The Influence of Socioeconomic Status, Self-efficacy, and Anxiety on Mathematics Achievement in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA[*]

Dilara Bakan Kalaycıoğlu[a]

Assessment, Selection and Placement Center

Abstract

The purpose of this study is to examine relationships among the such variables as socioeconomic status, math self-efficacy, anxiety, and mathematics achievement using structural equation modeling. A sample group of 8,806 students from England, Greece, Hong Kong, the Netherlands, Turkey, and the USA participated in the PISA 2012. To show how much variance on mathematics achievement can be attributed to the selected variables, separate structural equation modeling analyses were examined for each country. Next, multi-group structural equating modeling was determined to compare latent means. The results show that socioeconomic status has a significant effect on mathematics achievement. The relationship between socioeconomic status and mathematics achievement is highest in the Netherlands lowest in Hong Kong. For all six countries, the most important predictor of mathematics achievement is math self-efficacy. The relationship between math self-efficacy and mathematics achievement is highest for England. No statistically meaningful relationship between math anxiety and mathematics achievement is found for Hong Kong, England, or the Netherlands. A statistically meaningful relationship between math anxiety and mathematics achievement is observed in the countries that have low mathematics achievement; namely, Turkey, Greece, and the USA.

Keywords: Mathematics achievement • Measurement invariance • Socioeconomic status • Self-efficacy • Anxiety

International large-scale assessment studies provide comparative data to countries both to evaluate their education systems' performance and to give information about the factors related to students' achievement. The Programme for International Student Assessment (PISA) is one of these international assessment studies conducted by the Organization for Economic Co-operation and Development (OECD) every three year since 2000. The PISA target population consists of 15-year-old students and aims to assess students' ability to use their knowledge and skills in order to meet real-life challenges (OECD, 2013b). Mathematics literacy was the primary focus of the 2012 assessments. The PISA defines mathematical literacy as *"an individual's capacity to identify and understand the role that mathematics play in the world, to make well-founded judgments, and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned, and reflective citizen"* (OECD, 2009, p. 14).

Research on factors related to students' mathematics achievement and socioeconomic status (SES) has received broad attention. Various studies have examined socioeconomic status, finding it to be one of the strongest predictors of academic achievement (Caldas & Bankston, 1997; Papanastasiou, 2000; Sirin, 2005) in not only international large-scale assessments (Chiu, Chow, & Mcbridge-Chang, 2007; Chiu & Xihua, 2008), but also school level assessments (Engin-Demir, 2009; Ma & Klinger, 2000).

Researchers have also indicated that affective factors, such as math self-efficacy and math anxiety, play a crucial role in mathematics achievement. Social learning theorists define perceived self-efficacy as "to believe in one's capabilities to organize and execute the courses of action required to manage prospective situations" (Bandura, 1993, p. 2). When this theory is applied to subject specific self-efficacy, such as math self-efficacy, children with high math self-efficacy would be likely to demonstrate higher achievement in mathematics (Lent, Brown, & Gore, 1997). Furthermore, those who maintain a resilient sense of self-efficacy set challenging goals for themselves, make good use of analytic thinking skills, and have a firmer commitment to reach these goals. Having a high level of self-efficacy also regulates and reduces both stress and anxiety (Bandura, 1993; Bandura & Locke, 2003). Math anxiety in the PISA is defined as students' feelings of helplessness and stress when dealing with mathematics (OECD, 2013a). The components of math anxiety were found to be similar to those of identified in test anxiety (Wigfield & Meece, 1988) with some researchers describing math anxiety as a subject specific test anxiety (Bandalos, Yates, & Thorndike-Christ, 1995; Hembree, 1990). The predictive power of mathematic self-efficacy and math anxiety on students' mathematic**s** achievement has been well documented. Self-efficacy is positively associated with students' academic performance (Alcı, Erden, & Baykal, 2010; Chiu & Xihua, 2008; İş Güzel & Berberoğlu, 2010; Lee & Stankov, 2013), whereas anxiety is negatively associated with students' academic performance (Cassady & Johnson, 2002; Hembree, 1990; Ho et. al., 2000; Ma, 1999; Seipp, 1991; McCarthy & Goffin, 2005). However, just as math anxiety's effect on mathematics achievement tends to be relatively small, its debilitating effect neither direct nor significant (Meece, Wigfield, & Eccles, 1990).

At the country level, math self-efficacy and math anxiety are associated with mathematics performance. Across OECD countries, a 28% and 14% variation in students' performance in mathematics can be explained by differences in students' reported levels of math self-efficacy and math anxiety, respectively (OECD, 2013a). At the cross country level however, the relationship between math self-efficacy, anxiety, and achievement is relatively complicated (Bodas & Ollendick, 2005; Liu, 2009). One's perception of self-efficacy may have different meanings across different cultures. According to Schaubroeck, Lam, and Xie (2000), people's perception of self-efficacy differed between Western individualist cultures and Eastern collectivist cultures. Compared to self-efficacy however, math anxiety is well-defined across cultures (Bodas & Ollendick, 2005; Ho et al., 2000). Even though math self-efficacy is positively and math anxiety is negatively associated with mathematics performance, generally studies have shown that high performing Asian students' have lower self-efficacy levels and higher math anxiety compared to students of Western European countries and the USA (Lee, 2009; Morony, Kleitman, Lee, & Stankov, 2013; Shen & Pedulla, 2000). Furthermore, Lee (2009) concludes that while Asian and Eastern European countries demonstrate the strongest association of math self-efficacy and mathematics performance, there is a stronger association between the math anxiety and mathematics performance in European countries compared to Asian countries.

