Zhao, N., Valcke, M., Desoete, A. & Verhaeghe, J.P. (in press). A multilevel analysis on predicting mathematics performance in Chinese primary schools: Implications for practice. Accepted for publication in the *Asia-Pacific Education Researcher*.

A multilevel analysis on predicting mathematics performance in Chinese primary schools: Implications for practice

A multilevel analysis on predicting mathematics performance in Chinese primary schools: Implications for practice

Abstract

The acquisition of mathematical literacy in primary school is a complex process that is influenced by a large set of variables. A multilevel model was applied to identify significant predictors of mathematics performance in Chinese primary schools. Data were obtained from 10,959 students of six grades from primary schools in rural/urban areas, within five provinces with different developmental levels. At the school level, the aggregated socioeconomic status of school is a significant predictor of math performance (x^2=4.3, *df*=1, *p*<.05), until the individual reading level is included. At the class level, grade is a significant predictor. Teacher's graduation level predicts performance (x^2=4.84, *df*=1, *p*=.03), until individual students metacognition level is added. At the student level, reading performance (x^2=434.87, *df*=1, *p*<.00), mathematics self-efficacy (x^2=392.62, *df*=1, *p*<.00) and metacognition (x^2=756.62, *df*=1, *p*<.00) play a large and significant role. Socioeconomic status of family is a weak and polynomial predictor. The results reveal that individual background variables are important predictors and explain 46.67% of the total variance in math performance. After controlling for student characteristics, school and class level variables disappear as predictors, implying an interaction between contextual and individual variables. The present research findings have – next to theoretical implication – also policy implications for Chinese mathematics education: firstly, the educational quality between provinces seem to be balanced, but the school quality within a province does not seem to be balanced. Secondly, there seems to be a need for a quality control related to the output of open teacher training institutions. Thirdly, remedial or intervention programs have to be put in place, to be proactive as to difficulties of students with different language backgrounds.

Key words: Mathematics performance, Multilevel models, Chinese education, Primary school

Introduction

Mathematics is a key component of the primary school curriculum. But, the implementation of mathematics curricula does not automatically lead to an increase in mathematics performance. The latter appears to be the result of a complex interaction between factors related to learner, teaching approaches and the school setting. The literature about mathematics performance describes multiple factors affecting mathematics performance. A primary set is related to the basic capabilities of the learners (Russel & Ginsburg, 1984; Silver, Pennett, Black, Fair, & Balise, 1999). Most of these factors are difficult to be influenced, and resistant to educational interventions. Secondary factors, affecting mathematics performance, are related to (1) individual variables that can be influenced/changed, such as, math anxiety (Meece, Wigfield, & Eccles, 1990), self-efficacy (Pajare & Miller,1997; Pajares & Graham, 1999); (2) background variables, such as family socioeconomic status (SES) related variables (Sirin, 2005), such as home reading and homework support; (3) instructional environment variables, e.g., the nature of the mathematics curriculum, the handbook, the quality of educational interventions, a teacher’s professional status, time to task, the use of didactical tools, … (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). In the context of the present article, we focus on these secondary variables since – from an educational development point of view –primary variables are difficulty to influence or change.

A vast body of empirical evidence is available, studying the single and combined impact of sets of secondary variables. Unfortunately, limited empirical evidence is available in the context of Chinese primary education. Additionally, the available studies - set up in the Chinese context or involving Chinese learners in comparative studies - hardly focus on the complex interplay of the variables at the school, class and student level. In 2001, the *Basic Education Curriculum Reform Programme (Draft) was* issued by the Chinese Ministry of Education (MOE), aiming at improving the quality of education. But little empirical evidence is available to ground the changes in educational policy and to underpin the specific impact of concrete variables on mathematics performance (see e.g., Lim & Zhao, 2005).

This introduces the central research problem of the present study: to develop and test a model to explain mathematics learning performance in Chinese primary schools. The model incorporates the available theoretical and empirical base about the impact of secondary variables. The study builds on mathematics performance data from 10,959 Chinese pupils, from primary schools in five different Chinese provinces with five developmental level. The effects of school, class and student level variables on mathematics performance was analyzed by adopting a multilevel analysis approach. The purpose of the article is: (1) to examine student performance across provinces, schools and classes; (2) to find what variables are significant predictors of mathematics performance.

Theoretical Background

A variety of conceptual frameworks are available, describing and explaining the factors related to mathematics performance (e.g., Byrnes & Miller, 2007; Opdenakker & Van Damme, 2001). A central characteristic of current models is that variables are considered to play a role at different levels. In the next paragraphs we structure these studies along the following levels: school level, class level, and student level.

School level variables

There is plenty of evidence that “schools” matter in terms of mathematics performance. Research points at the impact of school policies, the size and social organization of schools (Bosker, Kremer, & Lugthart, 1990; Opdenakker & Van Damme, 2001; Sammons, Hillamn, & Moretimore, 1995). The latter factors are influenced by the geographical position of schools; this is of particular importance in developing countries such as China. Schools in “rich” regions could develop stronger policies, be better equipped, attract better teachers, … as compared to “poorer” regions. In the context of the present study, this regional position of schools will be operationalized with the variable “province” and the related Gross Domestic Regional Product (GDRP).

A meta-analysis of Bosker and Witziers (1996) shows that up to 18% the variance in academic performance can be attributed to school level variables. On the other hand, it is also necessary to mention that in available multilevel studies, the initial strong impact of school level variables shrinks when class level variables are taken into account (Scheerens & Creemer, 1989).

Also, some family related variables – such as SES – are contingent at the school level (Sirin, 2005; White, 1982). Since families live in a certain province with a certain Gross Domestic Regional Product (GDRP), family SES can be strongly dependent on the former variable. Students are therefore not randomly distributed between/within provinces and schools (Hanushek, Kain & Rivkin, 2004).

Class level variables

Class level variables have consistently been associated with factors affecting academic performance (Teddlie, 1994), such as the quality of the teacher, initial teacher preparation (degree level, professional level), experience with teaching and teaching related beliefs (Goldhaber & Brewer, 2000; Smith, Desimone, & Ueno, 2005). There is some controversy in the research literature whether teaching experience is a valid and significant predictor (see e.g., Kukla-Acevedo, 2009). Though teacher professional development is accepted as a significant predictor of student performance, some studies point at an interaction effect between this teacher characteristic and student related variables, such as learning style (Kukla-Acevedo, 2009). An additional class level variable is linked to the grade level of a class. There are six grade levels in primary school (also labeled “school year” in some countries). Since curricula differ according to the grade, we clearly can expect mathematics performance to be a significant predictor. Studies indeed show that grade is one of the most significant predictors to explain math performance (Kyriakides & Luyten, 2009).

Student level variables

Next to demographic variables (e.g., age, gender, language proficiency, ethnicity), a variety of personal characteristics have been studied that affect mathematics performance; e.g., family variables (e.g., parental involvement, socioeconomic status), student’s beliefs or self-efficacy, metacognition; etc. (Spelke & Ellison 2008; Tate, 1997).

Family variables. A particular set of student level variables is linked to the family setting. While age, gender, ethnicity and other background characteristics are basically related to math performance, family socioeconomic status (SES) supersedes most of the former variables (Fan, 2001; Reyes, & Stanic, 1988). The results suggest that students from low-SES backgrounds are more at risk in view of their mathematics performance (Borman & Overman, 2004; Coleman et al., 1966; Jeynes, 2005). Further meta-analysis studies about the influence of socioeconomic status (SES) on academic achievement – set up between 1990 and 2000 - consistently observed a medium to strong SES–achievement relationship, depending on the unit, the source, or the range of SES variable, and the type of achievement measure included in the studies. Interesting is the observation that the SES-performance relationship is also contingent at the school level, and levels related to a geographical location or region (Sirin, 2005; White, 1982). This suggests that an average family SES level should also be considered as an aggregated variable at other levels in a research model (e.g., at the school level).

SES is not a uni-dimensional variable. A first example of an SES related dimension builds on the educational level of the parents. Research of Alwin and Thornton (1984) has shown how both father’s and mother’s educational level are associated with student performance. In other studies, it is especially mother’s educational level that was found to be the most critical variable (OECD, 2009). Educational level is but one example of SES related variables. As becomes clear from Table 1, different authors and studies put forward a different set of variables to describe the multidimensional nature of the SES variable. In the context of the present study, we will try to integrate a number of SES related variable into a single SES score.

Table 1.

