International Journal of New Technology and Research (IJNTR)
ISSN:2454-4116, Volume-3, Issue-9, September 2017 Pages 87-102

A Multilevel Analysis to Analyse The Timss Data: A Comparison of The Lebanese And Singapore

Tarek Harmouch, Zaher Khraibani, Talal Atrissi

***Abstract*— The objective of this article is to model the variation in mathematics achievement of Lebanese and Singaporean eighth-graders as a function of student- and school-level factors by using the Trends in International Mathematics and Science Study (TIMSS) data.This article consists of abriefdescription of the educational system background of Lebanon and Singapore; this is followed by a review of research on school effectiveness and multilevel modelling, the theoretical framework of the study, and discussion of the factors explored in the study.**

***Index Terms*—TIMSS, Multilevel Analysis, Literature review, Unconditional model, Hierarchical Linear Modeling (HLM)**

I. INTRODUCTION

The main goal of a formal educational system is to realize a set of predetermined learning objectives [1]. Assessing the extent to which students achieve these objectives needs the implementation of assessments. Hence, assessment is an indispensable part of the teaching and learning process. International assessments, among the other types, such as the school-, district- and national-level assessment, became a common method of assessing the quality of education in the 1990s and 2000s [2]. This reflects the fact that the emphasis has shifted from academic input indicators (e.g., student participation rates, physical resources, teachers training), to the quality of educational outcomes, such as skills, knowledge, and attitudes [2]. Therefore, it is essential for educational policymakers to have comprehensive and timely information about the educational system performance to monitor the school activities and students' learning. Thus, it is proposed that educational evaluation should be based on output indicators, in which the entire educational system needs to be assessed rather than on input indicators, such as the identification of individual students learning difficulties, selection for further training and certification [2].

International assessments provide remarkable opportunities to ministries of education and educational practitioners to assess the quality of teaching and learning of key subjects, such as mathematics at the national level. The Trends in International Mathematics and Science Study (TIMSS) is the most ambitious comparative assessment of students' achievement in mathematics and science initiated by the International Association for Evaluation of Educational Achievement (IEA)[4]. The IEA offers extraordinary opportunities for researchers to investigate a wide variety of questions on mathematics teaching, and the learning process based on the TIMSS assessment concerning the variation in students' mathematics achievement within and between schools, whether mathematics education is driven largely by the student-level characteristics or those at the school level; and the level of mathematics achievement of different students groups, such as students from public and private schools, boys and girls, ethnic groups, students from rural and urban schools. The present study aims to model variation in mathematics achievement scores of Lebanese and Singaporean eighth-graders within and between schools based on TIMSS 2011 data using the multilevel linear modelling methodology.Lebanon is a developing country whilst Singapore is a developed country. There are similarities in the educational background indicators of these two countries. For example, the educational system is centralized, schooling age, and multilingual society is another common feature of both countries. As is documented at the beginning of the next section, however, there are differences between these two countries with regard to the education background indicators. Singapore invests a greater proportion of the total gross domestic product on public education than Lebanon, the national income per capita is higher in Singapore compared to Lebanon [10].

The main purposes of TIMSS are to assess students mathematics and science achievements; define the achievement in terms of concepts, processes, skills, and attitudes; describe the context in which the learning develops; provide international benchmarks that can assist policymakers to uncover the strengths and weaknesses of their educational systems compared to the other educational systems; and collect high-quality data that can increase policymakers' awareness and importance of students and school outcomes affecting factors [5].The literature review reveals that the TIMSS findings have been used in a wide variety of ways in different countries. For example, changes in revision of curricula is accelerated (Singapore and Czech Republic); major changes have been made in the area of school teaching, class organisation, teacher education, and target-setting for schools (Scotland); standards development, curriculum document development, mathematics and science teaching methodologies, and teacher studies affected by the TIMSS results (Slovak Republic); a centralized examination system was founded (Latvia); the results showed that there is gender gap and negative attitude towards science and mathematics, these two issues were used as a base for curriculum reform and teachers' professional development (Korea. Rep. of); new topics and contents were added to the mathematics and science curriculum (Romania and Spain); and national benchmarks were established in literacy and numeracy (Australia) [3]. This indicates the importance and role of

Tarek Harmouch, Doctorat institute of letters humanities and social sciences, Lebanese University, Beirut, Lebanon
Zaher khraibani, Applied Mathematics department, Lebanese University, Faculty of Sciences, Hadat,Lebanon
Talal Atrissi,, Institute of letters humanities and social sciences, Lebanese University, Beirut, Lebanon

the TIMSS findings in educational policy to improve the quality of school outcomes.As mentioned earlier, despite the similarities between Lebanon and Singapore in some of the educational background indicators, Singaporean students scored considerably higher in TIMSS assessments than their Lebanon peers. Lebanon eighth-graders with a mean score of 433 were ranked 31 in mathematics among 45 countries in 2003; also, with mean scores of 441 and 449 they were placed at 28 and 25 among 48 and 63 countries in 2007 and 2011, respectively. In contrast, Singaporean eighth-graders with a mean score of 605 ranked at 1st in 2003, and with mean scores of 593 and 611 they were ranked at 3rd and 2nd in 2007 and 2011, in that order [8]. The difference in overall mathematics achievement is a score of 162 on average in TIMSS 2003 and 2007, and 2011.

II. Educational System Background of Lebanon and Singapore

In order to compare students' achievement between Lebanon and Singapore, it is informative to contrast a few basic geographical, socio-economic and educational indicators that underlie the educational system background. Some of the important indicators of the two countries are presented in Table 1.

Table 1 Socio-economic indicators of Lebanon and Singapore

Indicators	Lebanon	Singapore	Difference
Population size (in millions)	4	5	-1
Area of country (square kilometres)	10,000	1,000	10
Population density (people per square kilometre)	313	7,125	6,812
Urban population (% of total)	87%	100%	-13%
Life expectancy at birth (years)	72	81	-9
Infant mortality rate (per 1,000 live births)	11	2	9
Gross National Income per capita (in US Dolores)	8,060	37,220	-29,160
GNI per capita (purchasing power parity)	13,400	49,780	-36,380
Public expenditure on education (% of GDP)	2	3	
Net enrolment ratio in primary education (% of relevant group)	90	100	-10
Net enrolment ratio in secondary education (% of relevant group)	75	98	-23
Primary pupil-teacher ratio	14	19	-5

Source:(Mullis et al, 2012, Vol, 1, p. 6)

Lebanon is a developing country. It is part of the Middle East and located along the Eastern of the Mediterranean Sea. In contrast, Singapore is a developed country located in South-West Asia. In terms of area, Lebanon is 10 times bigger than Singapore, and its population is more than Singapore by one million, however, the population density in Singapore is more dense (almost 23 times) than Lebanon. Singapore enjoys a higher (closely five times) per capita income, higher purchasing power parity (closely four times), far lower infant mortality rate, much longer life expectancy.

