UNIVERSITEIT GENT

FACULTEIT PSYCHOLOGIE EN PEDAGOGISCHE WETENSCHAPPEN

Mathematics learning performance and Mathematics learning difficulties in China

Ningning Zhao

Promotor: Prof. Dr. Martin Valcke
Co-promoter: Prof. Dr. Annemie Desoete

Proefschrift ingediend tot het behalen van de academische graad van Doctor in de Pedagogische Wetenschappen

2011

This Ph.D research project was funded by

Ghent University BOF Research Grant (BOF07/DOS/056)

Acknowledgements

There is still a long and indistinct way and I will keep on going to explore the unknown.

路漫漫其修远兮，吾将上下而求索。

- Qu Yuan (340-278 BC)

This dissertation would not have been possible unless so many persons contributed to it. The first person I should give my gratitude is Prof. dr. Cong Lixin in Beijing Normal University. It is she who recommanded me to my promoter - Prof. dr. Martin Valcke. Based on the cooperation contact between the two universities, I have the opportunity to start my journey in Ghent University. The fantasty journey started from Year 2007 gudided by the Prof. dr. Martin Valcke.

I am heartily thankful to my promoter Prof. dr. Martin Valcke and my co-promotor Prof. dr. Annemie Desoete, whose encouragement, supervision and support from the preliminary to the concluding level enabled me to carry on the research project. My deepest gratitude is to Prof. dr. Martin Valcke. I am not a smart student who always give him so much revision work. It is extremely fortunate for me to have a promoter who is characterized by energy, enthusiasm and patience. Some excellent ideas will arise from his brain sometimes which make us think his synaptic connection in his brain is very strong. His face is always full of real warmth on the research. His cheerful spirits and bubble laughter infect me and encourage me to get through the difficulties. And, he generously supported me to get access to the schools and gave lectures in return for the schools which I gathered the data. His mentorship was paramount in providing a well rounded experience consistent my long-term career goals.

And, I want to express my sincerely acknowledgement to my co-promotor Prof. dr. Annemie Desoete. I am deeply grateful to her for the suggestions she gave me on the mathematics education. She is the expert in the mathematics learning difficulties who inducted me into the mysterium of research. When I meet my limitations in the specific problem, she gave me lots of professional suggestions to expand the possibility for the theory development. She is a person of high efficiency who completes her work ahead of schedule which is a model for me. I hope that one day I would become a splendid female professor as her.

In the past four years, the two promoters introduced many knowledgeable experts from or out of our faculty to me. It is a pleasure to thank those who made this dissertation possible. I owe my deepest gratitude to the three experts of my guidance committees, Prof. dr. Bernadette van Hout-Wolters, Prof. dr. Pol Ghesquière and Prof. dr. Eric Broekaert. I thank Prof. Bernadette van Hout-Wolters for the revision work she did for me. I thank Prof. Pol Ghesquière for giving me lots of related references to expand my theoretical background. I thank Prof. Eric Broekaert for my logical structure of the research. I would like to thank the expert from our department - Mr. JeanPierre Verhaeghe - who gave me many good suggestions in the data analysis which reveal his professional ability in statistics things.

I owe my gratitude to the Dean Prof. dr. Geert De Soete, the colleagues in our faculty and the staffs of secretaries and others. It is they who construct a team-oriented, cooperative and supportive environment. I would like to thank my office mate Ms. Liesje De Backer who shared the happy and nervous with me, the mathematics team member of Ms. Elise Burny and Mr. Hendrik Van Steenbrugge who worked together with me, the warm-hearted friends Maria, Hester, Jo, Ruben, the previous staffs Dr. Goedele Verhaeghe, Mrs. Ilse Sinnaeve who showed me their virtue which impressed me so much. Of course, I have to thank to my Chinese colleagues who shared the long journey with me: Dr. Chang Zhu, Dr. Guoyuan Sang, Mrs. Qiaoyan He, Mr. Lin Wu and Ms. Diya Dou. I learned a lot from the elder bothers and elder sisters to improve myself. I thank for the Chinese friends in our faculty, Dr. Qing Cai, Dr. Qi Chen, Ms. Beiwen Chen, who shared the interesting ideas from different perspectives of research. I thank for the Chinese friends who shared the experience and supported me when I was upset, Baoyi, Mr. Shang Chen, Ms. Jianyun Wang, Mr. Nengye Liu, Ms. Jingsheng Peng. I thank for some Chinese friends who gave me support in my life in Gent, Mr. Chaobo Huang, Dr. Kai Chen, Mrs. Hong Liu, Mr. Jianjun Sun, Mr. Zanqun Liu, Mr. Lei Ding, Ms. Dangqing Ye, Mr. Baoyu Zhang, Mrs. Wei Wu, Mr. Yongbin Dai, Mrs. Xun Yan, Mrs. Liyan Wang, Mrs. Juan Ma.

Since the data is from twenty schools in five areas, I should show my appreciation to Education College of Beijing Normal University, Education Science College of South China Normal University and some Educational Bureaus in different cities. We thank Prof. Cong Lixin, Prof. Cai Yonghong, Dr. Zhang Qiulin, Prof. Zeng Xiaodong, Prof. Qi Chunxia, Dr. Li Minyi and Dr. Cheng Li from Beijing Normal University, Prof. Liu Quanli from Beijing Union University and Prof. Mo Lei, Prof. Huang Fuquan, Prof. Gao Linbiao and Mrs. Chen Dongmei, Mr. Xie Guangling, Mrs. Chen Caiyan from South China Normal University and all the students, parents, and teachers who cooperated so splendidly in data collection of this research.

Thanks for my father and mother who supports me during the long journey! I am a willful child who sticks to my own dream so long and make them worry. Thanks!

Ningning Zhao
120.045, Henri Dunantlaan 2, Ghent
September 2011

Contents

Part III

Chapter 1

General Introduction*

Abstract

Mathematics is a critical ability of human beings in modern society. Cross-cultural studies provide us with information about the way specific variables and processes contribute to mathematics performance in specific cultural contexts. In this introductory chapter, we present a literature review that summarizes the available research in this field. The aim is to develop a conceptual model that shows how the different studies in this doctoral thesis are interlinked. In the review of the available research two perspectives have been adopted: (a) a very broad perspective that builds on general instructional effectiveness studies, and (b) a specific perspective that centers on national and international research about predictors of mathematics performance. Next to the identification of available theoretical and empirical models that explain (mathematics) learning, this chapter will also build on a qualitative content analysis of available research about mathematics learning in China. This will result in a further delineation of variables that play a role to describe and explain mathematics learning and performance. The outcome of this combined approach is a first outline of a conceptual framework that will be helpful to direct the research, reported in this PhD dissertation.

1. Introduction

Mastery of mathematics is a key literacy component that influences children's success in education and in future society (Engle, Grantham-McGregor, Black, Walker, & Wachs, 2007). The focus on mathematics learning and mathematics ability development have been a recurrent topic in educational and psychological studies for over 100 years (Geary, 2006). In the early 20th century, psychologists started to study the children's understanding of number, arithmetics and specific mastery of mathematics elements via experimental research (Brownell, 1928; Thorndike, 1922; Thorndike & Woodworth, 1901). These studies contributed to our knowledge about mathematics learning from a psychological perspective. However, cross-national

---

* This chapter is based on the submitted paper of Zhao, N.N., Valcke, M., & Desoete, A. (submitted). Quantitative Content Analysis on the Studies of Mathematics Performance and related Predictors for Chinese Students From 1960s to 2010s. *Asia Pacific Education Review*

studies - since Husen (1967) - reveal that mathematics learning is also shaped by culture (Tang, Zhang, Chen, Feng, Ji, Shen, Reiman,& Liu, 2006). Also ongoing international performance indicator studies (such as PISA, TIMSS) focused researchers' interest on variables affecting mathematics performance from both psychological and socio-cultural perspectives (OECD, 2010; Mullis, Martin, & Foy, 2008).

A recurrent theme in cross-cultural studies is that Chinese students outperform learners from other countries in the mathematics domain (Geary & Salthouse, 1996; Imbo & LeFevre, 2009). The reasons behind this phenomenon seem to intrigue researchers. Many studies compare learner characteristics of children in China and other countries, and this at different levels in the educational system (Geary & Salthouse, 1996; Siegler & Mu, 2008). However, studies set up within the local Chinese context are rare (See the content analysis of research in the next paragraphs). Although mathematics education is considered to be very important in Chinese education – considering the high emphasis on mathematics summative assessment - limited empirical studies are available that explore the variables' affecting learning performance from a variety of perspectives. This lack of in-dept research might be due to barriers and limited resources, the limited power of local educational bureaus, and/or the limited attention paid to this type of research in developing countries (Li, 2006).The present PhD study tries to contribute to the research literature that fills this gap in the available empirical studies about Chinese mathematics teaching and learning. The gap in the literature is larger than initially expected since the discussion already start by looking at the available assessment instruments to determine mathematics performance. The gap widens when looking at the available comprehensive models to describe and explain mathematics learning and performance, and the gap is even larger when focusing on children at risk or underperforming in the mathematics domain.