Multi-dimensional perspectives on the SES variable

	Parental education	Parental occupation	Parental income	Home resources	Number of books	Reduced price programme	Neighbourhood	Home atmosphere
Baer (1999)	√	√						
Caldas & Bankston (1997)	√	√	√	√	√	√		√
Duncan, Brooks-Gunn & Klebanov (1994)			√					
Louis, & Zhao (2002)	√		√					
Olson, Martin, & Mullis (Eds.) (2008)	√			√	√			
OECD (2009) for PISA 2006 report	√	√		√				√
Sirin (2005)	√	√	√	√	√	√	√	

Demographic variables. Individual background variables, such as gender, ethnicity, … have been popular when studying mathematics performance (Scarr, 1988; Secada, 1992). Gender was central to the meta-analysis studies of Hyde, Fennema, and Lamon, (1990). Ethnicity was also identified as a relevant but positioned as an indirect predictor of mathematics performance. Lastly, some studies focused on the birth order of the child in the family as it also seems to affect in a significant way math performance (Zajonc & Markus, 1975). A first born child reflects – due to high parental expectations – higher achievement scores.

Individual characteristics. A large variety of individual characteristics have been linked to mathematics performance. We first discuss a range of non-academic variables and switch at the end to an academic variable. A first set of studies centred on the impact of motivational beliefs; e.g., motivation, attributions, regulation, participatory behaviours and engagement, self-concept (Elliott, DiPerna, Mroch, & Lang, 2004). Other researchers focused on the impact of engagement and study skills (DiPerna, Volpe, Elliott, 2002; 2005), motivation (Archer, 1996), confidence and anxiety (Hyde et al. 1990; Vermeer, Boekaerts, & Seegers,. 2000). Self-efficacy was found to play a critical role (Stevens, Olivarez, Lan , & Tallent-Runnels, 2004). Others found that metacognition significantly influences – next to reading, spelling and reading comprehension - mathematics performance (Desoete, Roeyers, & Buysse, 2001).

It was also observed that reading comprehension performance was a predictor of mathematics in primary school; suggesting that student with a higher level of reading comprehension in primary school achieve better (Grimm, 2008).

To summarize, the available theoretical and empirical research puts forward the need to adopt a complex model to describe and explain mathematics performance. Multiple sets of variables should be considered, such as found in the opportunity-propensity framework (Byrnes & Miller, 2007). In contrast to the latter approach in which hierarchical regression and structural equation modeling were adopted, we focus explicitly on the relationship between levels in the predictor variables. In a concrete educational setting, variables are nested into hierarchical levels. Individual learner variables are nested within class variables that are nested with school level variables (Raudenbush & Bryk, 2002). This requires the adoption of multilevel analysis techniques. In the following figure, we integrate the levels and variables as discussed before and included the variables that will be used in our study (See Figure 1).

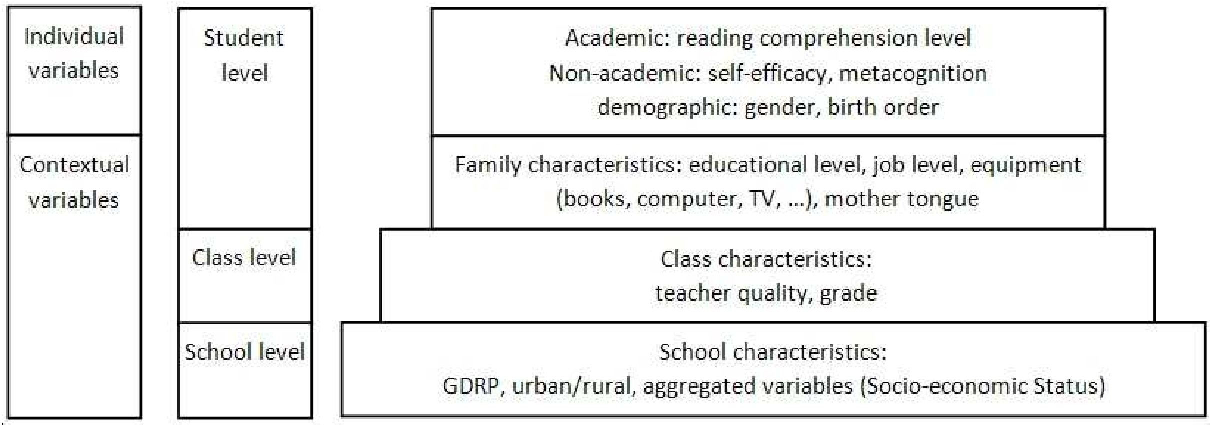

Figure 1. Multilevel model to position individual and contextual variables – a three levels - that are hypothesized to influence mathematics performance.

In a multilevel research, the data structure in the population is considered to be hierarchical, and the sample data are viewed as a multi-stage sample from this hierarchical population (Hox, 2002). Multilevel models assume that students and teachers are not randomly distributed in classes and schools, since they are clustered in classes with particular teachers (Lee & Bryk, 1989). The additional advantage of this approach is that the impact of variables - at the different levels – is studied simultaneously since interactions between the variables are considered. Multilevel modelling is therefore better suited to study real world phenomena that take into account the complex nature and impact of the social context, and the influence of mediating (e.g., instructional) processes and variables (Hox, 2002). As explained earlier, although multilevel studies focusing on mathematics performance are available in the literature (Opdenakker & Van Damme, 2001), thus far no extensive and large scale studies have been set up in the Chinese context. The present study aims at developing a baseline for this type of research.

Methods

Objectives

The present study aims at examining student mathematics performance across provinces, schools and classes. This will help to identify school level, class level and student level variables that significantly predict mathematics performance.

Research sample

Data for this study were obtained from 10,959 primary school pupils. A purpose-driven multi-stage stratification sampling approach was adopted. First, provinces were selected on the base of their GDRP level (6 levels).

In China, thirty one administrative provinces can be distinguished (excluding the Special Administrative Regions). Determination of the GDRP level was based on the 2005 report of the Chinese Economic Bureau (Level 1 = highest level). Given the very different way education is organized in the poorest provinces (e.g., multilingual education), no data were obtained from these provinces.

Second, sampling was based on the location of schools in a rural or urbanized area (labeled "region"). Thirdly, two schools were – with the support of the Educational Bureau - approached within each regional area to be involved in the study. The size of the selected schools ranged from 318 to 897 students (*M*=547.95, *SD*=140.19).

Thirdly, one class of each grade level in a school was randomly selected to be involved in the study. A minimum of 50 participants per grade per school was set forward. In cases the number of pupils in a class was lower than 50, a second complete class of the same grade was selected from this school. This implies that in total x classes were involved in the study.

Next to the pupils, also their classroom teachers were involved in the study. Of the 197 teachers, 73.4% were female. In some schools, the same teacher was responsible for two classes.

Table 2

Sample characteristics (N=10,959)

		Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Total	Percent
GDP	Level 1	497	548	498	502	488	490	3023	27.58%
	Level 2	293	272	282	279	255	270	1651	15.07%
	Level 3	349	367	364	363	369	367	2179	19.88%
	Level 4	320	282	271	309	248	348	1778	16.22%
	Level 5	398	386	390	380	389	385	2328	21.24%
Total		**1857**	**1855**	**1805**	**1833**	**1749**	**1860**	**10959**	**100%**
Region	Urban	988	992	931	995	913	1005	5824	53.14%
	Rural	869	863	874	838	836	855	5135	46.86%
Total		**1857**	**1855**	**1805**	**1833**	**1749**	**1860**	**10959**	**100%**

Research variables

Dependent variable Mathematics performance (MATH). The test used in the present study covers the three general mathematics domains that reflect the recent Chinese curriculum standards: number and algebra, shape and space, statistics and probability (MOE, 2001). The test items cover an series of mathematical building blocks: number reading skills, mathematical lexicon, knowledge, procedural knowledge, linguistic skills, mental representation, contextual skills, selecting relevant information, number sense skill, memory skills, visualization or mental representation skills and logical thinking (Desoete & Roeyers, 2005; Nunes et al, 2007; Zimmermann & Cunningham, 1991). Initial test versions were designed by Chinese mathematics teachers and mathematics experts, covering the curriculum for the six school grades. Next, the validity of the pool of test items was assessed by curriculum experts. At each grade level, students were presented with a test consisting of a core set of grade specific items and a set of anchor items. The latter overlap with items at an earlier or the next grade level and help to calibrate the complete set of items at primary school level. A pilot study and a main study were set up to test and calibrate the entire mathematics performance test. All the items and students from all grades were calibrated on the same continuum scale by Item Response Theory (See Figure 2). The internal consistency (Cronbach's α) of each grade level test ranges from 0.93 to 0.96. The test also helps to determine student abilities that range between -5.30 to 3.30 (M = .57, SE = .26). The test characteristics are summarized in Table 3 (See Authors, 2008).