In Singapore, the net enrolment ratio at both primary and secondary school is considerably higher than Lebanon. Based on the above indices, it seems that the implementation of the educational policies and planning in Lebanon is more difficult than in Singapore.

TIMSS is an international assessment of mathematics and science that has been conducted every four years since 1995. It assesses the mathematics and science achievement of students in their fourth and eighth years of formal schooling. At the fourth grade, the target grade is the grade that represents four years of schooling, counting from the first year based on UNESCO's International Standard Classification of Education (ISCED).

At the eighth grade, the target grade is the grade that represents eight years of schooling, counting from the first year of schooling [9].Participating countries may choose to assess one or both grades.Lebanon has participated in TIMSS since 2003 only at eighth grade, whereas Singapore joined TIMSS in 1995 for both fourth and eighth grades.

The present study was designed to compare the mathematics achievement of Lebanese and Singaporean students; hence the sample is limited to grade eight.To ensure that the data provided a representative sample of

the national students' population, TIMSS used a constant two-stage stratified cluster sampling design. In all countries schools are sampled with probability proportional to size (PPS) at the first stage, whereby larger schools are chosen with higher probability, and then one intact classroom was selected using a systematic random method. However, Singapore also had a third sampling stage, where students were sampled within classrooms.

The properties of the samples are presented in the following table.

Table 2. : The properties of the samples of the study

Country	Students			Classrooms					Schools			
	Total	Girls	Boys	No. Classes	Min class size	Max class size	Mean class size	SD* class size	No. Schools	Min school size	Max school size	Mean school size
Lebanon	3974	2128	1846	187	4	39	21	8.6	147	66	3702	742
Singapore	5927	2993	2934	330	11	19	18	1.3	165	408	4369	1364

* Standard deviation

Table 2 shows that Lebanon students were nested within 147 schools and the Singaporean were nested within 165 schools. Thus, the total number of students from each country was the student-level sample size. Of the total Lebanon students 53.5% were girls and 46.5% boys. In Singapore 49.5% were girls and 50.5% boys students. The mean age of the students was 14.3 and 14.4 in Lebanon and Singapore, respectively. The data were obtained from TIMSS 2011 online database (http://rms.iea-dpc.org/).

III. Population, Sampling Method and Sample Size

TIMSS is an international assessment of mathematics and science that has been conducted every four years since 1995. It assesses the mathematics and science achievement of students in their fourth and eighth years of formal schooling. At the fourth grade, the target grade is the grade that represents four years of schooling, counting from the first year based on UNESCO's International Standard Classification of Education (ISCED). At the eighth grade, the target grade is the grade that represents eight years of schooling, counting from the first year of schooling [7]. Participating countries may choose to assess one or both grades.Lebanon has participated in TIMSS since 2003 only at eighth grade, whereas Singapore joined TIMSS in 1995 for both fourth and eighth grades. The present study was designed to compare the mathematics achievement of Lebanese and Singaporean students; hence the sample is limited to grade eight.To ensure that the data provided a representative sample of the national students' population, TIMSS used a constant two-stage stratified cluster sampling design. In all countries schools are sampled with probability proportional to size (PPS) at the first stage, whereby larger schools are chosen with higher probability, and then one intact classroom was selected using a systematic random method. However, Singapore also had a third sampling stage, where students were sampled within classrooms. The properties of the samples are presented in Table 3.

Table 3. : The properties of the samples of the study

Country	Students			Classrooms					Schools			
	Total	Girls	Boys	No. Classes	Min class size	Max class size	Mean class size	SD* class size	No. Schools	Min school size	Max school size	Mean school size
Lebanon	3974	2128	1846	187	4	39	21	8.6	147	66	3702	742
Singapore	5927	2993	2934	330	11	19	18	1.3	165	408	4369	1364

*Standard deviation

Table 3 shows that Lebanon students were nested within 147 schools and the Singaporean were nested within 165 schools. Thus, the total number of students from each country was the student-level sample size. Of the total Lebanon students 53.5% were girls and 46.5% boys. In Singapore 49.5% were girls and 50.5% boys students. The mean age of the students was 14.3 and 14.4 in Lebanon and Singapore, respectively. The data were obtained from TIMSS 2011 online database (http://rms.iea-dpc.org/).

In Lebanon, one intact classroom was sampled per school. However, from 40 schools with 110 students and above, more than one classrooms were sampled. Thus, the school-level sample size was 147 schools (187 classrooms). In Singapore, on average, 18 students were selected at random from two classrooms within each school, consequently, the school-level sample size was 165 schools and the classroom-level sample size was 330 classrooms. Out of the total schools in Lebanon, 48 (32.7%), 94 (63.9%), and 5 (3.4%) were high performance public schools, high performance private schools, and low performance schools, respectively. Table 4 depicts the locations of the schools in Lebanon.

Table 4: The distribution of the sampled schools across Lebanon

Code	Location	No of schools	%
1	Beirut	13	8.8%
2	Bekaa	20	13.6%
3	Mont Liban	16	10.9%
4	Mont Liban (Banlieue)	35	23.8%
5	Nabayieh	11	7.5%
6	Nord	33	22.4%
7	Sud	19	12.9%
Total		147	100%

In Singapore, out the total schools, 162 (98.2%) were lower secondary schools and three (1.8%) were comprehensive school. Other than this, schools were not stratified because Singapore is a city state country. Census of all school participated in the study and there was not sampling at school level.

According to the definition of the target population in TIMSS 2011, all students who enrolled in the eighth-grade, regardless of their age, school type and location, potentially belong to the international desired target populations [11]. Participating countries were expected to include all the target grades in their definition of the population. "However, in some cases, political, organizational, or operational factors make complete national coverage difficult to attain" [11]. Consequently, in some rare situations, certain groups of schools and students may have to be excluded from the national target population. There were two levels of exclusion: school- and within-school exclusion. School-level exclusion refers to the exclusion of some schools from the national desirePopulation due to reasons, such as schools with very few students, special education schools and private schools. In Lebanon, school-level exclusion comprised schools with less than nine students are known to be very small schools, but there was no school with less than nine students, hence, there was no school level exclusion. In Singapore, school-level exclusions consisted of special education schools and private schools. In Singapore, there was also no school-level exclusion. The second type of exclusion was within-school exclusions. There was no within-school exclusion in either Lebanon or Singapore. Therefore, it can be concluded that the final sample of students was nationally representative of eighth-grade students in both country.