Three research objectives directed the different studies in this PhD study. First, we aim at developing a standardized assessment instrument to study in a valid and reliable way mathematics performance of Chinese primary school children. Second, by bringing together available evidence about variables and processes that predict mathematics learning and performance, we aim at studying the important predictors for mathematics learning performance in the Chinese context. Thirdly, we will centre on children at risk. The third research aim is therefore to identify the predictors of the students with learning problems in mathematics.

2. Theoretical background

In order to develop an overview of studies about mathematics learning, we first analyze a number of established theoretical models and link them next to mathematics learning. Next, we center on particular models that studied mathematics learning and look for influencing processes and variables (Brownell, 1928; Geary & Hoard, 2005; Thorndike & Woodworth, 1901; Thorndike, 1922). This approach helps to map a first set of relevant components of a model. However, as mentioned before, mathematics performance is also embedded in a cultural context. This will be added while exploring additional models.

2.1 General learning models

2.1.1 Walberg's educational productivity model: towards complex models of school learning

One of the first established comprehensive models trying to map what influences learning, was developed by Walberg and his colleagues. From the early 1980s, Walberg and colleagues started to elaborate their educational productivity model (Walberg 1981; 1982). It made explicit factors that were expected to contribute to learning outcomes (Reynolds & Walberg, 1992). Based on available evidence, they estimated the particular impact of particular (sets of) factors in a variety of school subjects.

Three sets of nine factors are proposed that are hypothesized to improve student achievement (Fraser, Walberg, Welch, & Hattie, 1987). First they point at student aptitude-attribute factors, including (a) ability or prior achievement, (b) age, (c) motivation or self-concept as indicated by personality tests or willingness to persevere on learning tasks. Second, they point at instructional factors, including (d) quantity of instruction, and (e) quality of the instructional experience. Third, the authors describe the educationally stimulating factors in the (f) home environment, (g) the classroom or school environment, (h) the peer group environment, and (i) the mass media (especially television). Figure 1 depicts the resulting "Model of School Learning" (McGrew, 2007).

The contribution of Walberg's studies is far-reaching since he clearly makes a distinction between three sets of factors: at the student level, at the instructional level and at the environment level. This reappears in later models that focus explicitly on mathematics learning, such as the Opportunity-propensity model (Byrnes & Miller, 2007; Byrnes & Wasik, 2009).

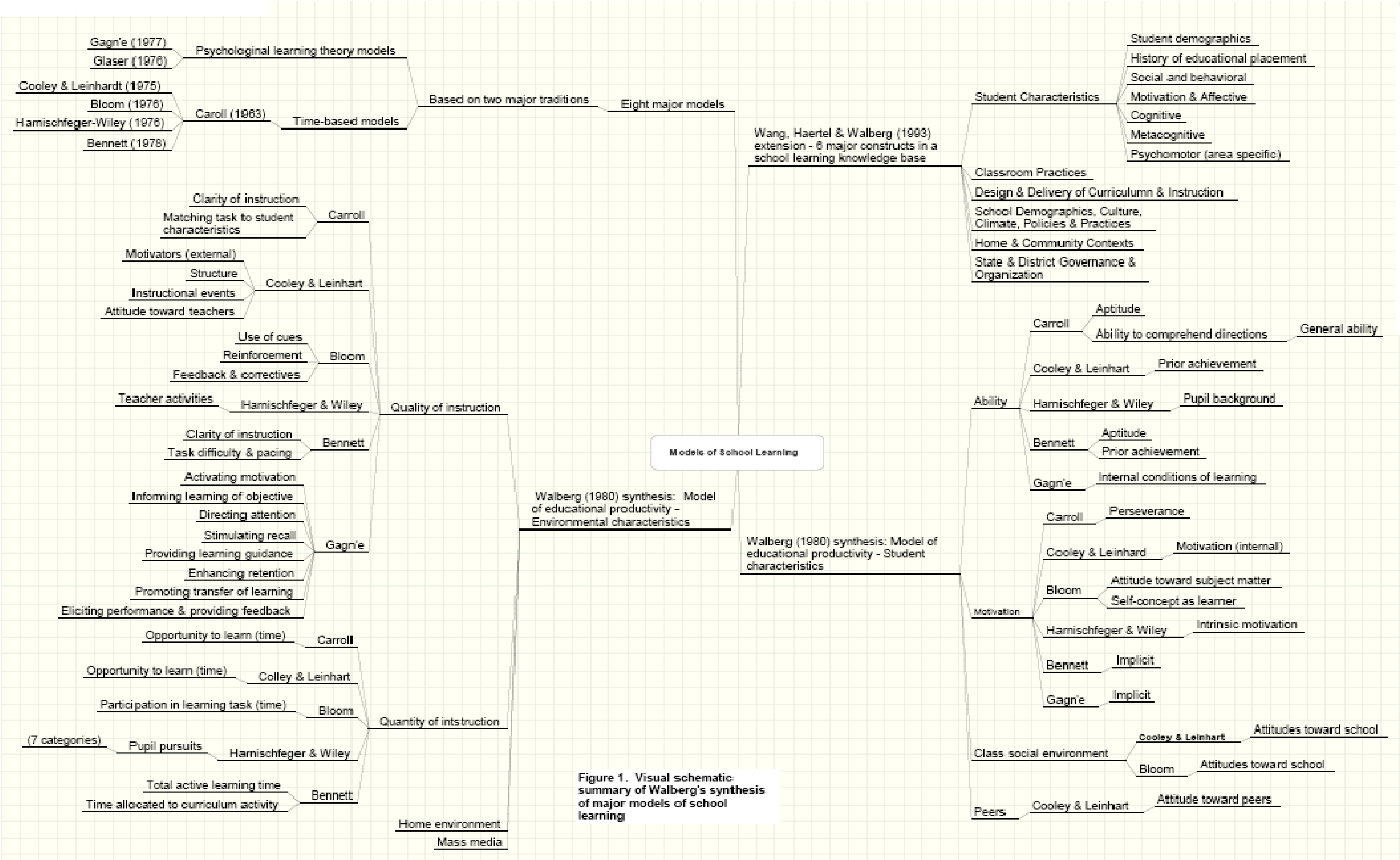


Figure 1. Walberg's synthesis of available research into an overview of "Models of School Learning" (based on McGrew, 2007)

The entire dissertation can be split up into two parts. The first part focuses on "normal" performing students. The second part centers on students with learning difficulties.

2.1.2 Creemer's educational effectiveness model: towards a nested hierarchical structure

Another model of relevance in the context of this introductory chapter, is Creemers' educational effectiveness model. This model started from the heavy debate about school effectiveness as a response to the Coleman report in the USA and the Plowden report (1967) in the United Kingdom. These studies questioned the added value of schools in coping with the dominant impact of the parents' background on learner achievement. Creemers and his colleagues reviewed the history of this debate and discuss the relationships between school effectiveness and school improvement in their paper "Educational Effectiveness and Improvement: The Development of the Field in Mainland Europe" (2007). In their model, they recognize the impact of social economic background (SES) variables, but additionally point at empirical research that underpins the impact on achievement of many other variables They go beyond a too direct and unidimensional relationship between SES and achievement.

In this model, Creemers distinguishes four levels to be taken into consideration: the student level, the classroom level, the school level and the context level. Based on these levels, key concepts and factors from the Carroll's learning model (1963) have been selected to further develop the model. Figure 2 depicts Creemers' framework of educational effectiveness. It is interesting to note that this model incorporates a feature not yet present in Walberg's model: the cross level interactions between the levels and factors.