Although many studies have shown that students' socioeconomic status, math self-efficacy, and math anxiety are associated with mathematics achievement, the relationship of these factors

with the students' achievement across countries is not well known. Some researchers have explained this association by stating that lower performing countries' have that low expectations and low academic standards (Shen & Pedulla, 2000; Shen & Tam, 2008) whereas other researchers make a connection with countries' cultural differences (Morony, et al., 2013; Stankov, 2010; Wilkins, 2004).

SES, math self-efficacy, and math anxiety variables are expected to predict students' mathematics achievement scores as described above. In the current study, we use PISA 2012 data to examine how students' SES, math self-efficacy, and math anxiety are associated with mathematics achievement in England, Greece, Hong Kong, Netherlands, Turkey, and the USA. These six countries were selected due to their different rankings based on mathematics achievement, SES, math self-efficacy, and math anxiety variables. According to PISA 2012 results, while the mathematics achievement scores of Hong Kong and the Netherlands are statistically significantly above the OECD average, and those of the USA, Turkey, and Greece are statistically significantly below the OECD average. The exception is England's mathematics achievement score, which is close to the OECD average. Among the countries chosen, while Turkey has the lowest SES index, Hong Kong's SES index is relatively low. England and the Netherlands are among the top countries based on their SES rankings. However, while SES has been demonstrated itself to be a stronger predictor of mathematic achievement in Turkey and the Netherlands, it was found to be a weaker predictor of mathematic achievement in Hong Kong. Furthermore, when compared to the relatively low performing countries of the USA and England, Hong Kong was found to have higher math anxiety levels and the Netherlands lower self-efficacy levels, despite the fact that the latter two countries were among high performing countries. Both Greece and Turkey were low performing countries with Greece having the lowest math self-efficacy level and Turkey the highest math anxiety level among the six countries.

Although there are many studies on PISA data, most of them either compare the results of high performing countries or give a comparison of all participating countries. In the literature, there is lack of work investigating high and low performing countries and a comparative evaluation of these countries. This study is important in terms of evaluating the influence of the affective factors on students' math performance levels from countries

with a competitive education system, such as Turkey, Hong Kong, and Greece, and those of students from the USA, the Netherlands, and the UK, all of which have a relatively less competitive education system than the former three countries. In addition little empirical research has been done directly linking math self-efficacy and math anxiety variables with mathematics achievement in Turkey.

Cross-cultural Studies and Measurement Invariance

International large-scale assessments' primary goal is to provide comparative data and to draw conclusions about target variables in different countries. However, in cross-country comparisons, there are many confounding variables that contribute to score differences, such as national curricula, characteristics of the language, translation mistakes and cultural-specific experiences. To compare the achievement levels of students who take different language versions of an assessment, the raw scores from each assessment should be transformed into a common scale (Hambleton, 1994; Sireci, 1997). Furthermore, in order to meaningfully compare constructs across countries, variables should be measured with a high degree of equivalence in all countries. More technically, establishing measurement invariance is a prerequisite of cross-country comparisons. Meredith (1993) defines measurement invariance as an observed score, stating that is said to be invariant if a person's probability of an observed score does not depend on his/her group membership, conditional on the true score. Thus, measurement invariance is the necessary condition for scores to be considered comparable in the first place. Meredith describes factorial measurement invariance as configural, weak, strong, and strict factorial invariance. Configural invariance is achieved when the same factor model fits across all groups No equality constraint is imposed in configural invariance. Weak invariance requires factor loadings to be identical across the groups. In addition to the equality constraints imposed by weak invariance, strong invariance requires intercepts to be equal across groups. Strict invariance, on the other hand, requires error variances in addition to previous models' constraints. However, equal factor loadings, equal intercepts, and equal error variances for all groups are unlikely to hold in practice. Byrne, Shavelson, and Muthén (1989) introduced the concept of "partial measurement invariance," in which only a subset of parameters

Table 1
Sample Sizes, Scale Means, and Standard Deviations of Six Countries Participating in the PISA 2012

Countries	N	SES		MATEFF		MATANX		MATACH	
		M	SD	M	SD	M	SD	M	SD
England	1316	.31	.78	18.14	3.62	11.49	3.20	502.89	92.54
Greece	1582	-.02	.98	17.01	4.02	12.37	3.41	460.92	85.94
Hong Kong	1496	-.83	.96	18.43	3.78	12.38	3.20	562.47	96.53
Netherlands	1344	.25	.80	17.59	3.55	10.94	2.96	526.84	85.64
Turkey	1504	-1.44	1.11	18.05	3.47	13.02	3.41	453.24	91.20
USA	1564	.19	.98	18.16	3.71	11.81	3.57	487.35	87.91

Note. SES: socioeconomic status; MATEFF: math self-efficacy; MATANX: math anxiety; MATACH: mathematics achievement.

must be fixed and the others are simply allowed to vary between groups.