A. Sample Size Considerations

Sample size is one of the important issues in multilevel linear modelling. This is because there is more than one sample size. In addition, the maximum likelihood estimation is the common method for estimating parameters in multilevel

linear modelling. This method requires a large sample size to estimate the parameters precisely [12]. In the present study, the data are a two-level hierarchy structured and there were two sample sizes: student-level sample size (3974 and 5927 students in Lebanon and Singapore) and school-level sample size (147 and 165 schools in Lebanon and Singapore). The estimated parameters are influenced by both individual- and group-levels sample sizes, however, the group-level sample size is much important than the individual-level sample size [12]. According to [13], having 30 groups (classrooms or schools) with at least 30 individuals is a sound advice if the researcher's interest is mostly in fixed parameters. However, researchers can modify this rule for particular applications. For example, if interest is in cross-level interaction, 50 groups with 20 individuals per group, and if there is strong interest in the random part, the variance and covariance components and their standard errors, the number of groups should be 100 with 10 individuals per group. Accordingly, the sample sizes of this study at the student and school levels; efficiently met the criterion has been suggested in the literature.

B. Sampling Weights

TIMSS used multistage stratified cluster sampling design. In multistage cluster sampling, individuals are nested within clusters within strata. Consequently, the chance of each unit in the population included in the sample is not

equal. When the multistage cluster sampling design is used, sampling weight must be taken into account to avoid bias in the estimated parameters [14]. The TIMSS centre has computed several weighting variables that should be considered by the researcher[15], [16]. These weighting variables are as follows:

1. TOTWGT is the sum of the sampling weights of all students in a country. It must be used when students' population parameters estimates are required. The advantage of using TOTWGT is to ensure whether several subgroups that comprise the sample are properly and proportionally represented in the computation of population estimates. However, in cross-country studies with unequal sample sizes, TOTWGT is not desirable, because it treats the unequal sample sizes equally.
2. SENWGT is the sum of the student sample weight in a sample size of 500 in each country regardless of the size of the students' population in each country.
3. HOUWGT is the sum of the student sample size in each country. To avoid the problems mentioned above, TIMSS provides HOUWGT, which is a transformation of TOTWGT that ensures the weighted sample corresponds to the actual sample size in each country.
4. To prevent the problem with using TOTWGT, TIMSS conveyed it into two different sampling weights of SENWGT and HOUWGT that can be used in cross-country analysis. Since, the present study involved two countries; the SENWGT was used at the student level.
5. In addition to the sampling weights at the student level, TIMSS also computed several sampling weights at the classroom and school level. MATWGT, mathematics teacher weight is an important sampling weight at the classroom level. It should be mentioned that many other sampling weights are calculated at the school level [15]. In this study the SCHWGT was used as the school-level weighting variable.

IV. MULTILEVEL LINEAR MODELLING

This section consists of three part. The results of the exploratory data analysis that was used to check the tenability of the assumptions underlying the multilevel linear modelling are presented, followed by the results of the main analysis for Lebanon and Singapore (overall sample) data. The data were obtained from 3,974 students nested within 147 schools from Lebanon (overall sample) and 5,927 students nested within 165 schools from Singapore who were involved in the 2011 TIMSS assessment were analysed using multilevel linear modelling technique. The data for public and private schools were obtained from 1,043 students nested within 48 public schools and 2,835 nested within 94 private schools. The HLM descriptive statistics for all the factors are provided in the following table.

The HLM descriptive statistics for all the factors

Table 5 Descriptive statistics for **student-level factors** (**Lebanon overall sample**)

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
Mathematics Achievement	3974	457.68	71.57	266.66	694.81
Attitudinal Factors					
Math Self-Concept	3974	10.58	1.94	3.18	15.82
Attitude Towards Math	3974	10.58	1.90	5.04	13.91
Math valuing	3974	10.44	2.06	3.41	13.71
Educational Aspiration	3974	2.10	1.01	0.00	3.00
Family Background Factors and gender					
Home educational Resources	3974	9.50	1.87	4.32	14.02
Home Language	3974	1.96	0.60	1.00	3.00
Students gender	3974	-	-	0.00	1.00

Table 6 Descriptive statistics for **school-level factors** (**Lebanon overall sample**)

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
Mathematics Achievement	147	448.56	50.85	347.91	589.24
School emphasis on academic success_ Teacher repost	147	10.31	2.00	4.99	14.58
School emphasis on academic success_ Principal repost	147	9.78	1.89	4.91	15.57
Teacher career satisfaction	147	10.05	1.78	6.17	13.80

Teacher confidence in teaching math	l47	10.53	1.49	5.07	11.99
Problem with working conditions	l47	10.91	1.96	3.10	14.89
Safe and orderly school	l47	10.06	1.89	4.58	13.22
Instruction affected by math resource shortage	l47	9.78	2.11	3.09	15.23
School discipline and safety	l47	10.33	2.13	3.98	13.94
Teachers gender	l47	-	-	0.00	1.00
Teachers years of experience	l47	0.53	0.52	0.00	1.62
Class size	l47	26	8	5	45
School size	l47	736	676	66	3702
Mean students math self-concept	l47	11.11	2.22	6.39	15.82
Mean students attitude towards math	l47	10.64	1.90	5.04	13.47
Mean students math valuing	l47	10.41	2.12	5.49	13.71
Mean students educational aspiration	l47	5.70	1.08	2.00	7.00
Mean students home educational resources	l47	10.17	2.01	4.32	14.02
Mean students home language	l47	2.16	0.91	1.00	4.00

Table 7 Descriptive statistics for **student-level factorsSingapore**

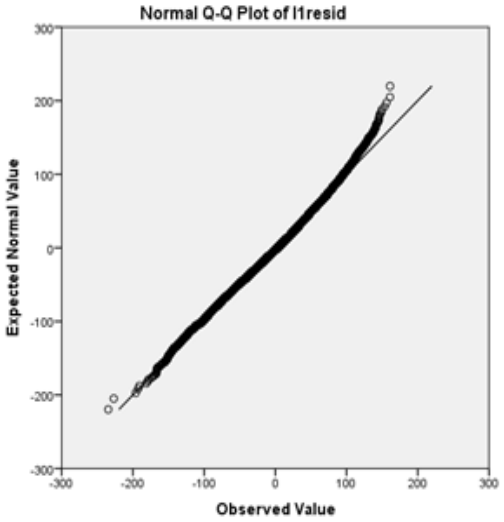
VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
Mathematics Achievement	5927	607.54	81.58	327.74	799.61
Attitudinal Factors					
Math Self-Concept	5927	9.95	2.06	3.18	15.82
Attitude Towards Math	5927	10.36	1.97	5.04	13.47
Math valuing	5927	9.96	1.79	3.41	13.71
Educational Aspiration	5927	2.43	1.00	0.00	3.00
Family Background Factors and gender					
Home educational Resources	5927	10.26	1.67	4.32	14.02
Home Language	5927	1.49	0.60	1.00	3.00
Students gender	5927	-	-	0.00	1.00