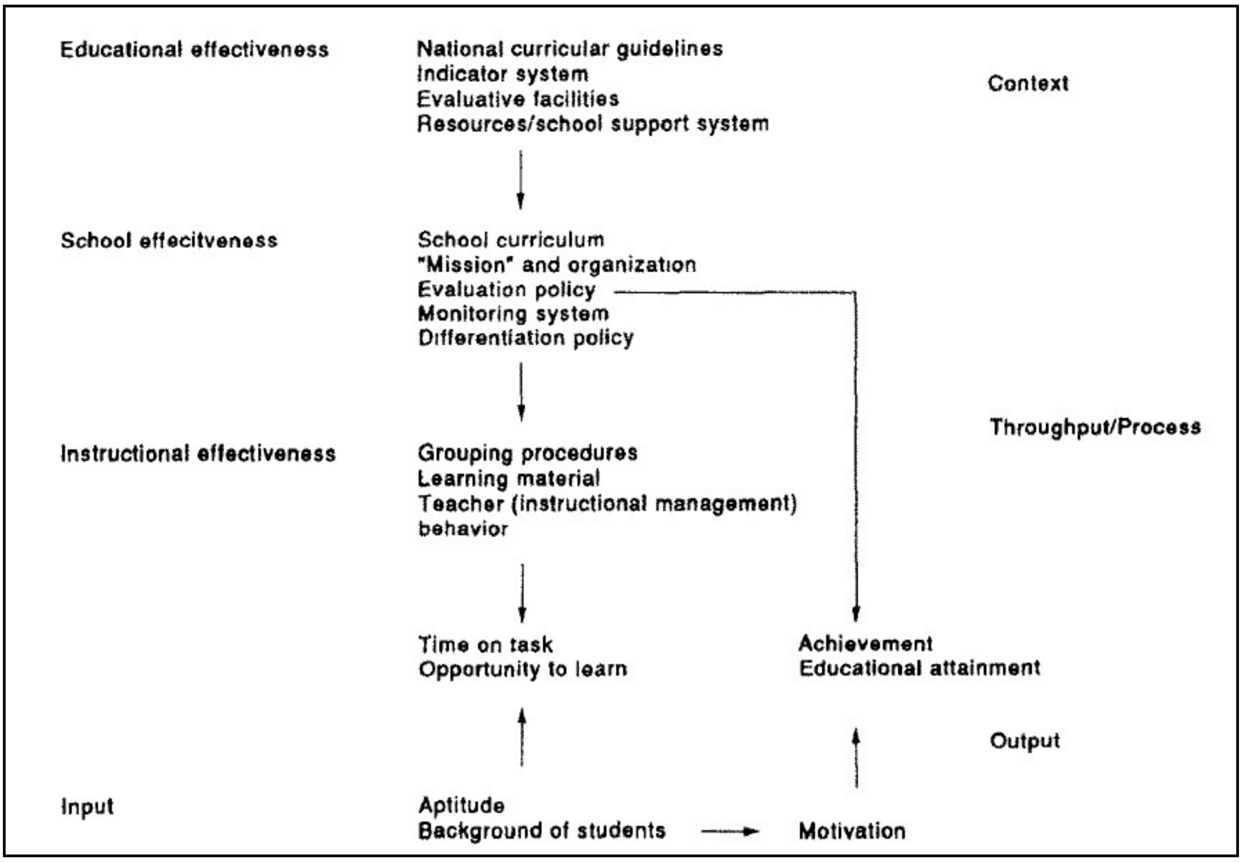


Figure 2. Creemers' Educational effectiveness model
(Creemers, & Scheerens, 1994, p. 132)

Creemers' model assumes that classroom- and school-level factors exert a joint influence on achievement, thus suggesting a multilevel structure in the way the different factors play a role. In recent years, with the development of more advanced statistical methodologies, Creemers's model has been evaluated by a number of researchers (De Jong, Westerhof, & Kruiter, 2004; Kyriakides, Campbell & Gagatsis, 2000). Compared to previous models, Creemers' model stresses an educational perspective on the academic achievement. As such, we can state that the model reflects to a larger extent the real educational situation of school base learning that recognizes the nested nature of a complex and interacting set of factors.

2.1.3 Geary's evolutionary theory – towards a more complex picture of the role of individual control mechanisms and adaptations to the ecological setting

Unlike previous models, Geary's theory is based on assumptions about individual development in relation to cultural and evolutionary influences. Geary claims that schools play the interface between evolution and culture. Thus, children learn through support that results from affective, conscious-psychological and cognitive mechanisms that are pushed by social, biological and physical modules (Geary, 2005) (See figure 3).

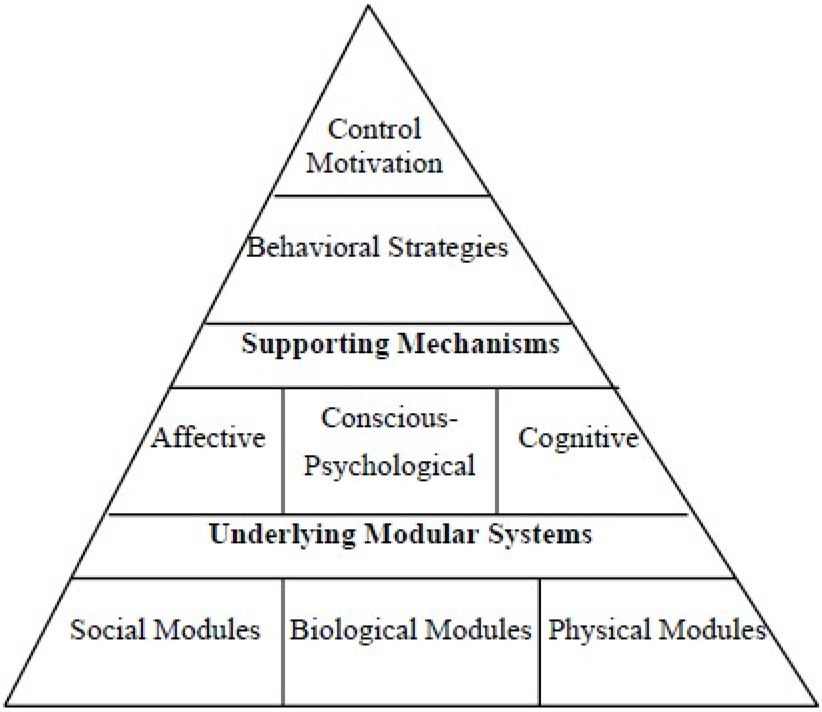


Figure 3 Geary’s evolutionary theory (Geary, 2007, p.386).

In Geary’s model, motivational control and behavioral strategies are highlighted as critical tools to solve the evolutionary pressure and the influence of the social, biological and physical modular systems (Geary, 2008). Compared to the previous two models, this model emphasizes (Geary & Bjorklund, 2000): (1) the development of the individual who makes use of specialized cognitive processing modules that have developed as a result of continuous problem solving attempts during his/her biological evolution; (2) the influence of mechanisms showing how the development of competencies is the result of adaptations to the local ecological setting (Siegler, 1996). The model hints at the combined impact of contextual factors and the way the individual learner controls motivational and cognitive resources to meets development needs..

Next to the three comprehensive models briefly outlined above, many others studies rather focus on one or two variables critical factors influencing learning performance. An exception is the model of Mcllrath and Huitt (1995). They reviewed and integrated available models into a heuristic teaching-learning process model. Figure 4 represents their effort that shows how variables in the student, class, school and context play together and affect student achievement.

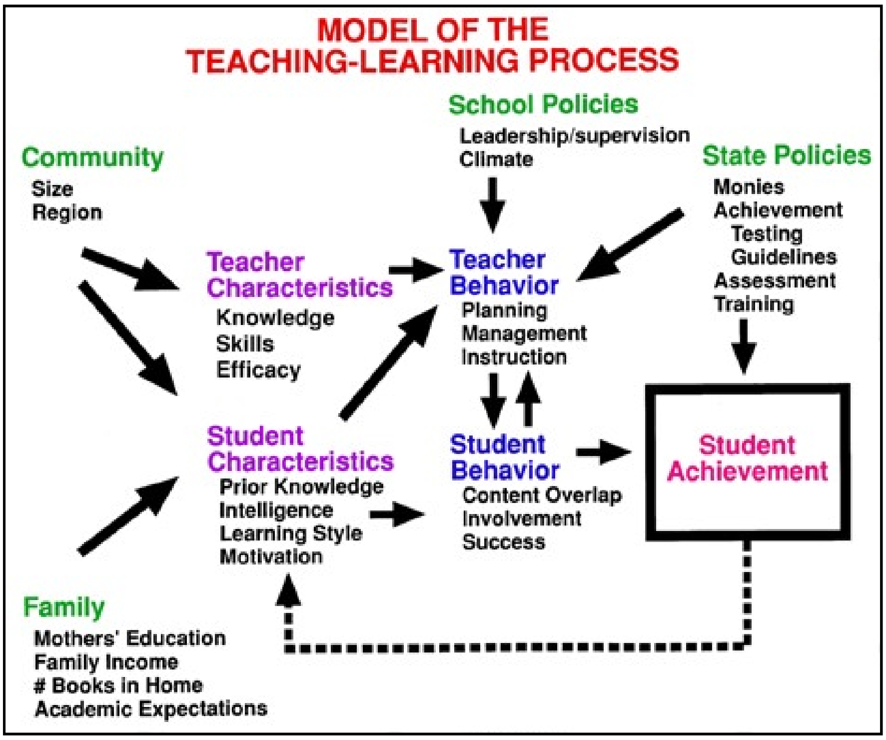


Figure 4. Mcllrath and Huitt's teaching-learning process Model (Mcllrath, & Huitt, 1995, Retrieved April 2008, from http://chiron.valdosta.edu/whuitt/papers/modeltch.html)

2.2 Models focusing on mathematics learning

This section highlights available learning models that have been set up and empirically tested in the context of mathematics learning.