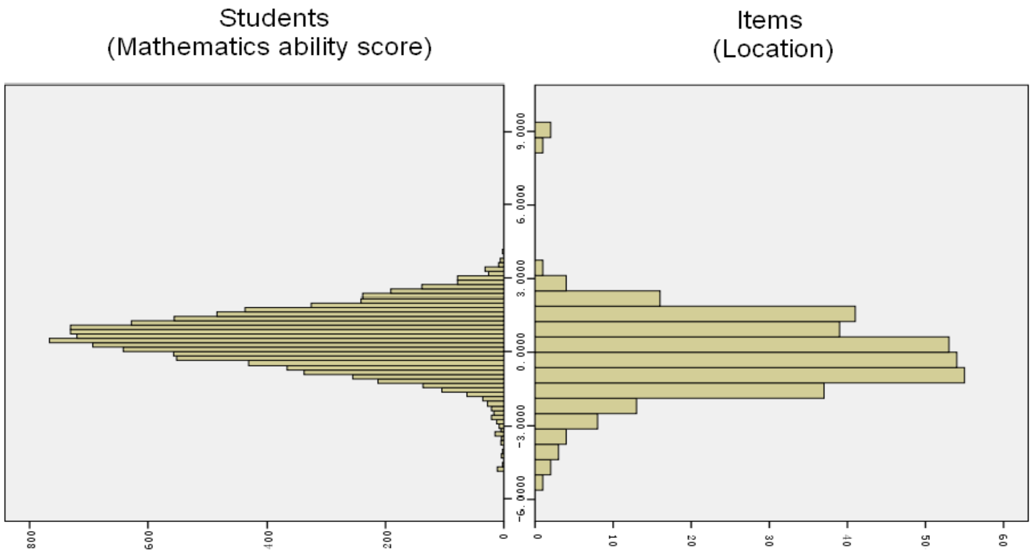

Figure 2. Distribution of students' ability estimations and item locations.

Table 3.

Mathematics test - IRT characteristics

Grade level	Reliability	Slope *Mean (SD)*	Location *Mean (SD)*
Grade 1	.94	.72 (.26)	-1.24 (1.29)
Grade 2	.96	.65 (.19)	-.89 (1.18)
Grade 3	.95	.63 (.21)	.05 (2.07)
Grade 4	.94	.74 (.57)	.18 (1.22)
Grade 5	.94	.67 (.19)	.70 (1.24)
Grade 6	.93	.88 (.34)	.83 (1.00)

Predictor variables at the school level.

GDP_P: Due to clear differences in social and economical development between Chinese provinces, the GDP_P variable was constructed on the base of the 2005 Gross Domestic Regional Product (GDRP) classification scheme of the Chinese Economic Bureau. This resulted in five GDRP levels ranging from 1 (highest GDRP level) to 5 (lowest GDRP level). As explained earlier, no schools form the lowest GDRP province levels were incorporated in the study.

MUR*:* A dichotomous variable was applied to refer to the urban (code 1) or rural location (code 0) of each school.

Aggregated variables were added to the school level to explore compositional effects (Raudenbush & Bryk, 2002):

SCFSES_J: This aggregated variable is based on the school mean of the students' socioeconomic status as reflected in their parents job.

SCFSES_E: This aggregated variable builds on the school means of the students' socioeconomic status as derived from their equipment level.

Predictor variables at the class level.

Grade: Grade level ranged from Grade 1 to Grade 6 in primary school. Grade 1 is the first schooling year in primary school. Students are typically 6 years old when they start in grade 1.

Teacher's Graduation School (TGRA): This variable checks the type of teacher education institute attended by the teacher. Two categories are considered: 0 for lower level teacher training institution (such as open university, distance education or independent learning system), 1 for higher level school (such as a normal university or teacher training college). The latter is considered to be the formal and predominant approach to become a teacher.

Predictor variables at the student level.

The students were asked to finish a survey focusing on the following variables. The family background information of the students was gathered from the school principal or the classroom teacher.

Family background variables.

Socio-economical status (FSES): Based on the overview in table 1, eight variables were used to determine the level of Socio-economic Status: educational level of father/mother, job level of father/mother based on a classification derived from Li (2005a; 2005b), and four variables related to wealth, income, and cultural possessions. Exploratory factor analysis (EFA) on the eight FSES related variables suggested a two factor solution, excluding father/mother educational levels. As a result, these two variables were kept separate. The results of the EFA were corroborated on the base of a confirmatory factor analysis (CFA). Satisfactory goodness-of-fit indexes are observed (x^2=265.80, *df*=8, *p* <.00, *CFI*=.98, *TLI*=.96, *RMSEA*=.048). This implies that two subscales can be distinguished in relation to this composite SES variable:

- *Family SES - Equipment (FSES_E):* This factor score reflects the number of wealth related equipments, present at home (television, refrigerate, washing machine, computer).
- *Family SES - Job (FSES_J):* Students reported their parents' jobs, that was subsequently coded into 27 levels (1=highest level) according to the Chinese classification scheme of Li (2005a, 2005b).

Since the father's and mother's educational level were excluded from the factor analysis, we retain them in our study as separate indicators of the family background.

Father's educational level (FEL) and mother's educational level (MEL): educational level was classified as: no schooling experience (code 1), primary school graduate (code 2), junior school graduate (code 3), senior school graduate (code 4), Pre-high school graduate (code 5), high school graduate (code 6), or postgraduate education or higher (code 7).

Language Background (LAN): Considering the importance of the mother tongue and how it related to the school language, pupils whose mother tongue equaled the school language received code 1, while pupils with a divergent mother tongue received code 0.

Student characteristics

Gender (SGender) was coded as 0 for girls and 1 for boys.

Birth order of the child in the family (ORCH): The birth order was used to give a specific value (first child or only child= 1).

Student mathematics self-efficacy (MSS): In order to control the influence of self-efficacy, a Likert-5-point mathematics self-efficacy scale (MSS; Marat, 2005) was administered. On the base of a pilot test, the original 85 item scale was reduced to 78 items. A confirmatory factor analysis reflected high goodness-of-fit indices (x^2= 41654.53, *df*=2818, *p* <.00, *GFI*=.91, *CFI*=.99, *RMSEA*=.04) and high reliability values (Cronbach's alpha .97).

*Metacogniton (META)**:*** Following the 'post diction paradigm' - students were invited to predict the level of their test performance (e.g., 'I think I will obtain 70/100 on this test'). A calibration-index was used to assess the difference between the actual performance score and the estimated mathematics performance score (see e.g., Desoete & Roeyers, 2006). The equation used to calculate META is: $\mathrm{META} = (\mathrm{Expected\ score} - \mathrm{Actual\ score})^2$. This equation implies that when the META score is not equal to 0, students' metacognition ability is lower.

Chinese language performance (CHI). To determine the mastery of the Chinese language, the mid-term examination or end-term examination test scores were used in the present study. Since it is difficult to compare these scores across different classes within the same school, the test scores were recoded on the base of the percentage of correct answers. Furthermore, differences between schools were also cancelled out by recoding the test scores into a categorical variable; five achievement levels: 1 for highest, 2 for high, 3 for average, 4 for low, 5 for the lowest achievement level.

Data Analysis

Multilevel analysis (MLWin) was applied to evaluate to what degree individual student variables, class variables and school variables influenced mathematics performance (Rasbash, Steele, Browne, & Prosser, 2004). A three-level hierarchical linear model was used to examine the independent association between the student-level variables, class-level variables and school-level variables and mathematics performance.

In a first step, the null model is tested, aiming at detecting significant differences in mathematics performance between schools; without predictors being considered. In the next models to be tested, school level, class level variables and finally individual student level variables are added to the model. Initially, all variables are included in the model as fixed effects, assuming that their impact did not vary from student to student or from class to class. The multilevel analysis specification equation is to be written as:

$$Y_{ijk}=\beta_{0ijk}+ \beta_{1i} X_{1i}+ \beta_{2ij} X_{2ij} + \beta_{2ijk} X_{2ijk} \ldots\ldots$$

where, $\beta_{0ijk} = \beta_0 + v_{0k} + u_{0jk} + e_{0ijk}$

- β_0 is the grand mean of math performance across all students, classes and schools;
- v_{0k} is the random effect at the school level, allowed-to-vary departure from the grand mean;
- u_{0jk} is the random effect at the class level, allowed-to-vary departure from the school effect;
- e_{0ijk} is the random effect at the student level, allowed-to-vary departure from the class effect within a school;
- β_{1i} is the coefficient of school-level variable X_{1i} ;
- β_{2ij} is the coefficient of class-level variable X_{2ij} ;
- β_{3ij} is the coefficient of student-level variable X_{3ijk} .

