4087	3.154378	0.00
Attitude 3	Memorization of definition	0.6133	0.5579	1.262069	0.10
Attitude 4	Application of definitions	0.6131	0.6212	-0.18721	0.43
Attitude 5	Primary math operations	0.4923	0.5733	-1.82451	0.03
Attitude 6	Advanced math operations	0.4935	0.5314	-0.85175	0.20
Attitude 7	Numerical calculations	0.5976	0.557	0.922622	0.18
Attitude 8	Application in real world problems	0.3535	0.4381	-1.93994	0.03
	Average probability of dominance	0.533975	0.5385875	-0.1039	0.46

Table 4 presents marginal probability of dominance on eight skills in boys and girls. According to results listed in the table, female subjects showed the highest probability of dominance on skills of dominance on definitions, application of definitions and numerical calculations, while male subjects showed the highest probability of dominance on application of definitions, basic math operations and dominance on definitions, respectively. Moreover, female subjects showed the lowest probability of dominance on application of learnings in real world problems, while male subjects showed the lowest probability of dominance on apprehension of definitions (0.464), advanced math operations (0.498) and application of learnings in real world problems (0.505). Probability of dominance and percentage of dominant people on eight skills for male and female students are shown in Figure 1 and Figure 2.

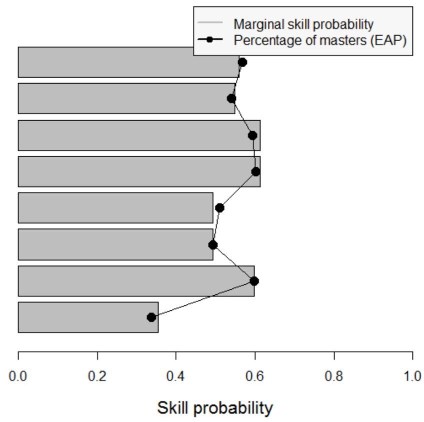

Figure 1: probability and percentage of dominance on skills in female subjects

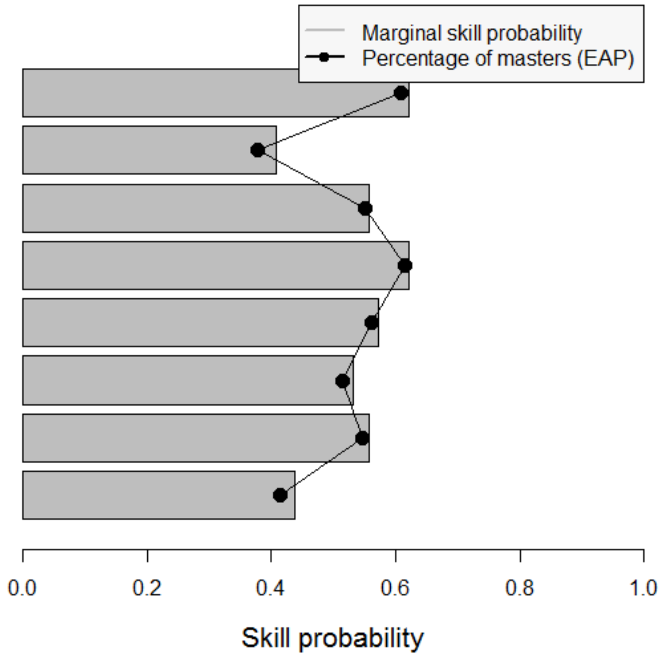

Figure 2: probability and percentage of dominance on skills in male subjects

DISCUSSION AND CONCLUSIONS

In this research, capabilities of CDMs was used to investigate gender differences in mathematics of first high-school grade based on empirical data obtained from 509 samples of high school students in Tehran (231 female students and 278 male students). The main advantage of these models is that, they are not only capable of investigating and comparing competency in mathematics as a total constituent, they also provide adequate detection information on details of differences in traits and basic skills forming this constituent.

In order to compare differences, a model, consisting of eight hierarchical attitudes, which was developed by Afzali (2014) for mathematics of first high-school grade, was used. To evaluate and compare the model fitness by experimental data as well as estimation and comparison of success probability in attributes and skills, the non-compensatory model DINA was applied.

According to data obtained from this model, no significant difference was found in guess and slip parameters of both genders. Besides, the mean parameters were almost equal for both groups. Performance of students in these two parameters was not considerably different for majority of items. That means there was no significant difference in probability of error for female and male students when they had skills required for responding an item (slip) as well as the probability of correct respond to an item without dominance on skills required for responding that item (guess). Based on results listed in the table, two average parameters of guess and slip were 0.398 and 0.18 for girls, 0.365 and 0.19 for boys.

In the case of fitness parameters including IDI and RMSEA, the general conditions are relatively similar in both genders. Items which were properly able to distinguish dominant and non-dominant girls were able to distinguish boys. There was no significant difference between the genders in different items. The average IDI and RMSEA as criteria for model fitness were similar in both groups; though IDI was slightly more for boys than for girls (0.44 versus 0.42) and RMSEA was slightly more for girls than boys (0.184 versus 0.16).