2.2.1 The opportunity-propensity model

The opportunity-propensity model is one of the distinct models being developed in recent year in the field of mathematics learning (Byrnes & Miller, 2007; Byrnes & Wasik, 2009). The model partly builds on Walberg's ideas, but it especially restructures a variety of factors and how they interact in the way they influence later achievement.

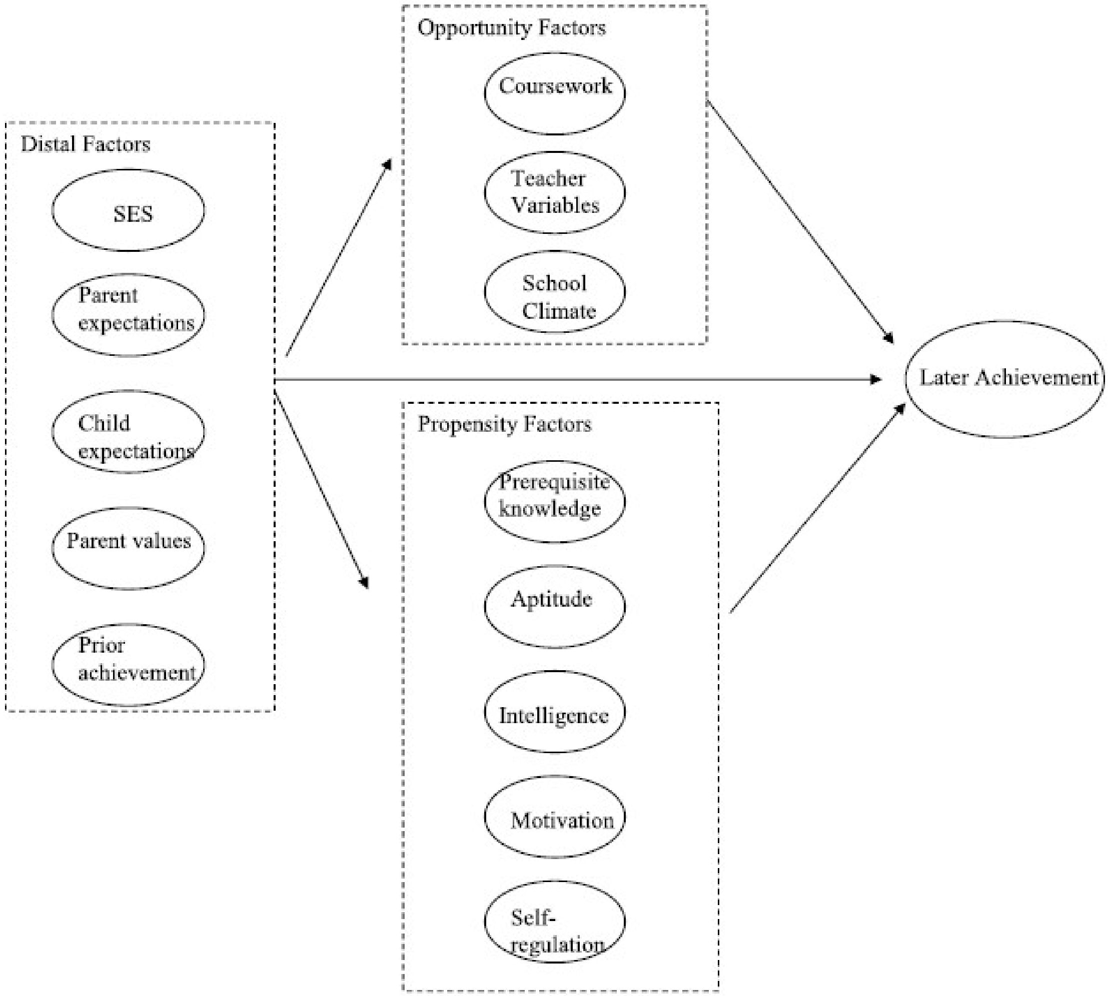


Figure 5. Opportunity-propensity model (Byrnes, & Miller, 2007, p. 602)

As we can derive from figure 5, in this model, there are three basic sets of factors (Byrnes & Miller, 2007). First, the authors distinguish “opportunities”, referring to elements in the culturally defined context in which an individual is presented with content to learn. Second, they distinguish “propensity factors” that refer to internal variables and processes that affect the ability to learn particular. Third, the authors make explicit “distal factors” that enable or explain the extent to which learners are affected by the opportunity factors, engage the propensity factors and/or directly influence later achievement. This model goes beyond limitations of the Walberg model (Byrnes & Wasik, 2009). The model expects researchers to combine the impact of opportunities (high or low), attitudes (willing to use or not) and ability (able or unable) in view of calculating the predictive impact on later achievement. In addition, the model presents a dynamic system that is to be tested over time. Current achievement has – therefore – to be entered as an additional predictor of later achievement. The model is therefore geared to longitudinal studies. Available empirical research with this model points at the propensity factors to be the most important predictors for achievement (Byrnes & Wasik, 2009).

2.2.2 Other models: emphasis on the nested nature of influencing factors

In the previous sections, we already mention Creemers' model of school effectiveness. Recently, other models followed the idea of Creemers and test this type of model in the mathematics domain (Opdenakker & Van Damme, 2001; Opdenakker, Van Damme, Defraine, Landeghem & Onghena, 2002). These studies reveal that the school- and class- level variables account for a large proportion in the variance of mathematics achievement.

At an international level, comparative studies know a long tradition and have been conducted since the 1950s. The most famous studies – in this context – are set up by the International Association for the Evaluation of Educational Achievement (IEA)'s project of *Trends in International Mathematics and Science Study (TIMSS)* from 1995 and by the Organization for Economic Co-operation and Development (OECD), the Programme for International Student Assessment (PISA) set up since 2000. Both studies are set up in a cyclic way and focus in part on mathematics achievement. They collect "rich" data from both students, parents, teachers and schools; thus mirroring a model that all related variables and processes influence learning and resulting mathematics achievement (See Figure 6).

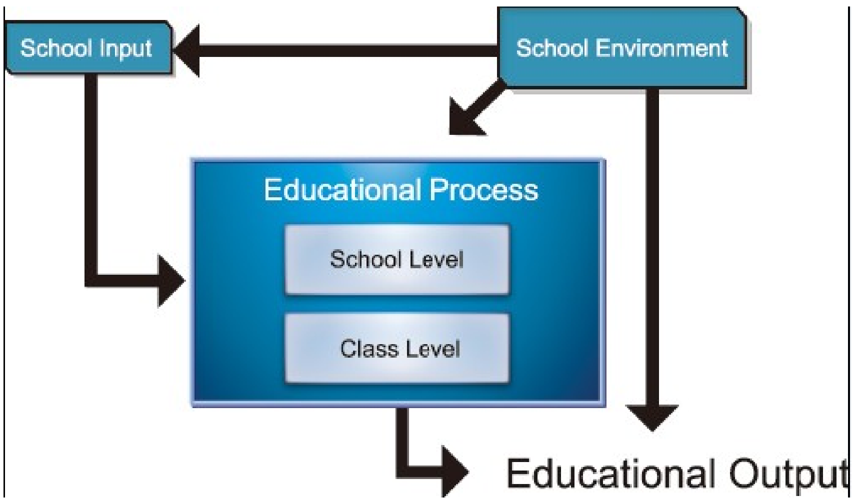


Figure 6. International Project of PISA and TIMSS

Figure 6 reflects the hypothetical structure adopted by both international comparative studies.

To conclude, several theoretical and empirical models present input to develop our own conceptual framework. These models already have an empirical base and reflect the history in the thinking about factors affecting learning and performance. What should be learn from these models in view of our own conceptual framework?

- (a) a comprehensive model should consider a variety of variables related to biological-primary, biological-secondary cognitive development influences;
- (b) the variables should be structured at different levels, while the cases are nested;
- (c) individuals do not merely respond to the context in a passive way, but also try to control the resources in the environment in view of their own development/ evolution;
- (d) during evolution/development, a dual learning process is activated that supports student development: the iterative development of performance and the related development of propensity variables;
- (e) interventions can be set up fostering the development and/or activation of students' propensity variables, thus improving disadvantaged situations at family and/or school level.

In the following section, we will try to construct a conceptual framework for our study.

3. Towards a comprehensive conceptual framework for mathematics learning of Chinese elementary school children

Based on the five key characteristics of available models in the literature, our conceptual framework will consider three levels in specific influencing variables: *individual level* variables related to the student and his/her family, *class level* variables related to the teacher, *school level* variables related to the location of the school (e.g., gross domestic product of the regional location of the school).