Given the theoretical background of mathematics achievement predictors, how they vary based on country, and the measurement invariance outlined above, we have sought answers to the following research questions regarding the interrelationship between mathematics achievement, SES, math self-efficacy, and math anxiety of students in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA.

1. Do math self-efficacy and math anxiety instruments measure the same constructs (measurement invariance) in the six sample countries?

2. To what extent are SES, math self-efficacy, and math anxiety variables related to mathematics achievement levels in the six sample countries?

3. To what extent do SES, math self-efficacy, math anxiety, and their relationship to mathematics achievement in the structural models path coefficients differ across the six sample countries?

Method

Research Design

The correlational research model was used to investigate relationship patterns among SES, math self-efficacy, math anxiety, and mathematics achievement. Structural Equation Modeling (SEM) makes it possible not only to assess latent variables, but to test a conceptual model with multiple dependent variables while using several regression equations simultaneously. In addition, multi-group structural equation models (MG-SEM) permit group differences to be tested for in specific parameter estimates.

Data and Sample

The data analyzed in this research were obtained from the PISA 2012 assessment. Approximately 510,000 students completed the PISA 2012 assessment from 65 participating countries. The data of England, Greece, Hong Kong, the Netherlands, Turkey, and the USA were taken into consideration. Three states (Florida, Connecticut, and Massachusetts) are not included in the study due to the fact that they participated as individual entities separate from the rest of the United States in the PISA 2012. The PISA employed a two-stage sampling procedure so as to ensure a representative sample of the target population for each country (OECD, 2013b).

Since participating students took different combinations of items and student questionnaires in the PISA 2012, we have included only those students who had completed data on the math self-efficacy and math anxiety scales.

Table 1 reports the sample sizes, scale means, and standard deviations of each country.

Measures

In addition to performance tests, students also answered a questionnaire. The stems and items for each scale are listed in Table 2.

Mathematics Achievement: The PISA measures students' ability to handle certain mathematics processes occurring in a real world context. The mathematics achievement test contains four content categories: quantity, uncertainty and data, change and relationship, space and shape. Some questions are multiple choice and others require students to construct their own response.

Since the PISA assessment uses an incomplete assessment design, students are required to answer a subset of the item pool. PISA estimates students' test scores as plausible values with each student having five plausible values for mathematics

performance. Plausible values represent the range of abilities that a student might reasonably possess based on a combination of their responses to the subset of items that they receive and other relevant background information (Mislevy, 1991; Wu, 2005). The PISA calculates students' mathematics achievement scores with a mean of 500 and standard deviation of 100. Five plausible values (PV1MATH-PV5MATH) for students' mathematics performance are considered as the outcome variables. A separate data analysis on mathematics achievement was run for each of the five plausible values and in order to provide unbiased estimates. The results have been averaged according to the 2009 OECD protocols.

Socioeconomic Status: Students' socioeconomic status is assessed using the educational, social, and cultural status (ESCS) index. The ESCS index is a composite index comprised of three separate indexes: parents' highest level of education, highest parental occupation, and cultural economic resources. The index maintains a mean of 0 and a standard deviation of 1 for students from OECD countries (OECD, 2009).

Math Self-efficacy: Self-efficacy items asked students to report on how confident they felt in their abilities to complete a range of mathematical tasks. Students responded to eight items, stating whether they felt "very confident," "confident," "not very confident," or "not at all confident." In this study, among the total eight self-efficacy items, only six are analyzed. The highest response code was indicative of a positive rating of self-efficacy.

Math Anxiety: Items gauging math anxiety includes five statements asking students to report to what extent they agree with the given statements. Responses consisted of the following values: "strongly agree," "agree," "disagree," and "strongly disagree." Similar to the previous scale, the highest response code indicated a positive rating of math anxiety.

Analysis

We employed both an exploratory (EFA) and confirmatory factor analysis (CFA) for each country using math self-efficacy and math anxiety variables to evaluate the measurement model's construct validity. The measurement invariance of these variables was then tested using a multi-group confirmatory factor analysis (MG-CFA). Separate SEM analyses were examined for each country to determine the amount of variance on student mathematics achievement that can be attributed to the selected factors. Finally, MG-SEM

are determined to compare latent means and path coefficients of countries'.