Following differences in model fitness and cognitive diagnostic measurement for both genders, the performance of male and female students was examined for mathematics of first high-school grade as a total constituent as well as attitudes and main skills required to this course. According to data listed in Table 5, there are no significant differences between boys and girls in terms of overall success probability in cognitive diagnostic assessment of mathematics. The average probability of dominance was 0.534 for female subjects and 0.538 for male subjects. These results are consistent with results of studies in which no relationship was found between gender and math achievement. However, significant differences were found in both groups when evaluating attitudes and basic skills. The greatest difference in success was related to apprehension of definitions; female students were significantly (at alpha 0.01) more dominant than boys were. In contrast, boys showed significant (0.05) dominance on skills of primary math operations, application of learnings in real world problems. These differences may indicate a tangible difference in the cognitive structure and mathematical problem solving patterns in the two groups. Moreover, major weakness of subjects, particularly girls, in application of math learnings is in problems in which the student is expected to solve a problem related to real world, as the Table 5 shows. The recent finding is consistent with Minaei (2013).

According to results obtained in this study, it can be concluded that consideration of cognitive and functional differences of boys and girls in educational planning for mathematics can effectively influence efficiency and effectiveness of these plans. It seems also necessary to train practical abilities alongside the theoretical concepts of mathematics, considering various application of math in other sciences and the real life. In conclusion, researchers are recommended to use capabilities of CDMs in their studies regarding interdisciplinary nature and particular capabilities of these models in different sciences and contexts.

REFERENCES

Afzali A. (2014). Cognitive Diagnostic Modelling for Mathematics of the First High-school Grade Based on Hirarchical Attitudes, Ph.D. Thesis, Allameh Tabatabai University, Tehran, Iran.

Caplan J. B. and Caplan P. J. (2005). The perseverative search for sex differences in mathematics abilities.In A. M. Gallagher and J. C. Kaufman (eds.) Gender Differences in Mathematics: An Integrative Psychological Approach. Cambridge: Cambridge University Press

De la Torre J. and Douglas J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika.* 69(3): 333-353.

Gallagher A. M. (1990). Sex differences in the performance of high-scoring examinees on the SAT-M. NewYork: College Entrance Examination Board.

Gallagher A. M. and DeLisi R. (1994). Gender differences in Scholastic Aptitude Test – mathematics problem solving among high-ability students. *J. Educational Psychol.* 86: 204-211.

Halpern D. F. (2000). Sex Differences in Cognitive Abilities (3rd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Henson R., Templin J. and Willse J. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika.* 74(2): 191-210.

Hyde J. S., Fennema E. and Lamon S. J. (1990). Gender Differences in Mathematics Performance: A Meta-Analysis. *Psychol. Bull.* 107(2): 139-155.

Junker B. W. and Sijtsma K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychol. Measurement.* 25:258–272.

Kabiri M. (2003). Role of Math Self-efficacy along with Personal Variables of Male and Female Students, Master Thesis, Kharzmi University, Tehran, Iran.

Kiamanesh A. and Pourasghar N., 2009. the Difference between Girls and Boys in Variables Related to Math Performance and Its Role on Math Progres. Psychology, 4(13), 12-22.

Maccoby E. E. and Jacklin C. N. (1974). The Psychology of Sex Differences. Stanford, CA: Standford University Press.

Minaei A., (2012). Cognitive Diagnostic Modelling for Tames Items 2007 in Eight Grade Students, Iran, Ph.D. Thesis, Allameh Tabatabai University, Tehran, Iran.

Mislevy R. J., Steinberg L. and Almond R. (2003). On the structure of educational assessments. Measurement: Interdisciplinary Research and Perspective. Hillsdale, NJ: Lawrence Erlbaum Associates.Nichols, P. D. (1994).

Ravand H. Barati H. and Widhiarso W. (2013). Exploring Diagnostic Capacity of a High Stakes Reading Comprehension Test: A Pedagogical Demonstration. *Iranian J. Language Testing.* 3(1).

Robitzsch A., Kiefer T. George A. and Uenlue A. (2013). Package 'CDM'. https://sites.google.com/site/alexanderrobitzsch/software
Royer J. M., Tronsky L. N., Chan Y., Jackson S. J. and Marchant H. I. (1999). Math-Fact retrieval as the cognitive mechanism underlying gender differences in math test performance. *Contemp. Edu. Psychol.* 24: 181-266.
Verma S., (2001). Curriculum and standards framework. Tehran: Metacognition of Thought publication, Tehran, Iran.
Voyer D., Voyer S. and Bryden M. P. (1995). Magnitude of sex differences in spatial abilities: A metaanalysis and consideration of critical variables. *Psychol. Bull.* 117: 250–270.
Willingham W. W. and Cole N. S. (1997). Gender and Fair Assessment. Mahwah, NJ: Erlbaum.
Young-sun L. and Yoon Soo P. (2011). A cognitive diagnosthc modeling of attribute mastery in Massachusetts, Minnesota, and the U.S. National sample using the TIMSS 2007. *Int. J. Testing.* 11:144-177.
Zhu Zheng (2007). Gender differences in mathematical problem solving patterns: A review of literature. *Int. Edu. J.* 8(2): 187-203.