After testing the model with fixed effects, a second test was carried out in which parameter coefficients of the variables were allowed to vary randomly across schools, classes within a school and students within a class. Finally, testing the full model implies the full set of predictor variables being entered in the random-coefficient regression analysis. In the latter case, the within-school individual learner variables are considered fixed within each school, but the Level-1 slope for mathematics performance is considered to vary randomly between schools. This implies that the effect of student ability and family background on mathematics performance varies across schools.

Results

Descriptives

The descriptive analysis centers on testing bivariate correlations by calculating Kendall's tau between all variables at the different levels in the analysis. Table 4 summarizes the correlation coefficients, means and standard deviations of variables included in the final model. The analysis results help to conclude that all variables can be included in the multilevel analysis .

Table 4.

Correlations [a], Means and Standard Deviation of variables (N=10,959)

Variables [b]	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) MATH	-											
(2) SCFSES_J	.16**	-										
(3) GRADE	.36**	-.01	-									
(4) TGRA	.13**	-.19**	.05**	-								
(5) FSES_J	-.10**	.47**	-.00	-.12**	-							
(6) FSES_E	.10**	.49**	-.00	.13**	-.62**	-						
(7) LAN	.03**	.13**	.04**	-.06**	.12**	-.16**	-					
(8) GENDER	-.02	.01	-.00	.00	.02	-.01	.02*	-				
(9) ORCH	.02*	.18**	.03**	.01	.13**	-.16**	.10**	.06**	-			
(10) MSS	.20**	-.16**	.14**	.04**	-.12**	.14**	-.15**	-.05**	-.04**	-		
(11) META	-.48**	.10**	.15**	-.12**	-.06**	-.04**	.00	.03**	.03**	-.05**	-	
(12) CHI	.13**	-.27**	-.01	.00	-.14**	.14**	-.13**	-.09*	-.14**	.23**	-.06**	-
M	.57	.08	3.49	.75	.07	-.01	.38	.52	1.32	296.89	.09	3.23
SD	1.11	1.38	1.72	.43	2.83	.32	.49	.50	.63	44.02	.11	.92

[a]** p: p<.01; *p: .01<p<.05

[b] MATH=mathematics performance, SCFESE_J=means of family SES at the school level, GRADE=grade, TGRA=teacher graduation school level, FSES_J= family SES based on job level, FSES_E= family SES based on wealthy possessions, LAN=mother tongue is Chinese or not, GENDER=student gender, ORCH=birth order in family, MSS = Mathematics self-efficacy, META= metacognition, CHI= Chinese language performance.

Multilevel Analysis Results

Null model. Table 5, 6 and 7 summarize the results of the consecutive model testing steps, by using an iterative generalized least squares (IGLS) estimation procedure. Model 0 is a fully unconditional three-level null model (See Table 5) without the inclusion of specific predictors. The intercept of .56 in this model indicates the estimated overall school average in mathematics performance of all students in all schools. The total variance is further decomposed into between-school, between-class and between student variance. The random part of the null model reveals that the variance at student level, class level and school level is significantly different from zero. School-level factors account for 18.55% of the overall variance in mathematics performance. The largest proportion of the variance (41.94%) is related to differences between classes within schools. And 39.52% of the variance is attributed to differences between students within classes within schools.

The analysis of the three-level null model reveals that the differences between students in mathematics performance far outweigh the differences between groups (schools, classes). Nevertheless, differences linked to school and class variables seem to be sufficiently important to explain variance in mathematics performance. These findings are in line with the results of the multilevel study of Opdenakker, Van Damme, De Fraine, Van Landeghem and Onghena (2002) who claimed that school effects play a significant role in learner performance.

Hierarchical model testing. Step-by-step, variables were added to the null model at the school level, class level and student level. All predictor variables were first centered around the grand mean at their corresponding level before being added to the model (means = 0). Since parsimonious models are preferred, only significant predictors and ameliorated models have been reported in the table.

School level variables. Firstly, at the school level, the provincial development level of the school (GDP_P) and the rural/urban location of the school (MUR) were entered into the model. This did not result in a significant improvement of the model. GDP_P and MUR do not appear to be significant predictors of mathematics performance. As stated earlier, some aggregated variables were added to the school level to explore compositional effects of socio-economic status. School average SES, based on the job level of the parents, seems to play a significant role. SCSES_J coefficient varies from -.15 to -.16, implying that a lower school mean of the student's parents job position, student performance will be lower. As will be discussed later, this SES-related contribution is overruled by student level variables (See model 5b).

Class level variables. Secondly, grade is added as a class level variable (see model 2 in Table 5). The average estimate of the overall school mathematics performance is 0, considering grade one is used as the reference category. Compared to students in grade one, students in grade two attain the same performance level as grade one. But students in grade three, four, five and six attain significantly higher mathematics performance levels (respectively .46, .62, 1.00, 1.09) when compared to grade one. Next, the characteristics of the teachers were added to the model. Teacher's graduation school level (TGRA) was found to result in a significant improvement in model 3 (x^2=4.84, *df*=1, *p*=.03). As will be explained later, the impact of this teacher variable is overruled by student level variables (See model 8a).

Table 5.
Multilevel regression analysis of school and class level variables, including the Null model

Predictor	M0	M1	M2	M3
FIXED				
Intercept	.56 (.12)	.54 (.11)	.01 (.13)	-15 (.15)
SCFSES_J		-.16 (.07)	-.16 (.07)	-.15 (.07)
Grade2			.11 (.13)	.14 (.13)
Grade3			.43 (.13)	.46 (.13)
Grade4			.59 (.13)	.62 (.13)
Grade5			.99 (.13)	1.00 (.13)
Grade6			1.10 (.13)	1.09 (.12)
TGRA				.21 (.10)
Random Part				
Level: School				
σ^2	.23 (.09)	.18 (.07)	.19 (.07)	.19 (.07)
Level: Class				
σ^2	.52 (.05)	.52 (.05)	.33 (.03)	.33 (.03)
Level: Student				
σ^2	.49 (.01)	.49 (.01)	.49 (.01)	.49 (.01)
-2LL	24186.83	24182.53	24083.63	24078.79
x^2		4.3	98.9	4.84
df		1	5	1
p		.04	<.001	.03
Reference		M0	M1	M2

* *Note*: Values in parentheses are standard error

Student level variables. Thirdly, student level variables were included in the model (See Table 6).

In a first step, student family background characteristics were entered in the model 4 and 5. We observe that the quadric of FSES_J (x^2=11.49, *df*=1, *p*<.001) is a better predictor than FSES_J (x^2=4.84, *df*=1, *p*=.03), implying a polynomial relationship between job-based family SES and mathematics performance. But at the same time, the SCFSES_J (aggregated school level variable) has no longer a significant impact (See model 4b). Adding FSES_E (SES based on equipment indicators), also results in a significant improvement of the model (x^2=12.91, *df*=1, *p*<.001). This finding is in line with Willms's (2003) hypothesis that there is a non-linear relationship between SES and academic performance.

The analysis results further indicate that the language spoken at home LAN

(x^2=10.48, *df*=1, p<.005) is a critical predictor for mathematics performance. Children of families not speaking the school language achieve .30 units lower in mathematics performance. By allowing LAN, FSES_E, FSES_J to vary randomly across schools and classes, no significant improvement of the model is observed. This means that within the same school and same class, the mathematics performance hardly varies for students with a comparable family socioeconomic status level.