The most commonly used goodness of fit indices are RMSEA (Root mean square error of approximation), CFI (Comparative fit index), NFI (Normed Fit Index), and GFI (Goodness of Fit Index). A RMSEA value of less than .05 to .08 indicates a close fit. CFI, NFI, GFI values close to .90 or .95 reflect a good fit (Hu & Bentler, 1998; Marsh, Hau, & Wen, 2004; Schumacker & Lomax, 2010). Although χ^2 test statistics are provided, RMSEA, CFI, NFI, and GFI fit indices were interpreted due to the χ^2 test's sensitivity to sample size to assess model fit. Since χ^2 is a function of a sample, it may reject trivial model-data differences when it is used with a large sample size (Browne & Cudeck, 1993; Wu, Li, & Zumbo, 2007).

Following Wu et al.'s (2007) recommendation which is $\Delta\chi^2$ is too strict and CFI is too lenient, a change in CFI of less than .01 (ΔCFI$\leq$ -0.01) decision rule is used to evaluate weak, strong, and strict invariance models in multi-group measurement invariance .

The violation of normality assumption of analysis procedures can affect estimates and significance tests in ordinal variables. Since the observed variables represent responses to a set of items in the form of a four-category Likert scale, we used PRELIS to estimate the polychoric correlations and asymptotic covariance matrix. We then used these matrices in LISREL to estimate the weighted least squares (WLS) in CFA and MG-CFA (Jöreskog, 2002). We also used robust maximum likelihood procedure in SEM analyses. In addition, to evaluate the significance of path coefficients, a *t*-test was conducted.

All analyses reported in this paper were carried out using PRELIS and LISREL 8.8 for Windows (Jöreskog & Sörbom, 2006).

Results

Using an EFA to derive factor loadings for each country allows researchers to ascertain to what extent items load onto scales differently between countries. According to the EFA's results, two math self-efficacy items were deleted from the math self-efficacy scale because they loaded different factors for Greece and Turkey's data. Factor analyses were repeated for the remaining math self-efficacy and anxiety items with the data of every country in the study. In order to save space, we have presented the EFA's results for the data of combined six countries in Table 2.

Table 2
Standardized Factor Loadings of the EFA

Items	Factor 1	Factor 2
Math self-efficacy (MATEFF)		
How confident do you feel about having to do the following mathematics tasks?		
St37Q01. Using a train timetable to work out how long it would take to get from one place to another	.704	
St37Q02. Calculating how much cheaper a TV would be after a 30% discount	.726	
St37Q03.Calculating how many square meters of tiles you need to cover a floor	.779	
St37Q04. Understanding graphs presented in newspapers	.701	
St37Q06. Finding the actual distance between two places on a map with a 1:10,000 scale	.716	
St37Q08.Calculating petrol consumption rate of a car	.688	
Math anxiety (MATANX)		
Thinking about studying mathematics: to what extent do you agree with the following statements?		
St42Q01.I often worry that it will be difficult for me in mathematics classes		.786
St42Q03. I get very tense when I have to do mathematics homework		.810
St42Q05. I get very nervous doing mathematics problems		.791
St42Q08. I feel helpless when doing a mathematics problem		.772
St42Q10. I worry that I will get poor grades in mathematics		.720

As can be seen in Table 2, the two extracted factors correspond to the math self-efficacy and anxiety factors. The EFA conducted on the selected 11 items derived two factors, accounting for a total variance of 57%. Six items were included in the math self-efficacy factor and five items on the math anxiety factor. The composite reliability was found to be .87 for math self-efficacy and .92 for math anxiety. Since composite reliability is calculated from factor loadings, it produces a precise estimation of reliability. The Cronbach's coefficients alpha for the math self-efficacy factor was .82, .80, .87, .82, .77, and .84 for England, Greece, Hong Kong, the Netherlands, Turkey, and the USA, respectively. Similarly, the Cronbach's coefficients alpha for the math anxiety factor was .85, .82, .85, .85, .80, and .88 for England, Greece, Hong Kong, the Netherlands, Turkey, and the USA, respectively. The math self-efficacy and math anxiety scales appeared to have good internal consistency for all six countries.

Variables' correlation values have been presented in Table 3.

Table 3
Bivariate Correlations among Scales (N = 8,806)

	SES	MATEFF	MATANX	MATACH
SES	1.000			
Math self-efficacy	.153	1.000		
Math anxiety	-.187	-.318	1.000	
Mathematics achievement	.289	.459	-.340	1.000

Note. p < .01. All correlations are significant at a level of .01.

Correlations among the scales showed that the signs and patterns of all correlations are within theoretical expectations. Small and moderate sized correlations are found between SES, math self-efficacy, and math anxiety. These correlations indicate that although the constructs are independent, they are related to each other. Table 3 also presents correlations between the predictive variables and PISA 2012 mathematics achievement scores. As illustrated in Table 3, the highest positive correlation is found between math self-efficacy and mathematics achievement.

To confirm the structure obtained through the EFA, we conducted a CFA for the math self-efficacy and math anxiety scales separately for each country. The results of the CFA models have been presented in Table 4.

Table 4
Fit Indices for Confirmatory Factor Analysis

Country	χ^2	RMSEA	CFI	NFI	GFI
England	178.04	.051	.97	.96	.99
Greece	220.13	.051	.95	.94	.99
Hong Kong	197.01	.049	.97	.97	.99
Netherlands	207.21	.051	.96	.94	.99
Turkey	180.19	.046	.96	.94	.99
USA	169.15	.043	.98	.97	.99

Note. df = 43.