The three levels incorporate a number of sub-constructs. The selection of these sub-constructs/variables could build on the available model-related literature. In the present chapter, we adopt a different approach. We combine the analysis of available model-related literature with the analysis of a China related corpus of empirical research. This will help to contextualize our modeling activity and answer the need for research that considers the cultural setting when studying learning and related performance.

3.1 Mathematics learning performance

In the literature, the term of "performance" is used in parallel to other concepts, such as "achievement", "outcome", "result", "output", "productivity", and many others. Often, there are connotations and denotations linked to these terms: it only refers to students' outcomes mathematics test scores as measured with a specific instrument and neglects the full complexity of the processes involved in resulting in particular "score". Therefore, in the present study, we will approach the definition of "performance" in a careful way. First, we focus on the debate in China about mathematics performance. Next, we center on its measurement history. Finally we make a decision as to the basic operational definition of the concept in the context of our studies.

Since the curriculum reform of 2001 about "what should be included in the curriculum", a nationwide debate started among Beijing and Shanghai scholars. This debate reflected a discussion between a focus on "Zhishi" versus "Nengli"; knowledge versus abilities (See, Wang, 2004). Some educational researchers criticized previous teaching, and curriculum approaches to be too knowledge-oriented and advocated a change towards an ability-orientation (Huang, 2004). Other researchers build on the latter, but state that "ability" is grounded in a sound knowledge base. As such, we cannot discuss ability without stressing the central position of knowledge acquisition in the context of elementary education (Wang, 2004). Nevertheless, a strong movement remains active that strives for an assessment reform changing a knowledge-orientation to an ability-orientation (Zhong, 2006). Although new assessment approaches and new instruments have been introduced in elementary education, the traditional paper-and-pencil assessment of performance is predominant in China (Cui, 2010). This neglects a focus on complex performance that goes beyond mere knowledge assessment and opens ways to study ability. As an example of the way to move forward, Chinese researchers point at the PISA approach of assessment that studies students' ability to formulate, employ and interpret mathematics in a variety context; thus going beyond the assessment of knowledge itself (OECD, 2010).

The former discussion can also be approached from a different perspective. We can briefly study the history of assessment and adopt current trends in our own assessment approach. At the beginning of the 20th century, Binet and Simon distinguished between three types of assessment. First, they distinguish a medical approach focusing on physiology and pathology. Second, they recognize a pedagogical approach stressing the knowledge base. Third, they refer to the psychological approach that tries to build on direct observations of intelligence (Binet & Simon, 1908/1961). As to the latter, they claim to study "pure" individual intelligence excluding the impact of instruction. No doubt, Binet and Simon's idea is a historical milestone in the assessment and measurement traditions. For instance, Fiske and Butler (1963) were proud to present a "pure" intelligent tests that is more stable than scholastic performance tests, and independent of other environmental influences. Intelligence tests were clearly set apart and aimed at measuring a subject's maximal performance or ability (Cronbach, 1949). More and more intelligence tests appeared that aimed at studying the structure of this underlying ability; for example, Wechsler Intelligence Scale for Children (WISC) (Wechsler, 1949), Cognitive Ability Test (CAT) II - UK (Thorndike, Hagen & France, 1986). With the development of specific intelligence tests, researchers also start to reflect on the relationship between intelligence and the results of scholastic test. Research points at clear correlation between academic performance and intelligence (IQ). Correlations are reported to be on average.50 (Baade & Schoenberg, 2004; Brody, 1997; Petrill & Wikerson, 2000). General cognitive abilities (*g*) are shown to be related to scholastic achievement (Frey, & Detterman, 2004). This brief discussion affects the above discussion about the nature of mathematics performance and its measurement. In our studies, we aim at studying the academic outcomes of mathematics learning processes. In addition, we aim at studying/estimating the mathematics abilities of our research subjects.

To conclude, in the present study, we start from the debate about knowledge and ability when constructing a new mathematics test. From a pedagogical perspective, we construct a test that reflects the different components of the mathematics curriculum. From a psychological perspective, we adopt an approach that makes inferences about the underlying abilities of Chinese primary school children. More details about the construction of this new test are provided in Chapter 2.

3.2 Variables related to the mathematics performance

Previous theoretical models already present a variety of variables structured at different levels. However, the amount and variety of variables is so large that it is difficult to decide which to incorporate in a particular new model. In his review about "what works" in teaching and learning, Carpenter (2000), for example, states that on average 36 new "good ideas" are published per year per journal between 1987 to 1997.

To direct our framework development, we start from the key observation that education is embedded in the local culture and how this affects local curricula, local teaching approaches, and local learning processes. This implies that we build on available empirical research about mathematics learning and performance, set up in the Chinese context; both by the researchers in or outside China. Next, building on research about differences between Chinese learners and learners with another cultural background, we incorporate studies aiming at explaining these differences in performance of learners in primary school. Lastly, we will select variables for our conceptual framework on the base of the available theoretical grounding, the extent to which they have been linked to educational interventions, and the extent they are grounded in international and national studies. The result of this specific analysis of the literature will be a structured list of variables and processes that are expected to be of relevance for studying mathematics learning and learning performance in the Chinese primary school context.

The analysis of the literature in the following sections offers a comprehensive overview of studies about variables that contribute to Chinese mathematics performance. Content analysis is used as a method to screen research articles published between 1950 and 2011. Our aims with this analysis are: (1) to identify articles related to the Chinese mathematics learning performance; (2) to give an overview of the trends in the studies over time and to present the attributes of these studies, (3) to compare the different variables studied in these articles, (4) to choose particular variables worth to be incorporated in our own conceptual framework and the studies reported in this PhD dissertation.

3.2.1 Method

3.2.1.1 Quantitative content analysis

Content analysis is a method developed in the social sciences; in particular in the field of mass communication studies (Berelson, 1952). It has been defined as "a research technique for the objective, systematic, and quantitative description of the manifest content of communication" (Berelson, 1952, p. 18). It is used to study messages in mass media and other sources (Krippendorff, 2004). Quantitative content analysis aims to "identify and count the occurrence of specified characteristics or dimensions of texts, and through this, to be able to say something about the messages, images, representations of such texts and their wider social significance" (Hansen, Cottle, Negrine & Newbold, 1998, p. 95). In a quantitative content analysis, frequencies are used to present and understand trends by extracting categories (Altheide, 1996).

3.2.1.2 Procedure

Search Strategy. A multistage process was used to identify relevant articles by building on the following keywords: "performance", "achievement", "outcome", "result", "output" and "productivity" referring to the student mathematics learning. The search was carried out in international and in national (Chinese) scientitifc databases. In a first step, we developed this sufficiently comprehensive set of search terms to be able to collect the relevant studies about Chinese mathematics education. The search involved the usage of the following electronic databases: (1) ISI Web of Science and ERIC at OVID; (2) the China Knowledge Resource Integrated Database (CNKI) by using the terms "shuxue" and "chengji" or "shuxue" and "chengjiu" in "topic" and "abstract". After deleting overlapping articles, we obtained a list of 817 citations in the international database and 687 citations in the national database. We imported all citations in Endnote to manage the coding.

Article selection. Secondly, inclusion and exclusion criterion are applied to further identify relevant articles. Firstly, the title and the abstract were reviewed. Studies remained included when meeting the following criteria:

- (a) the primary focus of the study is on mathematics education, involving Chinese students in China or comparative studies between China and other countries;
- (b) the studies focus on processes/variables in relation to the learners.

As a result, 573 international articles and 468 national article were selected for further examination. After this first scrutinizing effort, three criteria were applied set to select the 1041 articles:

- (c) the studies focus on students' mathematics outcomes or mathematics development;
- (d) the report is about a quantitative study that is reported with sufficient statistical detail;
- (f) participants belong to grade 1 to 6 in primary school. Adopting this criterion resulted in an extreme drop in the number of relevant articles (See Figure 8); therefore this last criterion was dropped.

The former procedure resulted in a data set of 110 articles from the national database, and 120 articles from the international database.

Article Review. In a next step, the actual content analysis was set up to explore the characteristics of each article. At the beginning, descriptors were defined to map in detail the range of ideas, approaches, … adopted when studying mathematics performance in of Chinese students during the past fifty years. The following information was recorded for each article: descriptions about the articles (publication year, journal title, etc.), region of the article (national or international), school level of the sample involved in the research, research methods (quantitative analysis or other), the specific variables explored in the articles. In a next stage, the characteristics of each variable being studied were scrutinized and coded. Lastly, based on the frequency of occurrence and as they could be linked to comprehensive frameworks, variables were selected and integrated into a conceptual framework for the present Phd study.