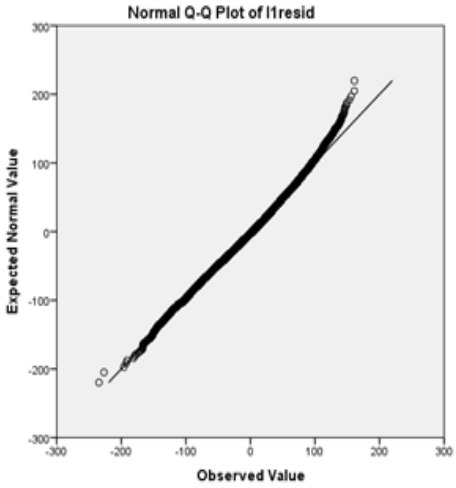
Table 8 Descriptive statistics for **school-level factors Singapore**

VARIABLE NAME	N	MEAN	SD	MINIMUM	MAXIMUM
Mathematics Achievement	165	607.04	56.13	438.85	736.86
School emphasis on academic success_ Teacher repost	165	10.46	1.77	4.99	16.21
School emphasis on academic success_ Principal repost	165	10.69	2.05	4.91	15.57
Teacher career satisfaction	165	9.41	1.90	1.82	13.80
Teacher confidence in teaching math	165	9.32	2.04	5.07	11.99
Problem with working conditions	165	10.95	1.90	7.01	14.89
Safe and orderly school	165	11.04	2.03	4.58	13.22
Instruction affected by math resource shortage	165	11.67	2.93	3.09	15.23
School discipline and safety	165	10.89	1.40	8.40	13.94
Teachers gender	165	-	-	0.00	1.00
Teachers years of experience	165	3.08	0.97	1.00	4.00
Class size	165	38	5	19	45
School size	165	1360	496	408	4369
Mean students math self-concept	165	9.95	0.58	8.57	11.86
Mean students attitude towards math	165	10.37	0.47	8.93	11.75
Mean students math valuing	165	9.96	0.40	8.91	10.98
Mean students educational aspiration	165	2.43	0.25	1.77	2.92
Mean students home educational resources	165	10.26	0.73	8.61	12.27
Mean students home language	165	1.49	0.21	1.03	1.94

A.Normality

Multilevel linear models assume that the level-1 random errors (r_{ij}) are normally distributed [19]. The HLM computer package, after fitting the final model, provides a means of checking the distributional assumptions of the model by producing residual files. The level-1 residual file contains the differences between the observed and the fitted values for the outcome variable. The level-1 residual file was used to produce a probability plot (Q-Q) to check the level-1 normality distribution of random errors. If the Q-Q plot resembles a 45-degree line, it shows that the random errors are normally distributed [19]. The following figures contains respectively two Q-Q plots of Math for Lebanon (overall sample) and Singapore that approximately are linear.





This indicates that the normality distributions of the random errors in mathematics achievement scores are tenable in both Lebanon and Singapore.

B.Multicollinearity

Multicollinearity refers to a high inter-correlation among the predictors or the independent variables. One way of detecting the multicollinearity is performing a correlation matrix among the predictors. Any correlation coefficient exceeding 0.80 [20] or 0.90 [21] can be considered a multicollinearity problem. Several correlation matrices were performed to check the multicollinearity among the predictors at the student level.

Table 9 Correlation matrix for factors at student level for Lebanon (overall sample)

	1	2	3	4	5	6	7
1	1						
2	.403**	1					
3	.268**	.652**	1				
4	.180**	.453**	.544**	1			
5	.293**	.192**	.146**	.139**	1		
6	.379**	.136**	.053**	.086**	.236**	1	
7	-.124**	-.060**	-.030	-.045**	-.151**	-.198**	1
n	3974	3974	3974	3974	3974	3974	3974

Note. **p<0.01; *p<0.05

1. Math achievement	5. Educational aspiration
2. Math self-concept	6. Home educational resources
3. Attitude towards math	7. Home language
4. Math valuing	

Table 10 Correlation matrix for factors at student level for Singapore

	1	2	3	4	5	6	7
1	1						
2	.412**	1					
3	.276**	.672**	1				
4	.124**	.420**	.557**	1			
5	.197**	.122**	.135**	.162**	1		
6	.365**	.165**	.056**	.052**	.190**	1	
7	-.146**	-.016	-.053*	-.029*	-.099**	-.308**	1
n	5927	5927	5927	5927	5927	5927	5927

Note. **p<0.01; *p<0.05

1. Math achievement
2. Math self-concept
3. Attitude towards math
4. Math valuing
5. Educational aspiration
6. Home educational resources
7. Home language

The correlation coefficients between mathematics achievement and the predictors for Lebanese and Singapore students (overall sample) at the student level are presented in Table 9 and Table 10. The correlations of all the predictors with mathematics achievement were positively significant ($p< 0.001$), except for "home language" which was negatively significant. The correlation coefficients varied between -0.124 for "home language" and 0.403 for "mathematics self-concept".

For Singaporean students, the correlation coefficients between mathematics achievement and the predictors at the student level are provided in Table 10. The correlations of all the predictors with mathematics achievement were positively significant ($p< 0.001$), except for "home language" which was negatively significant ($p< 0.001$). The correlation coefficients ranged between 0.124 for "mathematics valuing" and 0.412 for "mathematics self-concept".

The correlations between the predictors with mathematics achievement was positively significant only for "school emphasis on academic success - Principal repost", "school discipline and safety", and "school size" ($p< 0.001$) and it was positively significant between mathematics achievement and "teacher career satisfaction" and "problem with working conditions" at ($p< 0.05$) level. The correlation coefficients ranged between -0.004 ($p> 0.05$) for "school emphasis on academic success - Teacher repost" and 0.491 for "school size". Of the aggregated predictors, the correlations of all the predictors with mathematics achievement were significant, except for "mean students math valuing" which it was not significant($p> 0.05$). The correlation was positive for "mean students mathematics self-concept", "mean students attitude towards mathematics", "mean students educational aspiration" and "mean students home educational resources" ($p< 0.001$), but it was negative for "mean students home language" ($p< 0.001$). The correlation coefficients ranged between 0.002 for "mean students math valuing" and 0.789 for "mean students home educational resources".