As observed in Table 4, RMSEA values ranged from .043 to .051, and all CFI, NFI, and GFI values were either close to or greater than .95. These results indicate an acceptable model fit based on the above-mentioned cut-off criteria. All standardized factor loadings in the model are significant at a level of α = .05. The results of the CFAs therefore suggest that math self-efficacy and math anxiety variables were both sufficiently identified and reliably measured for the six countries.

In order to evaluate measurement invariance with ordinal variables, a set of thresholds were defined for each variable and rendered the same for each country (Jöreskog, 2002). Table 5 displays the fit indices for the configural, weak, partial weak,

strong, partial strong, and strict invariance models that tested measurement invariance.

Invariance Model	χ^2	Df	RMSEA	CFI	ΔCFI
Configural	1169.92	258	.049	.966	-
Weak	1658.86	303	.055	.949	-.02
Partial weak	1525.65	298	.053	.954	-.01
Strong	2682.49	348	.067	.912	-.04
Partial strong	1914.32	317	.058	.940	-.01
Strict	6603.75	358	.109	.764	-.15

The first model assessed configural invariance and yielded an acceptable fit (*RMSEA* = .049, *CFI* = .97). After configural invariance, the model is eligible to be tested for weak invariance. Testing for weak invariance failed (*RMSEA* = .049, *CFI* = .949, *ΔCFI* = -0.02) as the CFI decreased significantly ΔCFI≥ -0.01. Relaxing constraints for four of the items' factor loadings in the data for Turkey and one item for the data in the Netherlands, a partial factorial invariance model yielded a non-significant difference compared to the configural invariance model (*RMSEA* = .049, *CFI* = .954, *ΔCFI* = -0.01). The next model, testing for strong invariance, failed (*RMSEA* = .067, *CFI* = .912, *ΔCFI* = -0.04) as the CFI decreased significantly ΔCFI≥ -0.01. To obtain a partially strong invariance model, we relaxed the constraints for three intercepts in the data of Greece, the Netherlands, and Turkey and four intercepts for the data of Hong Kong and the USA. As a result, the existence of partial strong invariance was established (*RMSEA* = .058, *CFI* = .940, *ΔCFI* = -0.01). Finally, we tested the strict invariance model,

and it also failed (*RMSEA*= .109, *CFI* = .764, *ΔCFI* = -0.15). Since the partial strong invariance was established for the measurement model, latent means can be compared across countries. Table 6 presents the means and standard deviations of math self-efficacy and math anxiety factors.

	MATEFF		MATANX	
Countries	*M*	*SD*	*M*	*SD*
England	.00	.75	.00	.69
Greece	-.23	.92	.14	.80
Hong Kong	.08	.79	.19	.70
Netherlands	-.15	.69	-.13	.62
Turkey	-.04	.73	.31	.85
USA	-.03	.74	.04	.79

The means of the latent variables are assumed to be zero in the data for England. For math self-efficacy, Hong Kong has the highest mean value and Greece the lowest. With regard to math anxiety, the Netherlands has the lowest mean value and Turkey the highest. The order of ranks for latent means and raw score means are almost the same.

After the factor means are compared, the data were examined separately for each country to determine how much variance student mathematics achievement can be attributed to the factors. We employed a structural model using SES, math self-efficacy, and math anxiety variables as predictors of the students' mathematics performance. SEM values are presented in Figure 1; however, they
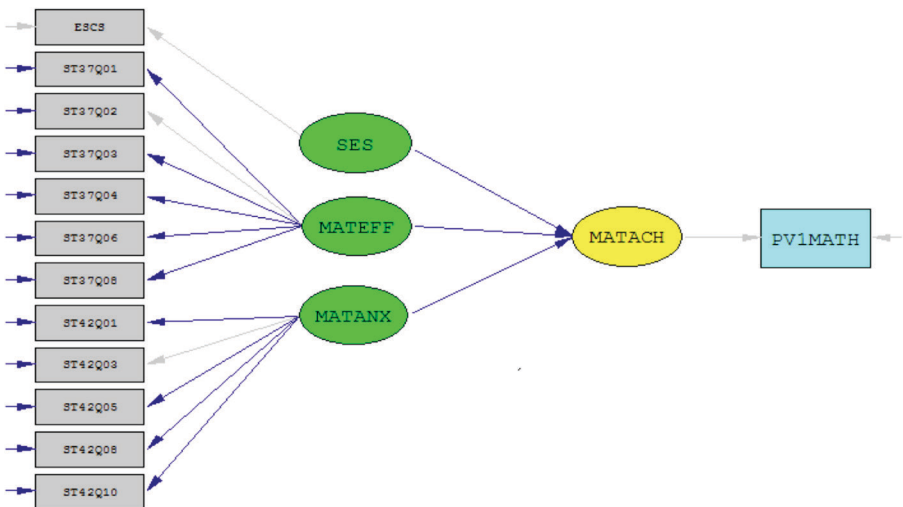

Figure 1: Structural equation model of SES, math self-efficacy, math anxiety, and PISA 2012 mathematics score (PV1MATH only).