3.2.2 Results and discussion

3.2.2.1 Analysis of studies about mathematics performance published during the last fifty years

3.2.2.1.1 Trends in international and national studies

The analysis points out that we observe a clear increase in the interest for studying variables affecting mathematics performance in China. Figure 7 illustrates the total number of studies that meet our first three criteria, set out over time. The number of the national articles increases from 1957 to 2011, while the number of international articles stays nearly constant. Cross-tabulation reveals a significantly change over time when comparing publication output for subsequent decades (x^2=264.448, *df*=45, *p*<.001). From 2002 on, the number of the national articles exceeds the number of the international articles. More and more Chinese researchers study and publish about mathematics performance of learners in the Chinese setting. A t-test (*t*=24.539, *df*=621.664, p<.001) shows that the number of 573 national articles significantly exceeds the 468 international articles. However, after applying the selection criterion that further centers on quantitative studies, the t-test (*t*=7.915, *df*=150.101, *p*<.001.) is no longer significant (110 national articles versus 120 international articles).

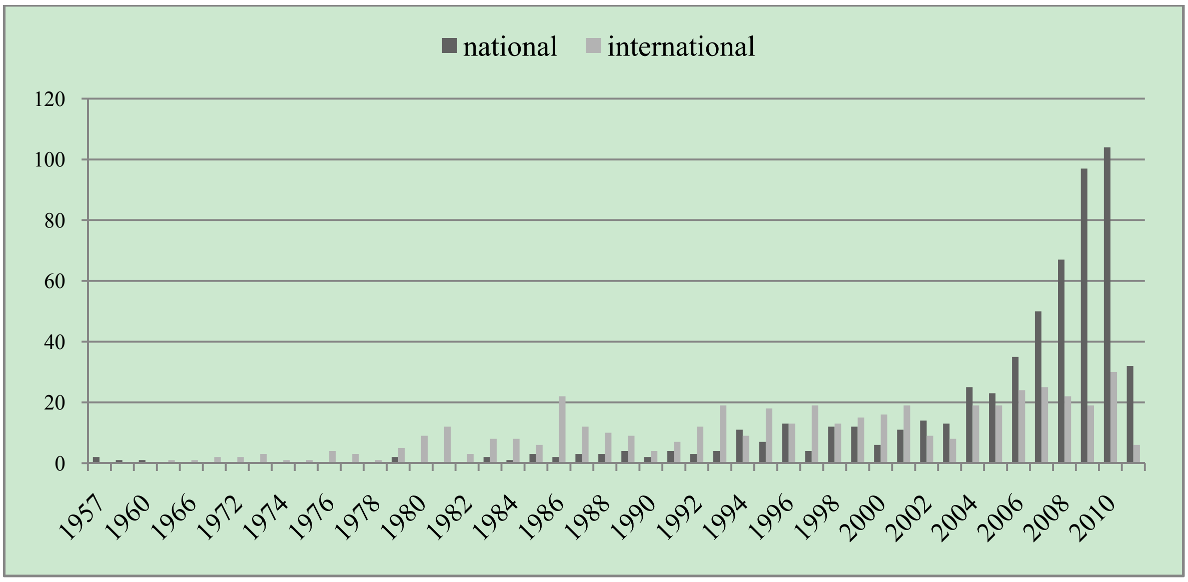


Figure 7
Number of national and international studies over time (1957-2011) (*n=1,041*).

3.2.2.1.2 Differences in the research samples between international and national studies

When focusing on the 230 quantitative studies, we understand that the objective of the national and international researchers might be different. National articles report about the exploration of local variables that are related to mathematics performance. The primary aim seems to be to improve mathematics teaching and resulting mathematics performance. International articles aim at comparing differences in learning and performance of Chinese and local students. As can be derived from Table 1, 89 articles (about 74.17%) compare Chinese students with other students in internationally published articles, while only 31 articles are published exploring models between variables and mathematics performance focusing solely on Chinese students.

Table 1

Nature of samples involved in studies from national and international research

Decade	**National**	**Studies**	**International Studies**	
	Single sample	Comparative sample	Single sample	Comparative sample
Before 1970s	0	0	0	1
1970s	0	0	0	2
1980s	3	1	3	6
1990s	9	0	8	37
2000s-	96	1	20	43
Total	108	3	31	89

We also perceive differences in the age groups, being focused upon by national and international researchers ($F_{(5,229)}$=134.839, p<.001, η^2=.784). As shown in Figure 8, the distribution of the age group is different in national and international articles. The majority of the international articles (87 articles, among 72.5%) focus on the primary school, while the majority of the national articles focus on the post-secondary education (38 articles, among 34.55%). However, the attention paid to local students at primary school, junior school, senior school and cross-stage schooling level is equal in national articles; while less attention is paid to older Chinese students. Compared to international studies, national researchers pay higher attention to higher school levels, while international researchers focus more on younger Chinese learners. Ninety-one international studies and eighteen national studies focus explicitly on the relationships between the variables predicting mathematics performance at primary school level. In the next section, a further analysis of these variables in relation to mathematics performance will be carried out; this will in part be based on the entire sample of studies and partly on a subset of studies that incorporate complex models.

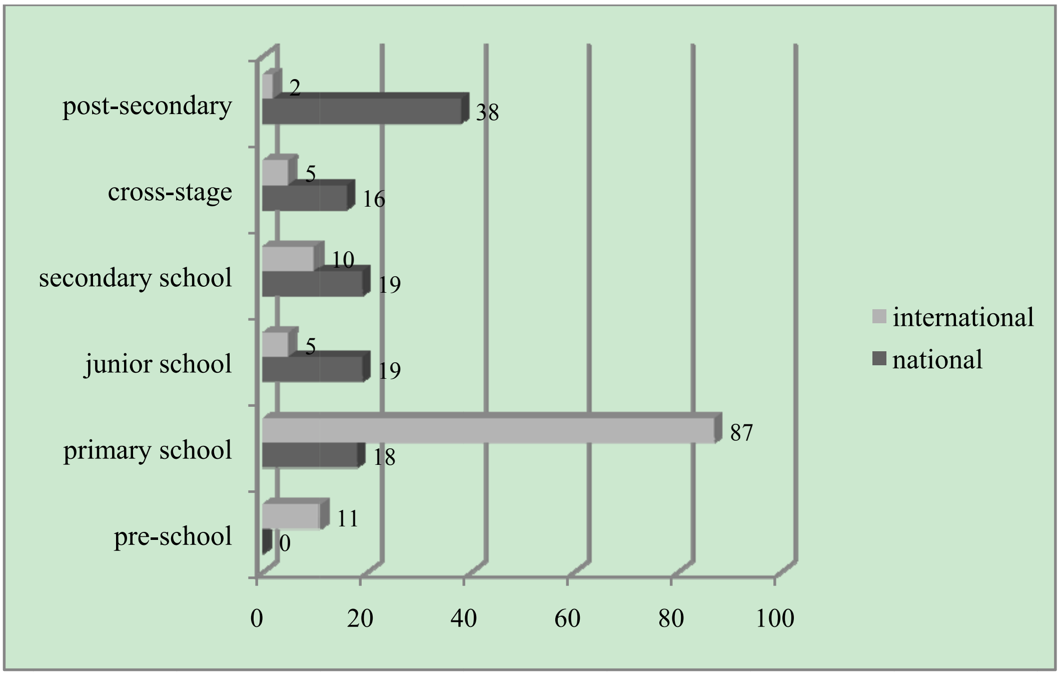


Figure 8
The distribution of attention paid to school levels in national and international journal articles.

3.2.2.2 Variables contributing to mathematics performance

On the base of a subsequent content analysis, a frequency list could be developed of recurrent variables studied in the journal articles. Frequencies for each coded variable are reported in table 2 and are in addition broken down for national and international studies. In sum, 96 different variables/processes were coded. In a number of cases, the coding can refer to partly overlapping constructs. A distinction is also made between a coding that centers on all school levels; next the coding is limited to studies focusing on primary school level. The variables which is not ranked in the top-10 are not presented here. The difference between national and international studies can been found from the Table 2. As the variables which is popular in the national studies which is not paid so much attention in the international studies. Later, one-tail Fisher test will be carried out to check the gap of the studies in national and international context. This will give us a hint to choose for the variables for our own dissertation.