Generally, having obtained high correlation coefficients between the outcome variable and the predictors is desired. The correlation coefficients between most of the predictors and the outcome variable were statistically significant at the ($p< 0.001$) level. The inter-correlations among a few predictors themselves were higher than the correlation between those predictors with the outcome variables. For example, the inter-correlations among the attitudinal factors were higher than the correlation between these factors with mathematics and science achievement. However, these inter-correlations did not exceed the criterion limit (0.80) that suggested in the literature [20]. Therefore, it was concluded that multicollinearity was not a problem in the present article.

V. THE RESULTS

A. *Research Question 1: Does the mathematics achievement of Lebanese and Singaporean eighth-graders differ across schools?*

The unconditional, null or One-Way random effect ANOVA model [19] was estimated to answer the first research question. The relationships among the components at each level in this model are presented in Equations 4.1 and the results are provided in Table 11.

The Null or One-Way random effect ANOVA model

Level-1 model:

$Y_{ij} = \beta_{0j} + r_{ij}$, $r_{ij} \sim N(0, \sigma^2), \sigma^2 =$ level-1 variance

Level-2 model: $\beta_{0j} = \gamma_{00} + u_{0j}$, $u_{0j} \sim N(0, \tau_{00}), \tau_{00} =$ level-2 variance

Combined model: $Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$ (Equations 1)

Table 11: The Null or One-Way random effect ANOVA model

Parameter	Lebanon	Singapore
Fixed effects		
Intercept	455.75**	605.98**
SE	4.09	4.38
t-ratio	111.178	138.11
P	0.000	0.000
Random effects		
School-level variance(τ_{00})	1571.00	3075.39
Student-level variance(σ^2)	3367.60	3624.84
χ^2	2136.46	5212.36
df	146	164
P	0.000	0.000
Intra-Class Correlation	0.3181	0.4589
Reliability	0.93	0.97
Deviance	43956	65962
Number of parameter	3	3
AIC	43962	65968
BIC	43971	65977

Note.*SE* = standard error; *df* = degrees of freedom; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; *p< 0.05; **p< 0.001

Table 11 presents the variance components and the relevant information for Lebanese and Singaporean students. For the Lebanese, the total variance in mathematics achievement is $(r_{ij}) = \sigma^2 = 3367.60$at the student level and it is $(u_{0j})\tau_{00} = 1571.00$ at school level. The ICC is computed through the following Equation.

$$\tau_{00}/(\tau_{00} + \sigma^2) \quad \text{(Equation 2)}$$

□ = 1571.00/ (1571.00+ 3367.60) = 0.3181.

This indicates that, 31.81% of the total variance in mathematics achievement was at the school level and 68.19% at the student level. The null hypothesis associated with the unconditional model is, H_0: $\tau_{00} = 0$, there is no difference in mathematics achievement across schools. This hypothesis was tested using Equation 3.

$$H = \sum n_j (\bar{Y}_{.j} - \hat{y}_{00})^2/\sigma^2, \quad \text{(Equation 3)}$$

The null hypothesis has a χ^2 distribution with *J*-1 degrees of freedom. The χ^2 associated with the variance components was 2136.46 with 146 (*J* = 147 schools) degrees of freedom, *p* = 0.000. The null hypothesis was rejected, which means mathematics achievement varied significantly across schools. The estimated intercept or grand mean (□$_{00}$) was 455.75, and the *t*-ratio (*df*, 146) = 111.17, *p* = 0.000 indicating that the grand mean differed significantly across schools. The grand mean with standard error of 4.18 yielded a 95% confidence interval of, 455.75□ 1.96(4.09) = (447.73, 463.76).

Once the result showed that mathematics achievement varied significantly across schools, the plausible values of the achievement within a 95% confidence interval were computed to determine the extent to what achievement scores vary across schools. The range of the plausible values within a 95% confidence interval were $455.75 \pm 1.96*(1571.00)^{1/2}$ = (378.07, 533.43), this shows a wide range in mathematics average. Similarly, an estimate of the reliability of the sample mean score in each school (□$_{0j}$) is derived from the model using the following Equation:

$$\lambda_j = \text{reliability}\,(\bar{Y}_{\cdot j}) = \tau_{00}\,[\tau_{00} + (\sigma^2/n_j)]; \qquad \text{(Equation 4)}$$

where n_j is the sample size for school j. The higher reliability coefficient indicates how well schools are differentiated in mathematics average. The reliability coefficient was 0.93, indicates that the school sample mean scores are reliably indicators of the true school mean scores.

For Singaporean students, the total variance in mathematics achievement is (r_{ij}) = σ^2 = 3624.84 at the student level and (u_{0j}) = τ_{00} = 3075.39 at the school level. The ICC was □ = 3075.39/ (3075.39+ 3624.84) = 0.4589, indicating that, 45.89% of the total variance in mathematics achievement was at the school level and 54.11% at the student level. The grand mean (□$_{00}$) was 605.98, and the *t*-ratio was (*df*, 164) = 138.11, p = 0.000. This shows that the grand mean differed significantly across the schools. The standard error of the grand mean was 4.38 yielded a 95% confidence interval of 605.98□ 1.96(4.38) = (597.39, 614.56). The plausible values of the mathematics achievement within a 95% confidence interval were computed, 605.98 ± 1.96*(3075.39)$^{1/2}$ = (497.30, 714.66). This indicates a wide range in mathematics average across schools. The estimated reliability coefficient of the sample mean scores was 0.97, indicates that the differences in mathematics achievement among schools were estimated with a high degree of the reliability.

Once the results of the unconditional model indicated that mathematics achievement varied significantly across schools in both countries, the subsequent models were estimated with the predictors at student and school level to explain the variance in mathematics achievement.

B. The Student-Level Models for Lebanon (Overall Sample) and Singapore

Research Question 2:

How much of the variance in mathematics achievement of Lebanese and Singaporean eighth-graders is associated with student-level factors – attitudinal factors and family background?

Before presenting the models' results, it is important to note that in order to assess the absolute effect of each group of the predictors on the outcome variables; the predictors were added to the model sequentially. That is, the predictors that introduced to the model they remained in the model while the new predictors were added. To facilitate the comparison between the two countries, the main estimated parameters for each group of the predictors are provided here and the detail information of the estimated parameters by each model is provided in Table 12.