Table 7
Fit indices and standardized path coefficients of structural equation models

Countries	SES	MATEFF	MATANX	χ^2	RMSEA	CFI	NFI	GFI	R^2
England	.19*	.48*	-.14	171.14	.038	.99	.99	.93	.43
Greece	.24*	.36*	-.27*	280.70	.048	.98	.98	.94	.38
Hong Kong	.15*	.50*	-.14	255.21	.046	.99	.99	.92	.40
Netherlands	.25*	.44*	-.07	250.76	.048	.99	.98	.93	.31
Turkey	.27*	.35*	-.18*	220.27	.042	.99	.98	.95	.32
USA	.23*	.38*	-.25*	222.74	.041	.99	.99	.94	.42

Note. df = 61. *p < .01.

represent PV1MATH only. For each country, five different SEMs were employed using five different mathematics plausible values.

Table 7 shows the standardized path coefficients and fit indices of the SEMs that estimate the model displayed in Figure 1 for each country. Estimates were based on five separate SEMs conducted with different plausible values (PV1math-PV5math) and then averaged for each country.

As can be seen from Table 7, RMSEA values ranged from .038 to .048 and all CFI, NFI, and GFI values were greater than .92. Thus, fit indices indicate a well-fitting model for all six countries. Variance in mathematics achievement of countries can be accounted for the SES, math self-efficacy and math anxiety scales for a range of 31% to 43%.

According to the path coefficients, it is apparent that the most important predictor of mathematics achievement for all countries is math self-efficacy. The importance of SES and math anxiety variables as predictors of mathematics achievement differs from country to country. For instance, while the math anxiety coefficient is higher than that of SES for Greece it is the opposite for the Netherlands. Moreover, while math anxiety has significant path coefficients for Greece, Turkey, and the USA, it has non-significant path coefficients for England, Hong Kong, and the Netherlands. Since countries were analyzed separately, path coefficients are not of the same scale and therefore not comparable. These results should be interpreted separately for each

country. In order to compare the means of latent variables and path coefficients for each country, we employed the MG-SEM described in the introduction and shown in Figure 2. Standardized path coefficients, estimated regression coefficients standard error, and squared multiple correlations from the MG-SEM have been reproduced in Table 8.

The MG-SEM resulted in $\chi^2(381)$ = 1335.06, RMSEA = .042, CFI = .99, NFI = .99, and GFI = .94. Overall, model fit seems good based on the fit indexes of the values selected. As illustrated in Table 8, SES, math self-efficacy, and math anxiety explain 32% to 43% of the variation in mathematics achievement for six countries. Among the six countries compared, the Netherlands exhibits the highest association between SES and mathematics achievement and Hong Kong the lowest. In all six countries, math self-efficacy has the strongest effect on mathematics achievement, and math anxiety is negatively related to mathematics achievement. When the path coefficients are compared, the highest coefficient for math efficacy belongs to England and the highest coefficient for math anxiety belongs to Greece. Math anxiety path coefficients have non-significant t values for Hong Kong, England, and the Netherlands.

Discussion

In this study, the 2012 PISA data is considered to determine how students' SES, math self-efficacy

Table 8
Standardized path coefficients, estimated path coefficients, standard errors, and squared multiple correlations

Countries	SES			MATEFF			MATANX			
	β	b	SE	β	b	SE	β	b	SE	R^2
England	.24*	23.07	4.87	.60*	59.90	18.53	-.16	-22.57	18.22	.43
Greece	.23*	21.94	1.95	.36*	36.00	6.40	-.25*	-36.61	7.07	.39
Hong Kong	.15*	13.82	2.94	.40*	40.00	10.03	-.19	-26.96	15.61	.37
Netherlands	.28*	26.58	3.46	.39*	39.40	10.61	-.08	-11.44	17.46	.34
Turkey	.24*	22.46	2.43	.41*	40.61	10.88	-.14*	-20.21	7.99	.32
USA	.22*	20.98	2.40	.44*	44.44	11.19	-.22*	-32.49	9.39	.40

Note. df = 381. *p < .01.

and math anxiety are associated with mathematics achievement among in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA. The exploratory and confirmatory factor analyses' results support the construct validity of the Student Questionnaire's math self-efficacy and anxiety items for the six countries. Measurement invariance of these variables was tested using a multi-group confirmatory factor analysis (MG-CFA) and partial strong invariance was established for the measurement model, rendering latent means comparable. The theoretical model fits are acceptable for all countries regarding the separate SEM analyses. Therefore, unified data from six countries were examined by the MG-SEM.