Table 6.
Multilevel regression analysis of student level variables (family background variables)

Predictor	M4a	M4b	M4c	M4d	M4e
FIXED					
Intercept	-.15 (.15)	-.15 (.16)	-.17 (.15)	-.17 (.16)	-.14 (.16)
SCFSES_J	-.141 (.073)				
Grade2	.14 (.13)	.14 (.13)	.14 (.13)	.14 (.13)	.14 (.13)
Grade3	.46 (.13)	.46 (.13)	.45 (.13)	.45 (.13)	.45 (.13)
Grade4	.62 (.13)	.62 (.13)	.62 (.13)	.62 (.13)	.62 (.13)
Grade5	1.00 (.13)	1.00 (.13)	.99 (.13)	.99 (.13)	.99 (.13)
Grade6	1.09 (.12)	1.09 (.12)	1.08 (.12)	1.08 (.12)	1.08 (.12)
TGRA	.20 (.10)	.22 (.10)	.22 (.10)	.22 (.10)	.22 (.10)
FSES_J	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)
FSES_J* FSES_J			.002 (.001)	.002 (.001)	.002 (.001)
FSES_E				-.09 (.03)	-.09 (.03)
LAN					-.06 (.02)
Random Part					
Level: School					
σ^2_{v0}	.19 (.07)	.23 (.08)	.23 (.08)	.23 (.08)	.23 (.08)
Level: Class					
σ^2_{u0}	.33 (.03)	.33 (.03)	.32 (.03)	.33 (.03)	.33 (.03)
Level: Student					
σ^2_{e0}	.49 (.01)	.49 (.01)	.49 (.01)	.48 (.01)	.48 (.01)
-2LL	24067.30	24070.72	24057.81	24049.42	24038.94
x^2	11.49	+3.42	12.91	8.39	10.48
df	+	+1	1	1	1
p	<.001	.06	<.001	<.005	<.005
Reference	M3a	M4a	M4b	M4c	M4d

* *Note*: Values in parentheses are standard errors

In a second step, demographic student variables were added to the model. Only a few variables were found to play a significant role. As can be derived from Table 5, the inclusion of gender results in a significant improvement in model 5 (x^2=15.58, *df*=1, *p*<.001). The results show that girls significantly outperform boys in mathematics. Adding birth order of the child, we notice an improvement in model 6 (x^2=13.16, *df*=1, *p*<.005). According to the Chinese one-child-per-family policy of 1979, the order of the child was centered by one. Considering the order of a pupil in the family, mathematics performance will decrease by .03 units for a next position in birth order. We can explain this by pointing at additional responsibilities that have to be adopted by first children, and higher parent expectations in relation to the first child, that result in higher school performance.

As a third step, other student characteristics were entered into the model. The difference in deviance between the consecutive models is statistically significant. When self-efficacy is added to the model (model 7), this results in a significant improvement (x^2=392.62, *df*=1, *p*<.001). Lastly, in model 8a, metacognition is added to the model. This seems to be a very important predictor of mathematics performance. The deviance of the model 8a compared to model 7 is 3411.89 (*df*=1, *p*<.001). An increase of one unit in the metacognition score (META), results in an decrease of 4.13 units in mathematics performance (0 refers to a high metacognition ability). It is critical to point out that by adding this student level variable, the coefficient of the teachers' graduation school and student gender is no longer significant. Therefore, teacher's graduation school level is excluded in model 8a (x^2=1.68, *df*=1, *p*=.19) and gender is excluded in model 8c (x^2=0.40, *df*=1, *p*=.53). When META is allowed to vary across the school level and class level (model 8d and 8e), there is again an improvement in the model (x^2=248.49, *df*=1, *p*<.001; x^2=756.62, *df*=1, *p*<.001).

In a fourth and final step, the Chinese language performance (CHI) was added to the model. CHI seems to contribute in a significant way to mathematics performance in model 9 (x^2=434.87, *df*=1, *p*<.001). The IGLS deviance drops consistently. This result implies that a higher Chinese language mastery level is related to higher mathematics performance scores.

Table 7. Multilevel regression analysis of student level variables (student characteristics)

Predictor	M5	M6	M7	M8a	M8b	M8c	M8d	M8e	M9	ES
FIXED										
Intercept	-.14 (.16)	-.14 (.16)	-.10 (.16)	-.15 (.13)	-.15 (.11)	-.15 (.10)	-.18 (.11)	-.18 (.10)	-.19 (.10)	-
Grade2	.14 (.13)	.14 (.10)	.10 (.13)	.27 (.10)	.25 (.10)	.25 (.10)	.25 (.10)	.25 (.10)	.24 (.10)	.22
Grade3	.45 (.13)	.46 (.13)	.41 (.13)	.50 (.10)	.48 (.10)	.48 (.10)	.49 (.10)	.53 (.10)	.53 (.10)	.48
Grade4	.62 (.13)	.63 (.13)	.56 (.13)	.74 (.10)	.73 (.10)	.73 (.10)	.74 (.10)	.77 (.10)	.80 (.10)	.72
Grade5	.99 (.13)	.99 (.13)	.92 (.13)	1.17 (.10)	1.17 (.10)	1.17 (.10)	1.18 (.10)	1.24 (.10)	1.25 (.10)	1.13
Grade6	1.08 (.13)	1.08 (.12)	1.01 (.10)	1.24 (.10)	1.25 (.10)	1.25 (.10)	1.25 (.10)	1.34 (.10)	1.33 (.10)	1.2
TGRA	.22 (.10)	.22 (.10)	.22 (.10)	.10 (.08)						
FSES_J	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.01 (.00)	-.03
FSES_J* FSES_J	.002 (.001)	.002 (.001)	.002 (.001)	.002 (.001)	.002 (.001)	.002 (.001)	.002 (.001)	.002 (.001)	.001 (.000)	.01
FSES_E	-.09 (.03)	-.09 (.03)	-.12 (.03)	-.09 (.03)	-.09 (.03)	-.09 (.03)	-.09 (.02)	-.09 (.02)	-.09 (.02)	.03
LAN	-.06 (.02)	-.05 (.02)	-.03 (.02)	-.04 (.02)	-.04 (.02)	-.04 (.02)	-.04 (.01)	-.04 (.01)	-.03 (.01)	-.03
SGENDER	-.05 (.01)	-.05 (.01)	-.04 (.01)	-.007 (.011)	-.007 (.011)					
ORCH		-.04 (.01)	-.04 (.01)	-.05 (.01)	-.05 (.01)	-.05 (.01)	-.04 (.01)	-.04 (.01)	-.03 (.01)	-.02
MSS			.004 (.000)	.003 (.000)	.003 (.000)	.003 (.000)	.003 (.000)	.003 (.000)	.003 (.000)	.12
META				-4.13 (.07)	-4.13 (.07)	-4.13 (.07)	-4.31 (.29)	-5.37 (.38)	-5.17 (.39)	-.51
CHI									.17 (.01)	.14
Random Part										
Level: School										
σ^2_{v0}	.23 (.08)	.24 (.05)	.22 (.08)	.14 (.05)	.14 (.05)	.14 (.05)	.14 (.04)	.11 (.04)	.11 (.04)	
$\sigma_{v0,META}$							-.30 (.13)	-.50 (.17)	-.50 (.18)	
σ^2_{META}							1.56 (.52)	2.12 (.92)	2.31 (.97)	
Level: Class										
σ^2_{u0}	.33 (.03)	.33 (.02)	.33 (.02)	.20 (.02)	.20 (.02)	.20 (.02)	.20 (.02)	.20 (.02)	.20 (.02)	
$\sigma_{u0,META}$								-.10 (.10)	-.10 (.10)	
σ^2_{META}								7.99 (.91)	7.40 (.86)	
Level: Student										
σ^2_{e0}	.48 (.01)	.48 (.01)	.47 (.01)	.34 (.01)	.34 (.01)	.34 (.01)	.33 (.00)	.28 (.00)	.28 (.00)	

-2LL	24023.36	24010.20	23617.58	20205.69	20207.37	20207.77	19962.28	19205.68	18770.81
x^2	15.58	13.16	392.62	3411.89	1.68	0.40	248.49	756.6	434.87
df	1	1	1	1	+1	+1	2	2	1
p	<.001	<.001	<.001	<.001	.19	.53	<.001	<.001	<.001
Reference	M4e	M5	M6	M7	M8a	M8b	M8c	M8d	M8e

* *Note*: Values in parentheses are standard errors

$\sigma_{u0*uMETA}$ and $\sigma_{v0*vMETA}$ represent the coefficients of the covariance between the random part and intercepts and slopes of metacognition variable in school level and class level.

Full model. In the full model, individual student characteristics have a dominant impact on mathematics performance. In order to develop a better understanding of the results, effect sizes were calculated and added to the output in Table 7. Next to the class level variable “grade”, metacognition (META, ES=-.51) has the strongest impact on mathematics performance in primary school as compared to other variables in the model. Next, self-efficacy (MSS, ES=.12), and Chinese language mastery (CHI, ES=.14) strongly affect mathematics performance. Family SES background and other variables do not seem to play an important role; though they still should be considered as relevant background variables.