Model 1: One way ANCOVA random intercept with the attitudinal factors

Model 1 was estimated by adding the attitudinal factors "mathematics self-concept", "attitude toward mathematics", "mathematics valuing", and "educational aspiration" to the Null Model. The relationships between these factors and mathematics achievement are presented in HLM Equations as follows:

Level-1 model: Y_{ij} = □$_{0j}$ +□$_1$(math self-concept $_{ij}$- ***mathself-concept ..***) +□$_2$ (attitude towards math $_{ij}$ – ***attitude towards math..***) + □$_3$ (math valuing $_{ij}$ - ***math valuing..***) + □$_4$ (educational aspiration $_{ij}$ – ***educationalaspiration..***) + r_{ij}
Level-2 model: □$_{0j}$ = □$_{00}$ + u_{0j} (Equations 5)

The bold and italic words in Equations 5 represent that the predictors are centered at the grand mean . The □$_1$ through □$_4$ are the regression coefficients or the expected change in mathematics achievement for a student is associated with one unit change in the values of the predictors. These four factors were treated as fixed effects, it were assumed that the effect of the factors are the same for all students across the schools. The results of the model are presented in the following table .

Table 12: Model 1 - One Way ANCOVA random effect with the attitudinal factors

Parameter	Lebanon		Singapore	
Fixed effects	Coefficient	SE	Coefficient	SE
Intercept	456.53**	3.64	606.10**	3.96
Mathematics self-concept	28.87**	1.5	13.94**	.92
Attitude towards math	4.38*	1.69	6.87**	1.15
Mathematics valuing	3.41*	1.42	0.67	1.09
Educational aspiration	5.29**	1.37	1.18	.85
Random effects		SD		SD
School-level variance (τ_{00})	1257.71	35.46	2514.93	50.14
School-level variance explained	19.94%	-	18.22%	-
Student-level variance (σ^2)	2748.06	52.42	3011.32	54.87
Student-level variance explained	18.39%	-	16.92%	-
Model fit index				
Reliability	0.93		0.97	
Deviance	43145		64860	
Number of parameter	7		7	
AIC	43159		64874	
BIC	43179		64895	

Note*.SE* = standard error; *SD* = standard deviation; AIC = Akaike InformationCriterion; BIC = Bayesian Information Criterion; *p< 0.05; **p< 0.001

Table 12 displays the results of Model 1. For the Lebanese, the estimated intercept was 456.53 with standard error of 3.64. This yielded a 95% confidence interval, 456.53 □ 1.96 (3.64) = (449.39, 463.66). The Null Model was used as a baseline to compute the proportion reduction in the variance. By comparing the variance components of Model 1 and those of the Null Model, the index of the proportion reduction in the variance or variance explained at the student level was developed using:

$$\left(\frac{\sigma^2 \text{ Null model} - \sigma^2 \text{ Model 1}}{\sigma^2 \text{ Null Model}}\right)$$

(Equation 6)

With the addition of the attitudinal factors, the variance components were substantially reduced at both student and school levels. The proportion reduction in the variance at the student level was (3367.60 - 2748.06)/ 3367.60 = 0.1839. The attitudinal factors accounted for 18.39% of the total student-level variance in mathematics achievement. Theoretically, it is possible that the student-level factors explain the school-level variance, but not vice versa. The proportion of the school-level variance explained by Model 1 was computed in the similar way of the student level by Equation 7.

$$\left(\frac{\tau_{00} \text{ Null model} - \tau_{00} \text{ Model 1}}{\tau_{00} \text{ Null Model}}\right)$$

(Equation7)

Accordingly, the proportion of the school-level variance explained by Model 1 was (1571.00 - 1257.71)/ 1571.00 = 0.1994. Overall, the attitudinal factors accounted for (18.39% + 19.94%) = 38.33% of the total variance in the Lebanese students mathematics achievement.

For the Singaporeans the, estimated intercept was 606.10 with standard error of 3.96, produced a 95% confidence interval, 606.10 ± 1.96(3.96) = (598.33, 613.86). The proportions of the explained variances were (3624.84 - 3011.32)/3624.84= 0.1692 at the student level and (3075.39 - 2514.93)/ 3075.39 = 0.1822 at the school level. Overall, the attitudinal factors accounted for 35.14% of the variance in the Singaporean students' mathematics achievement.

The mathematics self-concept was significantly linked to achievement in both countries. The effect of this factor was (β_1 = 10.44, p = 0.000) for the Lebanese, whereas it was (β_1 = 11.19, p = 0.000) for the Singaporeans. On average, one scale-point increasing in mathematics self-concept increased achievement by 10.44 points for the Lebanese and 11.19 points for the Singaporeans. Mathematics self-concept was the strongest predictor of mathematics achievement in both countries; however it

was slightly stronger for the Singaporeans students than for the Lebanese.

Attitude towards mathematics did not yielded a significant association with achievement for Lebanese, but the link was significant for Singaporeans. The effect of attitude was (β_2 = 1.77, p = 0.060) for the Lebanese and it was (β_2 = 1.74, p = 0.008) for the Singaporeans. One scale-point increase in attitude towards mathematics increased mathematics achievement by 1.77 points for the Lebanese and 1.74 points for the Singaporeans.

The association between mathematics valuing and mathematics achievement was not significant either for Lebanese students (β_3 = -0.37, p = 0.602) or for the Singaporeans (β_3 = -0.85, p = 0.106).

The link between educational aspiration and mathematics achievement was significant for both Lebanese students (β_4 = 9.74, p = 0.000) and Singaporeans (β_4 = 5.04, p = 0.000). Students who aspire for further education tended to achieve higher scores in mathematics. One scale-point increase in educational aspiration increased mathematics achievement by 9.74 points for the Lebanese and 5.04 points for the Singaporeans.

Model fit indices for Model 1, the deviance, AIC and BIC are given in Table 12. The difference between the deviance in Model 1 and the Null Model was (43956 - 43145 = 811) with degrees of freedom equal to difference between the number of parameters in Model 1 minus the number of parameters in the Null Model, which is (7-3=4). The χ^2 was (df, 4, α = 0.01) = 13.277, the difference in the deviances is far greater than the value of the Chi-square (p< 0.001). The lower the deviance the better the model fits the data. The difference between the AIC was calculated by subtracting the AIC of the Null Model from that of Model 1 (AIC= D+2P), where D refers to the deviance and p represents the number of parameters in the model. It was (43956 + 2*3 = 43962) for the Null Model and (43145 + 2*7 = 43159) for Model 1. The difference was (44760 – 43557 = 803), indicating that the AIC of Model 1 was much smaller than that of the Null Model. Similarly, the BIC of the two models were computed using: D+ Ln (n)*P, where n was treated as school-level sample size and it was the natural logarithm of the number of schools (147 schools from Lebanon) which was 4.99. Thus, the BIC was (43956 + 3*4.99 = 43971) for the Null Model and (43145 + 7*4.99 = 43179) for Model 1. The difference between these the two measures was (43971 – 43179 = 972) which provided a strong evidence [22] in favour of Model 1 over the Null Model. These measures were computed in the same way for the Singaporean students.