Table 2

Frequency of predictor variables identified in national and international journal articles

National Variable	Freq.	**International** Variable	Freq.	**Total** Variable	Freq.
Section I - All school levels					
Total count	*110*	*Total count*	*120*	*Total count*	*230*
learning strategies	15	culture	25	culture	25
anxiety	14	teaching, classroom	14	learning strategies	20
self-efficacy	11	effort	12	anxiety	16
learning attitude	11	expectations	9	teaching, classroom	15
achievement motivation	10	homework	7	effort	14
metacognition	9	SES	7	gender	12
teacher quality	9	interaction parent-child	7	learning attitudes	12
gender	8	number sense	7	self-efficacy	11
learning interest	8	self-concept	6	expectations	11
learning motivation	7	perceived parental expectations	6	self-concept	10
		language[a]	6	learning motivation	10
				achievement motivation	10
				teacher quality[a]	10
Section II-Primary school level					
Total count	*18*	*Total count*	*91*	*Total count*	*109*
teacher quality	4	culture	19	culture	19
learning strategies	3	teaching, classroom	11	teaching, classroom	11
gender	2	effort	8	expectations	8
self-efficacy	2	number sense	7	effort	8
learning attitudes	2	expectations	6	learning strategies	7
expectations	2	SES	6	SES	7
cooperative learning[b]	2	ethnicity	5	number sense	7
		perceived parental expectations	5	ethnicity	5
		homework	5	self-concept	5
		language	5	perceived parental expectations	5
		Interaction parent-child[a]	5	homework	5
				language	5
				interactions parent-child[a]	5

*Note:

[a] When we ranked from the terms with the most mentioned time, there are some terms with the same frequence. Thus, here the top-10 coded Terms include the terms with the same frenquce until the ranking is just larger than 10.

[b] Since there are same frequency of "1" for the following variables, these variables are not mentioned here although they are ranked in top-10.

[c] The sum is not equal to the figures presented in this table because there are some figures which is not at the top of 10 which are ignored in this table. For example, the teacher quality in national studies got 5 time and is ranked at the top-10 but it is just mentioned once in international studies and got 1 time which is ignored in the Table 2.

First, as can be derived from the first section in Table 2 – focusing on all school levels - that other variables stand out in national and international research. This can partly be explained by what we already observed in table 1; international studies involve comparative samples; national studies mostly involve Chinese learners only. As a consequence, international studies dominantly refer to variables explaining the gap in mathematics performance between Chinese and other students. Culture is mostly adopted in these studies as a key variable (n = 25). But since culture is a complex concept, other variables appear that reflect sub-constructs that can be linked to the cluster concept "culture"; such as. student effort, homework, parental involvement, parent expectations and/or perceived parent expectations. Variables studied to a more limited extent in international research refer to particular abilities, such as number sense and language. In the same way, variables related to the school environment are found to a lesser extent in these international studies. And, more and more concerns are put on the teaching situation in China by the international researchers as we can see from the Table 2.

The difference between international and national studies is confirmed when we carry out a statistical analysis on the data. As reflected in Figure 9, a one-tailed Fishers' Exact Test shows significant differences in the frequencies of coded variables as observed in either national or international studies. The studies differ in relation to their emphasis on 19 of the 99 variables identified in the 230 studies. In the national studies, researchers focus to a larger extent on motivational variables, such as learning attitude, learning interests, learning motivation, self-efficacy, anxiety; learning strategies, and metacognition. However, also the teacher quality is important in the national studies. In the Chinese studies, also "classroom teaching" and "teacher quality" are emphasized. This can be linked to the strong emphasis in China on the prescriptive role that teacher handbooks play.

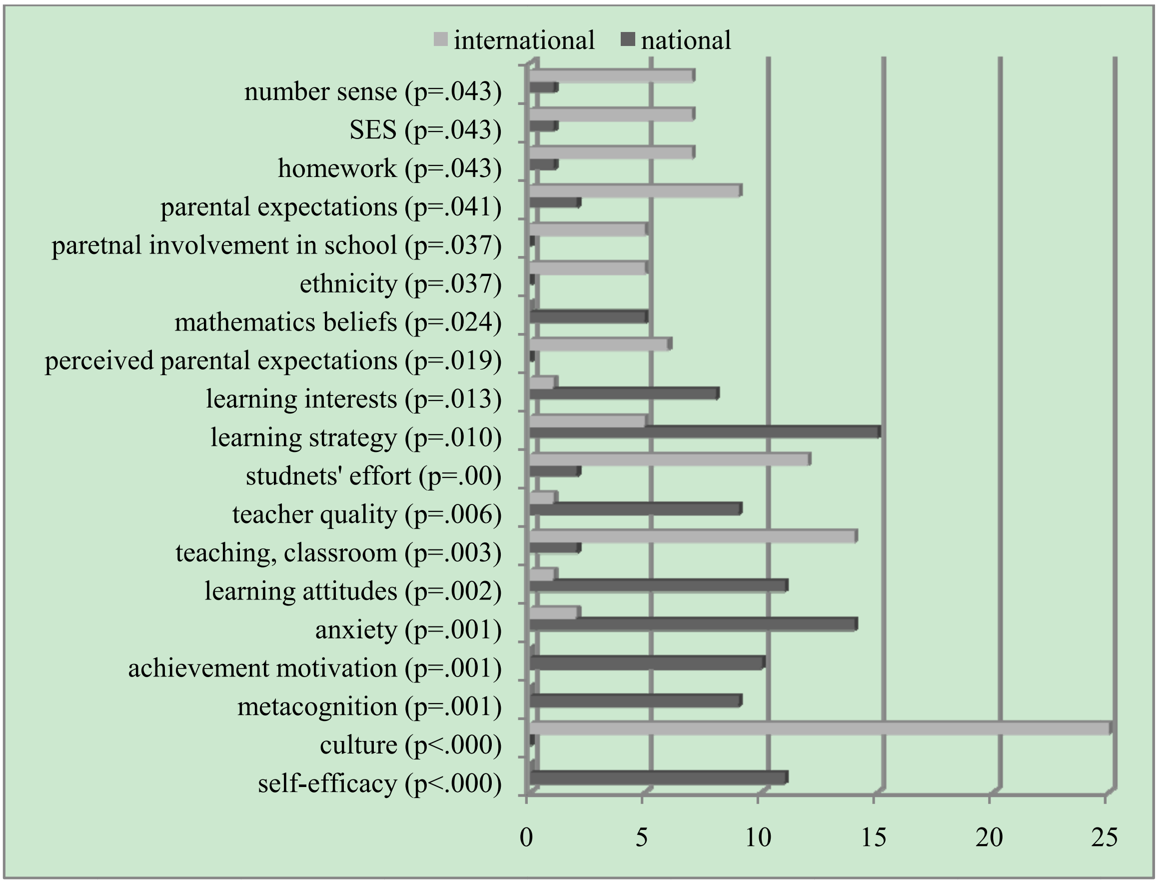


Figure 9

Fisher's Exact Test results in relation to differences in the occurrence of predictor variables in international and national studies.

When we turn our attention to section II in table 2 that presents the content analysis results of studies involving primary school children, a different picture emerges. The results of one-tailed Fishers' Exact Tests to check the differences in the frequencies of coded variables between national and international articles shows that only in relation to four variables, significant differences can be observed: culture (p=.019), time spent on textbook (p=.028), teacher quality (p=.001) and self-efficacy (p=.026). We also observe that - in the national studies - the variables anxiety, learning interest, learning motivation and metacognition do no longer reflect high frequencies in the ranking. Primary school children seem to be different from older children that face high-stake mathematics tests. The related research models therefore do not incorporate these particular affective variables. School level might also affect the stronger weight put on expectations, parent expectations, and perceived parent expectations in studies involving older learners. In the Chinese context both parents and children have high expectation as to attaining a higher level of schooling (called “wang zi cheng long”, meaning “to become a dragon”, someone important).

3.2.2.3 Summary of the variables contributing to mathematics performance and choices made in the context of our studies

As stated earlier, the aim of our content analysis of earlier mathematics research was to identify key variables to be included in a conceptual framework that helps to describe and explain mathematics learning performance of Chinese primary school learners. Table 3 brings together all variables that resulted from the content analysis; also those that were only observed to a very limited extent (n < 2). These variables have been clustered on the base of the clusters, already distinguished in Geary’s model (2005). In addition, they are clustered in line with the levels that were already found in Creemers’ model. The numbers between brackets refer to the attention paid to this particular variable in – first – all studies and – second – in primary education research. Variables printed in bold represent variables that are incorporated in our studies. Variables followed by an asterix (*), refer to variables were stressed in particular in comparative – international - studies.

The overview of the variables actually selected to be incorporated in our studies, immediately makes it clear that also a number of variables have not been selected and considered in this PhD study. Though we aimed at (1) selecting variables at different levels (student (individual and family), class, and school level and (2) selecting of variables that represent different clusters in Geary’s model, we stress that choices had to be made. These choices were influenced by the resources available to gather data from the target group (number of researchers involved, time and budget), the fact that gathering data about a larger set of variables would impose too high demands on the target group (available time, attention focus of respondents, fatigue); and the fact that some variables are too complex to measure in a large scale study based on classroom group assessment (e.g., math anxiety, learning strategies, …).