According to the SES factor, the relationship between SES and mathematics achievement is highest in the Netherlands and the lowest in Hong Kong. Grouping high SES students into a small number of schools makes the relationship between school SES and student achievement stronger. For example, low track schools tend to be dominated by socioeconomically disadvantaged students in many countries with educational academic tracking systems. Unsurprisingly, for all six countries, math self-efficacy has the strongest effect on mathematics achievement. Since students' perception of their own math self-efficacy represents how confident they are in performing mathematical tasks or in succeeding on mathematical cognitive skills, math self-efficacy is the most salient predictor of mathematics achievement. Results of this study support the importance of math self-efficacy on predicting mathematics achievement (Liu, 2009; McConney & Perry, 2010). While the highest mathematics achievement and math self-efficacy mean values belong to students from Hong Kong, the association between math self-efficacy and mathematics achievement is the highest for England. In England, portfolios are effectively used to evaluate students. These portfolios show what students can accomplish and how successful they are at different subject based on pre-determined standards. The students are also asked to evaluate themselves. Therefore, English students may be more aware of their own abilities and performance on specific mathematics tasks, leading them to give more reliable and consistent responses.

Finally, math anxiety is negatively related to mathematics achievement for all studied countries. The highest math anxiety mean value belongs to Turkey. Greece and Hong Kong also have relatively high math anxiety mean values. Math anxiety levels may be explained by the different educational

systems in these countries. Educational systems of developing countries, such as those of Turkey, Greece, and Hong Kong, are more centralized and more challenging for students. These countries education systems currently use high stake testing as a means of allocating students into academic high schools and students are ranked depending on their test scores. In Hong Kong, Turkey, and Greece, students are allowed access to specific schools based on their exam score starting in middle school. In these countries, shadow educational institutions (frontistrio-cram schools-dershane) are common to prepare students for high stake examinations. Since the PISA study is applied to the 15-years old students, these students' anxiety levels may be explained by the previous exams that they have taken before the PISA.

When the impact of math anxiety on mathematics achievement is considered, the highest impact was observed for Greece. Although students from Hong Kong experience high levels of math anxiety, the relationship between math anxiety and math achievement is not statistically significant. On the other hand, although students from the USA experience low levels of math anxiety, the relationship between math anxiety and math achievement is statistically significant. These results may be explained by the country's mathematics achievement results rather than the country's math anxiety mean values. Among the six countries, there is no statistically meaningful relationship between mathematics anxiety and mathematics success in those countries that have high mathematics achievement. Yet, in those countries that have low mathematics achievement such as Turkey, Greece and the USA, there is a statistically meaningful relationship between mathematics anxiety and mathematics achievement. Although it is difficult to generalize the results, it is believed that the effect of math anxiety on mathematics achievement decreases as mathematics achievement increases.

Each country's education system countries should be discussed according to the relationship between the discussed variables and mathematics achievement so as to decrease the negative influence of high competition on students' affective factors. There are examples of countries, such as England, in which students have both high academic achievements and self-efficacy with a reasonable level of anxiety. In order to increase math achievement levels, there is much to learn from countries that have a balance between high math self-efficacy and math anxiety levels.

Based on the study results, the following recommendations are presented for future research:

-Similar research can be conducted on additional countries and the results can be compared in order to generalize the findings.

-More research on the effects of math anxiety is needed in high school senior students. Since they are about to take university entrance examinations, they may experience more math anxiety compare to 15 years old students.

-Although this study did not examine math self-concept, math interest, enjoyment of mathematics, or learning strategies, the examination of these variables is an important area for a potential research study.

-Since it is the most influential variable on mathematics achievement in this study, further investigation is required to determine how to increase students' math self-efficacy.

References

Alcı, B., Erden, M., & Baykal, A. (2010). Üniversite öğrencilerinin matematik başarıları ile algıladıkları problem çözme becerileri, özyeterlik algıları, bilişüstü özdüzenleme stratejileri ve ÖSS sayısal puanları arasındaki açıklayıcı ve yordayıcı ilişkiler örüntüsü. *Boğaziçi Üniversitesi Eğitim Dergisi, 25*(2), 53–68.

Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology*, *87*(4), 611–623.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, *88*(1), 87–99.

Bodas, J., & Ollendick, T. H. (2005). Test anxiety: A cross-cultural perspective. *Clinical Child and Family Psychology Review*, *8*(1), 65–88.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*(3), 456–466.

Caldas, S. J., & Bankston, C. III (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, *90*, 269–277.

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, *27*(2), 270–295. doi:10.1006/ceps.2001.1094

Chiu, M. M., & Xihua, Z. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction, 18*, 321–336. doi:110.1016/j.learningstruc.2007.06.003

Chiu, M. M., Chow, B. W., & Mcbride-Chang, C. (2007). Universals and specifics in learning strategies: Explaining adolescent mathematics, science, and reading achievement across 34 countries. *Learning and Individual Differences, 17*, 344–365. doi:110.1016/j.lindif.2007.03.007

Engin-Demir, C. (2009). Factors influencing the academic achievement of the Turkish urban poor. *International Journal of Educational Development*, *29*, 17–29.

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10*, 229–240.

Hembree, R. (1990). The nature, effects and relief of math anxiety. *Journal for Research in Mathematics Education, 21*(1), 33–46.