Model 2: One way ANCOVA random intercept with students' family background factors and gender

The family background factors "home educational resources", "home language" and "students' gender" were added to Model 2.

Level-1 model: $Y_{ij} = \beta_{0j} + \beta_1$(math self-concept $_{ij}$- ***mathself-concept..***) $+\beta_2$ (attitude towards math $_{ij}$- ***attitude towards attitude..***) $+ \beta_3$(math valuing $_{ij}$- ***math valuing..***) $+ \beta_4$ (educational aspiration $_{ij}$- ***educationalaspiration..***) $+\beta_5$(home educational resource $_{ij}$ - ***home educational resource..***) $+ \beta_6$ (home language $_{ij}$- ***home language..***) $+ \beta_7$ (students gender $_{ij}$-***students gender..***) $+ r_{ij}$

Level-2 model: $\beta_{0j} = \beta_{00} + u_{0j}$ (Equations 8)

Table 13 Model 2- One way ANCOVA random intercept with family background factors and students gender

Parameter	Lebanon		Singapore	
Fixed effects	Coefficient	SE	Coefficient	SE
Intercept	456.96**	3.27	606.21**	3.68
Home educational resources	5.82**	0.81	6.10**	0.75
Home language	-3.40	5.80	-2.38	1.47
Students gender	4.29	2.60	-5.76**	1.59
Random effects		SD		SD
School-level variance (τ_{00})	1000.11	31.62	2164.68	46.52
School-level variance explained	20.48%	-	29.61%	-
Student-level variance (σ^2)	2671.84	51.68	2925.95	54.09
Student-level variance explained	2.77%	-	19.28%	-
Model fit index				
Reliability	0.91		0.96	
Deviance	43006		64670	
Number of parameter	10		10	
AIC	43026		64690	
BIC	43056		64721	

Note.*SE* = standard error; *SD* = standard deviation; AIC = Akaike InformationCriterion; BIC = Bayesian Information Criterion; *p< 0.05; **p< 0.001

Table 13summarizes the results for Model 2 with the family background factors and students gender. For Lebanese students, the intercept was 456.96 with a standard error of 3.27, with 95% confidence interval yielded 456.96 ± 1.96 (3.27) = (450.55, 463.36). The proportion reduction in the variance was computed using Equation 4.6, Model 1 total student-level variance (σ^2) subtracted from that of Model 2 (2748.06 - 2671.84)/ 2748.06 = 0.0277.Model 2 accounted for a very small portion of the student-level variance, whereas it was accounted for a great proportion of the school-level variance (1257.71 - 1000.11)/ 1257.71 = 0.2048. Altogether, Model 2 contributed to (0.2048 + 0.0022) = 23.25% of the total variance in mathematics achievement.

For Singaporean students, the intercept was 606.21 with a standard error 3.68 yielded a 95% confidence interval, 606.21 □ 1.96 (3.68) = (598.99, 613.42). By adding the family background factors and students gender, the variance components were reduced both at the student and school levels. The contribution of these three factors to the student-level variance was (3624.84 - 2925.95)/ 3624.84 = 0.1928 and it was (3075.39 - 2164.68)/ 3075.39 = 0.2961 to the school-level variance. Overall, Model 2 explained (0.1928 + 0.2961 = 48.89%) of the total variance in mathematics achievement.

Of the three factors "home educational resources" was positively linked to the Lebanese students mathematics achievement (β_5 = 5.82, p = 0.000), whereas, "home language" (β_6 = -3.40, p = 0.558), and "students gender" (β_7 = 4.29, p = 0.100), did not yield a significant link with mathematics achievement.

For Singaporeans, similarly, "home educational resources" was positively linked to mathematics achievement (β_5 = 6.10, p = 0.000), there was a significant association between students gender and mathematics achievement (β_7 = -5.76, p = 0.001), but there was no significant association between home language and mathematics achievement (β_6 = -2.38, p = 0.105).

On average, after controlling for all the other factors in Model 2, one scale-point increasing in home educational resources increased mathematics achievement by 5.82 points and 6.10 for Lebanese and Singaporean students, respectively

Research Question 3: Do the effects of the student-level factors on the mathematics achievement vary across the Lebanese and Singaporean schools?

In Model 1 and 2, it was assumed that the student-level factors have the same effect on mathematics achievement of all students across the schools, thus, the factors were examined as fixed effects [17], [18]. To answer the third research question, all factors that had a significant effect in Model 2 were allowed varying across the schools. The relationships of the factors in Model 3 are shown in Equations 9 and 10 and the results are presented in Table 6.

Model 3: Random intercept and random slopes (one way ANCOVA with random intercept and slopes)

For Lebanese students:

Level-1 model: $Y_{ij} = \beta_{0j} + \beta_1$(math self-concept $_{ij}$- ***mathself-concept..***) $+\beta_2$ (attitude towards math $_{ij}$- ***attitude towards attitude..***) $+ \beta_3$ (educational aspiration $_{ij}$- ***educationalaspiration..***) $+\beta_4$(home educational resource $_{ij}$ - ***home educational resource..***) $+ r_{ij}$

Level-2 model: $\beta_{0j} = \beta_{00} + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

$\beta_{2j} = \gamma_{20} + u_{2j}$

$\beta_{3j} = \gamma_{30} + u_{3j}$

$\beta_{4j} = \gamma_{40} + u_{4j}$

(Equations 9)

For Singaporean students:

Level-1 model: $Y_{ij} = \beta_{0j} + \beta_1$(math self-concept $_{ij}$- ***mathself-concept..***) $+\beta_2$ (attitude towards math $_{ij}$- ***attitude towards attitude..***) $+ \beta_3$ (math valuing $_{ij}$- **math valuing..**) $+\beta_4$ (educational aspiration $_{ij}$- ***educationalaspiration..***) $+ \beta_5$ (home educational resource $_{ij}$ - ***home educational resource..***) $+ \beta_6$ (students gender $_{ij}$-***students gender..***) $+ r_{ij}$

Level-2 model: $\beta_{0j} = \beta_{00} + u_{0j}$

$\beta_{1j} = \gamma_{10} + u_{1j}$

$\beta_{2j} = \gamma_{20} + u_{2j}$

$\beta_{3j} = \gamma_{30} + u_{3j}$

$\beta_{4j} = \gamma_{40} + u_{4j}$

$\beta_{5j} = \gamma_{50} + u_{5j}$

$\beta_{6j} = \gamma_{60} + u_{6j}$

(Equations 10)