Comparing the null model with the full model, a lower school level variance coefficient is observed (.23 vs. .11). But, this still explains 52.17% of the variance in mathematics performance between schools (See Figure 3). But focusing on the class level, a large change in proportion of explained variance (.52 vs. 20) is observed. This implies these variables help to explain 61.54% in mathematics performance between classes (See Figure 3). Also in figure 3, and focusing on the student level, a large proportion of variance can be explained (.49 vs. .28 in Table 6). This represents a 42.86% of explained variance between students in mathematics performance.

Our findings are largely in line with the results obtained in other developing countries (Ma, 1997). Variables at the individual learner level explain the largest proportion of variance in mathematics performance; even after the school level and class level variables have been controlled for. Remarkably, individual learner characteristics do not only explain differences between learners within classes, but also 16.14% of the remaining variance at the school level and 28.52% at the class level (See Figure 3).

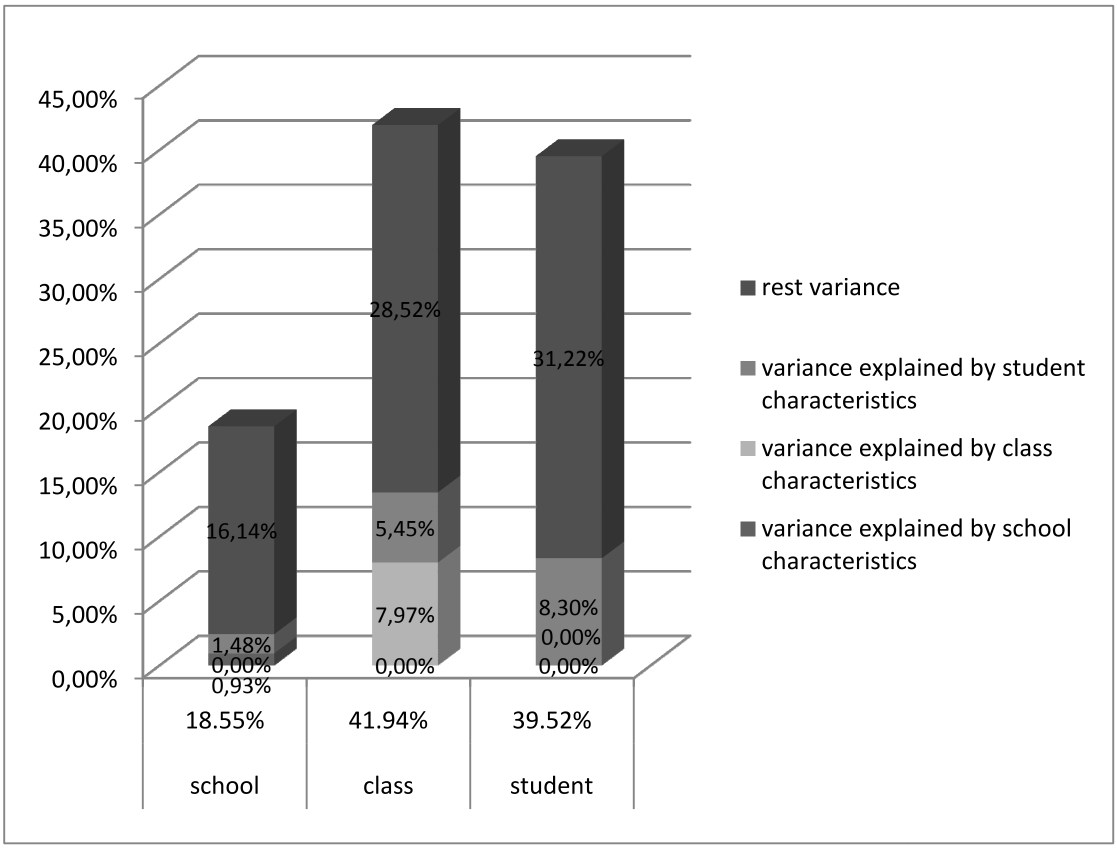

Figure 3. Proportion of explained variance in mathematics performance, as explained by significant school, class and individual student level characteristics.

Discussion, implications and conclusions

School level effects

During recent years, studies from developing countries center on a controversy between the findings of the Coleman report and the Heyneman-Loxley effect (Huang, 2010). Coleman stressed that mainly SES variables account for academic performance. In contrast, Heyneman and Loxley (1983) argued that in countries with different development levels, the relationship between SES and performance might be different. In low-income countries, it is expected that school effectiveness accounts for a larger proportion of the variance in academic performance. Differences in GDP_P and MUR are linked to such larger socio-economical differences. Gross Domestic Product (GDP_P) of the province and the urban or rural location of the schools (MUR) is therefore considered to explain differences in mathematics performance. But – in our study - both GDP_P and MUR do not appear to be significant predictors. This implies that – in the Chinese context - wherever the school is located, individual students still are able to attain high performance levels in mathematics. This can also interpreted that the education quality in different provinces is balanced, but that the school quality within a province is not balanced. As will be explained below, this can be confirmed in the Chinese context.

Adding – in an aggregated way – socio-economic variables at the school level shows that the job of parents still plays a certain role. But, when we allow socio-economical variables to vary across students within classes, mathematics performance is hardly different. Initially, the analysis results derived from model 1 reflect the so-called “compositional effect” of SES related variables (Van Damme, De Fraine, Van Landeghem, Opdenakker, & Onghena, 2002; Van Ewijk and Sleeger, 2010). Also in India, the same result is found. The school population composition reflects specific SES-levels and affects the performance level of this schools within a certain district (Venkatanarayana, 2005). Our results can easily be linked to particularities and the history of the Chinese educational system. Primary school performance differs within the same province, but hardly between provinces. In the Chinese context, these strong between-school differences can be explained by referring to the *Provincial Key School* policy of the late 1970s and the early 1980s (Organization of Educational Yearbook in China, 1984). After the Cultural Revolution, in view of making efficient use of the limited resources available, the Ministry of Education decided to invest most of the available resources in a limited number of schools in each region, resulting in a higher school performance. Students who wanted to enroll in these schools were selected on the base of an entrance examination. This resulted in a situation where – within the same province - next to high performing school, there are middle and low performing schools. Though the Ministry of Education banned the practice of "key schools" in the mid-1990s on the base of an unfair distribution of educational resources, unequal school performance is still a reality. In each province, despite difference in GDRP, there are still “key schools” that keep the performance of the province in balance.

Class level effects

The analysis results show that grades are a relevant class level predictor of mathematics performance. This is expected given the systematic impact of grade related school curricula. Nevertheless, we observe differences in the impact of the grade variable. It is interesting to compare the difference between grade one (reference category) and grade five with the difference between grade one and grade six. From this comparison we learn that the curriculum at each grade level might impose other challenges for students, resulting in a more difficult or less difficult curriculum. This result is in line with observations from previous studies (Basang, 2006; Liu, 2007).

As to teacher related classroom variables, we observe that only the quality of teacher preparation is important (model 3). Teachers graduating from the formal educational system are better able to help students attain high achievement levels, as compared to teachers from alternative teacher training system (e.g., open university teacher education). Teacher training has been a weak component of the Chinese educational system till the end of the Cultural Revolution (end 1970s). By 1985, still two-fifth or the primary teachers had not received appropriate pre-service teacher training (DPSEDC, 1986). This urged the authorities to upgrade these unqualified teachers and also to train more teachers to be appointed in rural areas. This resulted in the late 1980s in an alternative program delivered through the China Television Teachers' College and other types of open universities. The present research results reveal that teachers graduating from these institutions, attain lower performance levels in their students. This suggests that a specific in-service teacher training program should be put in place to close the gap between teachers graduating from formal and non-formal training systems. The results also suggest that there is a need for a quality control system related to the output of open teacher training institutions and expand teacher education at university level (Zeng, 2008).

Student level effects

At student level, the Chinese language attainment level is a critical predictor of mathematics performance. This result is in line with the findings of many authors, such as Dirks, Spyer, and Van Lieshout (2008). Gersten and their colleagues (Gersten, Jordan, Flojo, 2005) who also revealed a significant relationship between language (e.g., reading comprehension) and mathematics performance. This result suggests that remedial or intervention programs have to be put in place to be proactive as to difficulties of students with a different language background. This is regularly mentioned in relation to the learning outcomes within multilingual contexts (Pretorius & Currin, 2010).