Ho, H., Senturk, D., Lam, A. G., Zimmer, J. M., Hong, S., Okamoto, Y. … Wang, C. (2000). The affective and cognitive dimensions of math anxiety: A crossnational study. *Journal for Research in Mathematics Education, 31*(3), 362–379.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453.

İş Güzel, Ç., & Berberoğlu, G. (2010). Students' affective characteristics and their relation to mathematical literacy measures in the Programme for International Student Assessment (PISA) 2003. *Eurasian Journal of Educational Research, 40*, 93–113.

Jöreskog, K. G. (2002). *Structural equation modeling with ordinal variables using LISREL.* Retrieved from http://www.ssicentral.com/lisrel/ordinal.htm

Jöreskog, K. G., & Sörbom, D. (2006). LISREL 8.80 for Windows. Lincolnwood, IL: Scientific Software International, Inc.

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences*, *19*, 355–365.

Lee, J., & Stankov, L. (2013). Higher-order structure of noncognitive constructs and prediction of PISA 2003 mathematics achievement. *Learning and Individual Differences*, *26*, 119–130.

Lent, R. W., Brown, S. D., & Gore Jr, P. A. (1997). Discriminant and predictive validity of academic self-concept, academic self-efficacy, and mathematics-specific self-efficacy. *Journal of Counseling Psychology*, *44*(3), 307–315.

Liu, O. L. (2009). An investigation of factors affecting gender differences in standardized math performance: Results from US and Hong Kong 15 year olds. *International Journal of Testing*, *9*(3), 215–237.

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education, 30*(5), 520–540.

Ma, X., & Klinger, D. A. (2000). Hierarchical linear modeling of student and school effects on academic achievement. *Canadian Journal of Education*, *24*, 41–55.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320–341.

McCarthy, J. M., & Goffin, R. D. (2005). Selection test anxiety: Exploring tension and fear of failure across the sexes in simulated selection scenarios. *International Journal of Selection and Assessment, 13*(4), 282–295.

McConney, A., & Perry, L. B. (2010). Socioeconomic status, self-efficacy, and mathematics achievement in Australia: A secondary analysis. *Educational Research for Policy and Practice, 9*(2), 77–91.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology, 82*(1), 60–70.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 525–543.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, *56*(2), 177–196.

Morony, S., Kleitman, S., Lee, Y. P., & Stankov, L. (2013). Predicting achievement: Confidence vs self-efficacy, anxiety, and self-concept in Confucian and European countries. *International Journal of Educational Research, 58,* 79–96. doi:10.1016/j.ijer.2012.11.002

Organisation for Economic Co-operation and Development. (2009). *PISA data analysis manual SPSS* (2nd ed.). Paris: Author.

Organisation for Economic Co-operation and Development. (2013a). *PISA 2012 results: Ready to learn: Students' engagement, drive and self-beliefs* (Vol. III). Author. Retrieved from http://dx.doi.org/10.1787/97892642001170-en

Organisation for Economic Co-operation and Development. (2013b). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy.* Author. Retrieved from http://dx.doi.org/10.1787/9789264190511-en

Papanastasiou, C. (2000). Internal and external factors affecting achievement in mathematics: Some findings from TIMSS. *Studies in Educational Evaluation*, *26*, l–7.

Schaubroeck, J., Lam, S. S. K., & Xie, J. L. (2000). Collective efficacy versus self-efficacy in coping responses to stressors and control: A cross-cultural study. *Journal of Applied Psychology, 85*(4), 512–525.

Schumacker, R. E., & Lomax, R. G. (2010). *A Beginner's guide to structural equation modeling* (3rd ed.). New York, NY: Taylor and Francis Group.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research*, *4*, 27–41.

Shen, C., & Pedulla, J. J. (2000). The relationship between students' achievement and their self-perception of competence and rigour of mathematics and science: A cross-national analysis. *Assessment in Education*, *7*(2), 237–253.

Shen, C., & Tam, H. P. (2008). The paradoxical relationship between student achievement and self-perception: A cross-national analysis based on three waves of TIMSS data. *Educational Research and Evaluation: An International Journal on Theory and Practice, 14*(1), 87–100. doi:10.1080/13803610801896653

Sireci, S. G. (1997). Problems and issues in linking assessments across languages. *Educational Measurement: Issues and Practice*, *16*, 12–19. doi:10.1111/j.1745-3992.1997.tb00581.x

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417–453.

Stankov, L. (2010). Unforgiving Confucian culture: A breeding ground for high academic achievement, test anxiety and self-doubt? *Learning and Individual Differences, 20*, 555–563. doi:10.1016/j.lindif.2010.05.003

Wigfield, A., & Meece, J. L. (1988). Math anxiety in elementary and secondary school students. *Journal of Educational Psychology, 80*(2), 210–216.

Wilkins, J. L. (2004). Mathematics and science self-concept: An international investigation. *The Journal of Experimental Education, 72*(4), 331–346.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research and Evaluation*, *12*(3), 1–26.

Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, *31*(2), 114–128.