Table 14: Model 3- random intercept and random slopes

Parameter	Lebanon			Singapore		
Fixed effects	Coefficient	SE		Coefficient	SE	
Intercept	456.07**	3.40		609.50**	3.75	
Slopes	Coefficient	df	χ^2	Coefficient	df	χ^2
Math self-concept	14.36**	145	233.38	24.51**	138	229.38
Attitude towards math	33.47**	145	282.21	22.18**	138	235.20
Educational aspiration	12.42*	145	190.65	75.82**	138	248.26
Home educational resources	16.44**	145	226.71	51.95**	138	316.05
Math valuing	-	-	-	5.32	138	165.54
Students gender	-	-	-	81.70	138	130.86
Random effects		SD			SD	
School-level variance (τ_{00})	1040.98	32.26		2164.68	47.28	
School-level variance explained	0.00%	-		0.00%	-	
Student-level variance (σ^2)	2482.73	49.82		2597.34	50.96	
Student-level variance explained	7.08%	-		11.23%	-	
Model fit index						
Reliability	0.84			0.95		
Deviance	42942			64388		
Number of parameter	21			36		
AIC	43026			64460		
BIC	43056			64571		

Note.*SE* = standard error; *SD* = standard deviation; AIC = Akaike InformationCriterion; BIC = Bayesian Information Criterion; *p< 0.05; **p< 0.001

In Equations 9 and 10, in addition to the intercept, all the slopes of the predictors have been subscripted with *j*. This indicates, the effects of the predictors allowed to be varying across the schools. Table 6 depicts that the effects of all the four factors varied significantly across the Lebanon schools. Attitude towards mathematics had the greatest variation (u_{2j} =33.47, p = 0.000), this was followed by home educational resources (u_{4j} =16.44, p = 0.000), mathematics self-concept(u_{1j} =14.36, p = 0.000), and educational aspiration (u_{3j} =12.42, p = 0.007). The relationship between these factors with mathematics achievement differed significantly across the Lebanese schools.

For Singaporean students, the effects of four out of six factors varied significantly across the schools. The educational aspiration had the highest variation (u_{3j} =75.82, p = 0.000), this was followed by home educational resources (u_{4j} =51.95, p = 0.000), mathematics self-concept(u_{1j} =24.51, p = 0.001), and Attitude towards mathematics had the greatest variation (u_{2j} =22.18, p = 0.000). The effects of these factors on mathematics achievement differed significantly from school to school.

VI. CONCLUSION

This article aimed to explore how the variation in mathematics achievement of Lebanese and Singaporean eighth-grade students is distributed within and between schools. Exploring the pattern of variation in mathematics achievement of Lebanese eighth-grade students who attend public and private schools within and between schools was also one of the main objectives of this study. Assessing the contribution of several student-, and school-level factors to the variation in mathematics achievement was another main objective of the study. In addition, comparing and contrasting the similarities and differences in the sources of variations in the Lebanese, Singaporean, Lebanese public and private schools students' mathematics achievement, and the contribution of the individual factors on the achievement of students from the two countries and the two schools samples was also another objective of the study. The data analysed here obtained from a total of 3,974 and 5,927 eighth-grade students nested within 147 and 165 schools in Lebanon and Singapore, respectively. The data also are from 1,043, and 2,835 students nested within 48 and 94 public and private schools in Lebanon, in that order. The multilevel linear modelling technique was utilized using the HLM6.07 computer package.

REFERENCES

[1] Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching* (10 ed.): Pearson Educational International.

[2] Greaney, V., & Kellaghan, T. (2008). Assessing national achievement levels in education (Vol. 1). Washington, DC: The World Bank.

[3] Geary, D. C., & Hamson, C. O. (2000). Improving the mathematics and science achievement of American children: Psychology's role education directorate. *American Psychological Association.*

[4] Beaton, A. E., Postlethwaite, T. N., Ross, K. N., Spearritt, D., & Wolf, R. M. (1999). *The benefits and limitations of international educational achievement studies*. Paris:UNESCO International Institute for Educational Planning.

[5] Beaton, A. E., & O'Dwyer, L. M. (2002). Separating school, classroom, and student variances and their relationship to socio-economic status. In D. F. Robitaille and A. E. Beaton (Eds.), *Secondary Analysis of the TIMSS Data* (pp. 211-231): Kluwer Academic Publishers.

[6] Mullis, I. V.S., Martin, M. O., Minnich, C. A., Stanco, G. M., Arora, A., Centurino, V. A.S., & Castle, C. E. (2012). *TIMSS 2011 Encyclopaedia Education Policy and Curriculum in Mathematics and Science* (Vol. 1),TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

[7] Mullis, I. V.S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*, TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

[8] Mullis, I. V.S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 Assessment frameworks*: TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

[9] Mullis, I. V.S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 international mathematics report*: TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

[10] Mullis, I. V.S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international mathematics report*: TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

[11] Joncas, M., & Foy, P. (2012). *Sample design in TIMSS and PIRLS*. Retrieved from http://timssandpirls.bc.edu/methods/t-sample-design.html.

[12] Maas, C. J. M., & Hox, J. J. (2004). Robustness issues in multilevel regression analysis. *Statistica Neerlandica*, *58*, 127–137.

[13] Hox*, J. J. (2002). Multilevel analysis techniques and applications: Mahwah, NJ: Lawrence Erlbaum.*

[14] Dargatz, D. A., & Hill, G. W. (1996). Analysis of survey data. *Preventive Veterinary Medicine*, *28*, 225-237.

[15] Foy, P., & Olson, J. F. (2009). *TIMSS 2007 user guide for the international database*, TIMSS and PIRLS International Study Center, Lynch School of Education, Boston College.

[16] Asparouhov. T. (2005). Sampling weights in latent variable modelling. *Structural Equation Modelling*, *12*(3), 411–434.

[17] Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. Educational Researcher, 39, 142–151.

[18] Rabe-Hesketh, S., and Skrondal, A. (2006). Multilevel modeling of complex survey data. *Royal Statistical Society*, *169*, *Part 4*, 805–827.

[19] Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & Toit, M. (2004). *HLM 6: hierarchical linear and nonlinear modelling: Application and data analysis methods*: Lincolnwood, IL: SSI Scientific Software.

[20] Field, A. (2009). Discovering statistics using SPSS (3rd). London SAGE Publications.

[21] Orme, J. G., & Orme.T. (2009). *Multiple regression with discrete dependent variables*: Oxford University Press.

[22] Raftery, A. F. (1995). Bayesian model selection in social research. *Sociological methodology*, *25*, 111-169.