Metacognition and self-efficacy predict mathematics performance in a significant way. In the present study, metacognition is even the most important predictor. In model 8a, as compared to model 7, 36.37% of the variance at the student level was explained by metacognition. This is in line with the studies of Efklides (2006) and Veenman, Van Hout-Wolters and Afflerbach (2006). Some authors state in this context that weak mathematics learners suffer a dual burden, they make many mistakes and at the same time they are less able to build on metacognitive competences that might help them to monitor and evaluate their own performance (Kruger, 2002). Also, mathematics self-efficacy was found to influence academic performance. This reiterates the findings of many other studies (Bandura, 1977, 1997; Parjares & Miller, 1997).

The impact of students variables related to their family background presents an interesting picture. The effect of the socio-economic status of the students' family is not as high as reflected in the literature (Sirin, 2005; White, 1982). Our results suggest that the relationship between SES and mathematics performance is more complex than often explained. This should therefore be explored more deeply in future studies.

Summary, limitations and directions for further research

This study contributes to the limited understanding of the predictors for the mathematics performance in Chinese primary education. Strengths of our study are: a large number of students were involved in the study (N 10,959); the study took into account important geographical differences (GDRP, urban/rural areas); third, state-of-the-art mathematics tests were used to determine mathematics performance (application of IRT, aligned with the new curriculum, and covering the full range of mathematical building blocks); lastly a multilevel analysis perspective was adopted to analyze the impact of predictors at three levels (school, class, student) and to explore the interaction between variables.

Nevertheless, some limitations have to be stressed. First, the range of variables, considered at the school level and class level, might still have been insufficient to reflect their full impact. For example, we could have – additionally – considered the variables school culture or school atmosphere (Heck, 2007; Philips, 1997; Mackenzie, 1983). Other interesting variables can be derived from school policies (Lashway, 2002), educational beliefs of principals and teacher beliefs (Ross & Gray, 2006), etc. Second, though our sampling approach helped to develop a stratified sample, due to circumstances no data were gathered from the poorest provinces. Third, our multilevel analysis approach only focused on mathematics performance as the dependent variable. Also, it is less clear along what path the variables – at the different levels - have a direct or indirect impact on mathematics performance. Path analysis can be adopted to test the how the predictor variables are interrelated and how some play a mediating or interaction role, next to studying a causal relationship with mathematics performance. Fourth, additional research is needed to study whether the impact of predictors in the full model remains the same when we consider different levels of mathematics performance (e.g., when focusing on weak performers) or when we focus on a specific developmental level of a school.

In future research, attention could be paid to additional variables at the school and class level: school leadership, didactical approaches, handbooks used, etc. Also, the compositional effect of aggregated variables could be explored more deeply. Additionally, next to studying mathematics performance on the base of a quantitative test, alternative research designs should be adopted to corroborate the findings of the present study or to develop rich measures related to school, teacher or student related variables. For instance, video-based analysis could help to analyze teacher quality, student engagement, teacher and student beliefs, etc. Lastly, the results of the multi-level analysis only present a basic picture of what variables affect at different levels mathematics performance. As suggested path analysis can be adopted to study the interesting interplay between the research variables and to develop a fuller understanding of the nature and quality of mathematics teaching and learning in the Chinese primary education context.

References

Alwin, D. F.,& Thornton, A. (1984). Family origins and the schooling process: Early versus late influence of parental characteristics. *American Sociological Review*, *49*(6), 784-802.

Archer, J. (1996). Comparing women and men: what is being compared and why? *American Psychologist, 51*(2), 153-154.

Baer, J. (1999). The effects of family structure and SES on family processes in early adolescence. *Journal of Adolescence, 22*(3), 341-354.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1997). *Self-Efficacy: The exercise of control*. USA: W. H. Freeman and Company.

Basang, Z. (2006). Research on middle and primary school student's cognition level to statistics. Dissertation. Dongbei Normal University.

Borman, G. D., & Overman, L. T. (2004). Academic resilience in mathematics among poor and minority students. *Elementary School Journal. 104*(3),177-195.

Bosker, R.J., Kremers, E. J. J., & Lugthart, E. (1990). School and instruction effects on mathematics achievement. *School effectiveness and School Improvement, 1*(4), 233-248.

Bosker, R. J.,&Witziers, B. (1996, April). *The true size of school effects*. Paper presented at the AERA, New York.

Byrnes, J.P. & Miller, D.C.(2007). The relative importance of predictors of math and science achievement: An opportunity-propensity analysis. *Contemporary Educational Psychology, 32*(4), 599-629.

Caldas, S. J., & Bankston, C. L. (1997). The Effect of School Population Socioeconomic Status on Individual Student Academic Achievement. *Journal of Educational Research, 90*(5), 269-277.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M,; Weinfeld, F. D., & York, R. L.,(1966). *Equality of Educational Opportunity***,** Washington, D. C.: United States Government Printing Office.

Department of Planning, State Education Commission (DPSEDC) (1986). *Achievement of education in China: Statistics 1980-1985*. Beijing: People's Education Press.

Desoete, A., & Roeyers, H. (2005). Cognitive building blocks in mathematical problem solving in grade 3. *British Journal of Educational Psychology, 75,* 119-138.

Desoete, A., & Roeyers, H. (2006). Metacognitive macroevaluations in mathematical problem solving. *Learning and Instruction, 16*(1), 12-25.

Desoete, A., Roeyers, H., & Buysse, A. (2001). Metacognition and mathematical problem solving in grade 3. *Journal of Learning Disabilities, 34*(5), 435-449.

DiPerna, J.C., Volpe R.J., and Elliott, S. N..(2002) A model of academic enablers and elementary reading/language arts achievement, *School Psychology Review, 31*(3) , 298–312.

DiPerna, J.C., Volpe, R.J., & Elliott, S.N.(2005). A model of academic enabler and mathematics achievement in the elementary grades*, Journal of School Psychology. 43*(5), 375–392.

Dirks, E., Spyer, G., & Van Lieshout, E.C.D.M. (2008). Prevalence of combined reading and arithmetic disabilities, *Journal of Learning Disabilities, 41* (5), 460-473.

Duncan, G. J., Brooks-Gunn, J. , & Klebanov, P. K. (1994). Economic deprivation and early childhood development*. Child Development, 65(2)*, 296-318.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1(1)*, 3-14.

Elliott S.N., DiPerna, J.C., Mroch, A.A., & Lang, S.C. (2004). Prevalence and patterns of academic enabling behaviors: An analysis of teachers' and students' ratings for a national sample of learners, *School Psychology Review 33*(2), 302-309.

Fan, X. (2001). Parental involvement and students' academic achievement: A growth modeling analysis*. Journal of Experimental Education, 70(1)*, 27–61.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties, *Journal of learning disabilities, 38* (4), 293-304.

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129-145.

Grimm, K. J. (2008). Longitudinal Associations between reading and mathematics achievement. *Developmental Neuropsychology, 33*(3),410-426.

Hanushek, E. A., Kain, J. F. & Rivkin, S. G. (2004). Disruption versus Tiebout improvement: the costs and benefits of switching schools. *Journal of Public Economics, 88*(9-10), 1721-1746.

Heck, R. H. (2007). Examining the relationship between teacher quality as an organizational property of schools and students' achievement and growth rates. *Educational Administration Quarterly 43*(4), 399-432.

Heyneman,S.P. & Loxley, W. A. (1983). The effect of primary-school quality on academic achievement across twenty-nine high- and low-income countries. American Journal of Sociology, 88 (6), 1162-1194.

Hox, J. J. (2002). *Multilevel Analysis. Techniques and Applications*. Erlbaum, Mahwah, NJ.

Huang, F. L. (2010). The role of socioeconomic status and school quality in the Philippines: Revisiting the Heyneman-Loxley effect. *International Journal of Educational Development, 30*(3). 288-296.

Hyde, J. S., Fennema,, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139-155.

Jeynes, W. H. (2005). A meta-analysis of the relation of parental involvement to urban elementary school student academic achievement. *Urban education, 40* (3),237 -269.

Kruger, J. (2002). Unskilled and unaware – but why? A reply to Krueger and Mueller. *Journal of Personality and Social Psychology, 82*(2), 189-192.

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review, 28*(1), 49-57.

Kyriakides, L. & Luyten, H. (2009). The contribution of schooling to the cognitive development of secondary education students in Cyprus: an application of regression discontinuity with multiple cut-off points. *School Effectiveness and School Improvement, 20* (2), 167-186.

Lashway, L. (2002). *Using School Board Policy To Improve Student Achievement*. ERIC Digest, DeC 163. No. ED-99-C0-0011.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education, 62(3),*172-192.

Li C. L. (2005a). *Duanlie yu suipian: dangdai Zhongguo shehui jieceng fenhua shili fenxi (Cleavage and Fragmentation: An Empirical Analysis of Social Stratification in Contemporary China)*, Beijing: Shehui kexue chubanshe (*Sociology Science Publisher*).

Li, C. L. (2005b). Dangdai zhongguo shehui de shengwang fencing: Zhiye shengwang yu shehui jingji diwei zhishu celiang (Prestige Stratification in the Contemporary China: Occupational prestige measures and socio-economic index). *Shehuixue Yanjiu* (*Sociological Research), 2,* 74-102.

Lim, C. S., & Zhao, X. P.(2005). *Assessment and Examination System: A Comparative Study between Malaysia and Shanghai (China).* Paper presented at the Third East Asia Regional Conference on Mathematics Education, Shanghai, 5-12 August, 2005.

Liu, X. M. (2007). Study on the characteristics and laws of the development of the pupil's spatial sence. Dissertation. Dongbei Normal University.

Louis, V. V., & Zhao, S. Y. (2002). Effects of family structure, family SES, and adulthood experiences on life satisfaction. *Journal of Family Issue, 23(8)*, 986-1005.

Ma, X. (1997). A multiple regression analysis of mathematics achievement in the Dominican Republic. *International Journal of Educational Development, 17* (3), 313-321.

Mackenzie, D. E. (1983). Research for School Improvement: An Appraisal of Some Recent Trends. *Educational Researcher 12* (4), 5-7.

Marat, D.,(2005). Assessing mathematics self-efficacy of diverse students from secondary schools in Auckland: Implications for academic achievement.

Issues in Educational Research, 15, 37-68. Retrieved from http://www.iier.org.au/iier15/marat.html.

Meece, J. L., Wigfield, A., & Eccles, J. S. (1990). Predictors of Math Anxiety and Its Influence on Young Adolescents' Course Enrolment Intentions and Performance in Mathematics. *Journal of Educational Psychology, 82*(1), 60-70.

Ministry of Education in P.R. China. (2001). *Mathematics Curriculum Standards in the Phase of Full-time Compulsory Education - Experimental Manuscript.* Beijing: Beijing Normal University Publisher.

Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Development Psychology, 25(1)*, 147-166.

OECD (Organisation for Economic Co-operation and Development) (2009). *PISA 2006 Technical Report*. OECD Publishing.

Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). *TIMSS 2007 Technical Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timss.bc.edu/TIMSS2007/PDF/T07_TR_Chapter3.pdf.

Opdenakker, M. C., Van Damme, J., Fraine, B., Landeghem, G., &Onghena, P. (2002). The effect of schools and classes on mathematics achievement. *School effectiveness and school improvement. 13*(4), 399-427.

Opdenakker, M.C., & Van Damme, J. (2001). Relationship between school composition and characteristic of school process and their effect on mathematics achievement. *British Educational Research Journal, 27(4),* 407-432.

Organization of Educational Yearbook in China, (1984). *Educational Yearbook in China* (1981-1984-Key School). Beijing; Encyclopedia of China Publishing House.

Pajare, F., & Miller, M. D. (1997). Mathematics self-efficacy and mathematical problem solving: Implications of using different forms of assessment. *Journal of Experimental Education, 65*(3), 213-229.

Pajares, F. & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24,* 124-139.

Pretorius, E. J. & Currin, S. (2010). Do the rich get richer and the poor poorer? The effects of an intervention programme on reading in the home and school language in a high poverty multilingual context. *International Journal of Educational Development, 30* (1), 67-76.

Philips, M. (1997). What makes school effective? A comparison of the relationships of communitarian climate and academic climate to mathematics achievement and attendance during middle school. *American Educational Research Journal, 34*(4),633-662.

Rasbash, J., Steele, F., Browne, W., & Prosser, B.(2004*). A user's guide to*

MLwiN version 2.0. London, Institute of Education.
Raudenbush, S. W. & Bryk, A. S. (2002). *Hierachical linear models. Applications and data analysis methods* (2nd ed.). Thousand Oaks: Sage Publications.
Reyes, L.H., & Stanic, G.M.A.(1988). Race, sex, socioeconomic status, and mathematics. *Journal for Research in Mathematics Education, 19(1),* 26-43.
Ross, J. A., & Gray, P. (2006). School leadership and student achievement: The mediating effect of teacher beliefs. *Canadian Journal of Education, 29* (3), 798-822.
Russel, R.L. & Ginsburg, H. P. (1984). Cognitive analysis of children's mathematics difficulties. *Cognition and Instruction, 1(2),* 217-244.
Sammons, P., Hillamn, J., & Moretimore, P. (1995). *Key characteristics of effective schools: A review of school effectiveness research.* London: Office for standards in Education [OFSTED].
Scarr, S. (1988). Race and gender as psychological variables: Social and ethical issues. *American Psychologist, 43(1):* 56-59.
Scheerens, J. & Creemers, B. P. M. (1989). Conceptualizing school effectiveness. *International Journal of Educational Research, 13(7),* 691-706.
Secada, W. G. (1992). *Race ethnicity, social class, language, and achievement in mathematics*. In D. A. Groues (Ed.), Handbook of research on mathematics teaching and learning (pp. 623e660). New York: Macmillan.
Silver, D. H. Pennett, H. D., Black, J. L., Fair, G. W., & Balise, R. R.(1999). Stability of Arithmetic disabilities subtypes. *Journal of Learning Disabilities, 32(2)*, 108-119.
Sirin, S. R. (2005). Socioeconomic status and academic achievement: a meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.
Smith, T. M., Desimone, L. M., & Ueno, K. (2005). "Highly qualified" to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics. *Educational Evaluation and Policy Analysis, 27*(1), 75-109.
Spelke, E., & Ellison, K. (2008). *Gender, math and science.* Harvard University Psychology Department Working Paper.
Stigler, J., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A..(1999). *The TIMSS videotape classroom study. Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan and the United States*. Washington, DC: US Department of Education.
Tate, W. F. (1997). Race-ethnicity, SES, gender and language proficiency trends in mathematics achievement: an update. *Journal for Research in Mathematics Education, 28*(6), 652-679.
Teddlie, C. (1994). *The integration of classroom and school process data in school effectiveness research*. In Reynolds, D. B.P.M., Creemers, P.S.,

Nesselrodt, E.C. Schaffer, Stringfield, S. & Teddlie C. (Eds.), Advances in School effectiveness research and practice. (pp.111-132). Oxford: Pergamon Press.

Van Damme, J., De Fraine, B., Van Landeghem, G., Opdenakker, M. C., & Onghena, P. (2002). A new study on educational effectiveness in secondary schools in Flanders: An introduction. *School Effectiveness and School Improvement*, *13*(4), 383–397.

Van Ewijk, R., & Sleegers, P. (2010), The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review.* doi:10.1016/j.edurev.2010.02.001.

Veenman, M.V.J., Van Hout-Wolters, B.H.A.M., & Afflerbach, P. (2006). Metacognition and learning. Conceptual and methodological considerations. *Metacognition Learning, 1*, 3-14.

Venkatanarayana, M. (2005). *On The Non-Random Distribution of Educational Deprivation of Children in India.* Paper Submitted for A Workshop on “Normative and Quantitative Analysis of Educational Inequalities” held on 31 May 2005 in Belgium.

Vermeer, H., Boekaerts, M., & Seegers, G. (2000). Motivational and gender differences: Sixth grade students’ mathematical problem-solving behavior. *Journal of Educational Psychology, 92(2)*, 308-315.

White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin, 91*(3): 461-481.

Willms, J. D. (2003). *Ten hypotheses about socioeconomic gradients and community differences in children’s developmental outcomes*. Ottawa, Ontario, Canada: Applied Research Branch of Human Resources Development Canada.

Zajonc, R. B. , & Markus, G. B.(1975). Birth order and intellectual development. *Psychological Review. 82*(1), 74-88.

Zeng, X. D. (2008). Teacher Education in the context of social change in China, Submission to the international Alliance of Leading Education Institutes. Retrieve from : http://www.intlalliance.org/China13aug.pdf.

Zimmermann, W., & Cunningham, S., (Eds.). (1991). *Visualization in teaching and learning mathematics.* Washington, D.C.: American Mathematical Association.