Table 3 Overview of relevant variables to be incorporated in the dissertation

Individual level - student	Individual level - family	Classroom level	School level
Biological primary variables: **gender (12,3),** **IQ (8,3),** **age (1,1),** **scholastic ability (1,1)** **initial performance (6,1),** initial experience at home/school (2,1), **right/left handed (1,0),** weight (1,0),	**Primary variables:** **ethnicity (5,5)*,**	**Primary variables:** **textbook (4,2),** discipline of math (3,0),	**Primary variables:** **culture (25,19)*,** **language (8,5),** **schooling/education (4,3),** **number language (3,2),** **urban or rural (4,0),** school environment (2,0),
Biological secondary variable: **number sense (8,7)*,** **number retrieval (1,1),** number estimation (1,1), number representation (1,1), categorical ability (1,1), base-ten number (1,0),	**Secondary variable:** **SES (8,7)*,** **cultural capital (2,1),** **literacy (1,1),** **parent educational level (2,0),**	**Secondary variable:** **teacher quality (10,4)*,**	**Secondary variable:** **school type (1,1),** Physical environment (2,2),
Cognitive variables: **metacognition (9,1)*,** mathematics cognition (2,0), thinking style (2,0), working memory (1,0),	**Cognitive variables:**	**Cognitive variables:**	**Cognitive variables:**
Psychological variable: self-concept (10,5), learning attitudes (12,3)*, attributions (6,4), math anxiety (16,2)*, **self-efficacy (11,2)*,** mathematics beliefs(5,1)*, learning self-confidence (3,1), academic self-concept (1,1), peer acceptance (1,1), socialization (1,1), perceived scholastic competence (1,1), perceived performance (1,1), self-esteem (2,0)*, self-adaption (1,0), self-explore (1,0),	**Psychological variable:** mother's evaluation of childrens' competence (2,2), parental concerns (2,1), satisfaction (1,0),	**Psychological variable:**	**Psychological variable:**

spiritual (2,0), personality (1,0), learning choice ability (1,0), non-IQ (1,0),			
Affective variables: learning attitudes (12,3)*, aggression (1,1), emotion (4,0), defensive pessimist (1,0), self-handicapping (3,0),	**Affective variables:**	**Affective variables:**	**Affective variables:**
Motivational variables: **perceived parent expectations (6,5)*,** learning motivation (10,3), learning interests (9,2)*, goal (5,1), achievement motivation (10,1), value of the mathematics (2,0), child's willing to learn(1,0),	**Motivational variables:** **expectations (11,8)*,**	**Motivational variables:**	**Motivational variables:**
Behavior strategies: **effort (14,8)*,** learning strategies (20,7), **homework (8,5)*,** **time spent on learning (6,3),** self-regulative learning (4,0), self-control (4,0), learning methods (2,0),	**Behavior strategies:** interaction between parental and child (9,5), parental method (4,4), parental formulation (4,4), parental involvement in school (5,3), **time of parents used for teaching children to learn (3,3),** support (4,2), environment of family (1,0),	**Behavior strategies:** **teaching, classroom (16,11)*,** place value (2,2), cooperative learning (2,2), **teacher questioning (2,1),** **time spent on the textbook (1,1),** **teacher feedback (4,1),** ICT usage (1,1), classroom climate (3,0),	**Behavior strategies:** leadership (1,1), school autonomy (1,1), project of "school merger" (1,1),
Others: beauty of math (1,0), difficulties of test (1,0), practical work (1,0), simulation math ability (1,0),			

Note: Numbers between brackets refer - first - to the occurrence of this variable in all studies and – second – the attention paid to this variable in primary school research.

3.2.3 The conceptual framework

Based on the content and structure of the models presented in section 2, and the results of the content analysis summarized in section 3, we can delineate the conceptual framework adopted in our dissertation. In the next eight chapters, particular variables in the conceptual framework will be discussed in more detail and a substantive analysis of the literature will be presented to study these variables in relation to mathematics learning and resulting performance. In the next paragraphs, we limit our discussion to a first short positioning of these variables.

3.2.3.3 Students - Individual level

At the individual student level, primary biological variables play a role: age (Salili & Hau,1994), schooling age (Kyriakides & Luyten, 2009), gender (Wang, 2006), intelligence (Lynn, 2008), left/right handed (Zang, et al, 2008), initial performance (Marsh & Hau, 2002) have been found to affect mathematics performance in primary school. In addition, we can add a students' number facility ability related to basic mental calculations (Geary & Salthouse, 1996).

At the individual level, we can add meta-cognition, playing an important role in view of mathematics performance (Desoete, Roeyers, & Buysse, 2001). Also, students' self-efficacy has been shown to predict mathematics performance (Stevens, Olivarez, Lan , & Tallent-Runnels, 2004).

In relation to a motivational and behavioral strategies dimension, studies have revealed that perceptions about parents' expectations affect student motivation and result in an impact on mathematics performance (Mau, 1997). The perceived control of the learning context, will influence the effort students spend on homework, on time for learning in general and time for learning mathematics (Stevenson, Lee et al, 1986).

3.2.3.4 Parents - Individual level

The socio-economic status of the parents has been found to have – though sometimes weak - relationship with mathematics performance in primary school (Liu & Ke, 2008). However, research shows that SES is difficult to be measured (Sirin, 2005; White, 1982). Definitions of SES seem to differ widely; for instance ethnicity is not always considered to be a part of SES (Peverly, 2005). In the context of this dissertation, attention will be paid to a clear operational definition of SES. For instance, next to parents' educational level, all information about parents' job level and family wealth indicators will be included in our studies.

3.2.3.2 Teachers - classroom level

Research suggests that teacher quality is an important predictor of mathematics performance of learners (Stevenson, et al, 1990). Teacher quality is again a complex concept. In the context of this dissertation, next to teacher gender (Beilock, et al., 2010), also age of the teacher, years of experience in teaching, graduation level, diploma level, career level, experience in teaching a subject and teaching a particular grade, beliefs about mathematics are considered as relevant sub-concepts of teacher quality.

Next to these background variables, actual teaching behavior is also studied (House, 2002), such as the questioning approach in the classroom (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999), and the nature of feedback given to the learners (Salili & Hau,1994).

3.2.3.1 Culture school level

The dominant language in the cultural setting (Imbo & LeFevre, 2009), the number language adopted in the school(Geary, et al., 1993), and the educational syllabus adopted by teachers in the school will be considered as a primary set of operational ways to study culture at the school level and how this influences mathematics performance.

In addition, schools also differ in other ways. They are e.g., set in a rural or urban setting (Wang & Li, 2009), or they are located in a region with a particular gross-domestic product (GDP). Also, the school type (Peverly, 2005) can differ. This is stated to be - in the setting of developing countries – to be an important predictors of mathematics performance. School types differ in the number of teachers, ratios of the different level teachers, proportion teachers vs. number of students, time allocated for actual teaching.

4. Research design and overview of the dissertation

4.1 Research objectives

As stated earlier, three main research objectives direct the studies presented in this PhD study. The first research objective can be considered as a preliminary objective. In relation to the three research objectives, we present key research questions.

Research objective 1: the construction of a standardized assessment instrument to study mathematics performance of Chinese elementary school children.

Research question 1: What is the reliability and validity of a new mathematics test that has been developed building on item response theory?

Research objective 2: to exam the most important predictors of mathematics learning performance in Chinese primary schools. Four research questions are addressed to attain this objective:

Research question 2: What are the strongest predictors of mathematics performance at the school level, class level and individual student level?

Research question 3: How do individual student variables moderate between context variables and mathematics learning performance?

Research question 4: What is the relationship between family variables and mathematics learning performance ?

Research question 5: How do teaching approaches in the classroom contribute to students' mathematics learning performance?

Research question 6: How do teacher, parents and students related variables compensate for a disadvantaged learning environment; with an emphasis on different types of homework assignments?

Research objective 3: to exam the predictors of the students with learning problems in mathematics. This objective is approached with two research questions:

Research question 7: Which variables are significant predictors for learning difficulties in the Chinese context?

Research question 8: How do students – of different learning abilities levels - perform on a variety of mathematics tasks (e.g., fact retrieval, basic numerical exercises)?

Table 4 presents an overview of the research questions, the research methods adopted in the different studies, the variables being focused upon and the corresponding objectives.

Table 4 Research questions, research design, research methods and output for the different objectives.

Research objective	Research questions	Research Variables	Research methods	Output
Objective 1				
- Instrument	RQ1: What is the reliability and validity of a new mathematics test that has been developed building on item response theory?	- Mathematics syllabus - Mathematics ability	Item response Theory (IRT) Pilot study (N=3,002) Main Study (N=10,959)	Chapter 2
Objective 2				
- General	RQ2: What are the strongest predictors of mathematics performance at the school level, class level and individual student level?	- School level variables - Class level variables - Individual level variables	Multilevel analysis Main study (N=10,959)	Chapter 3
- Student	RQ3: How do individual student variables moderate between context variables and mathematics learning performance?	- Student's characteristics - Academic variables - Non-academic variables	Structure equation model Pilot study (N=1,749)	Chapter 4
- Family	RQ4: What is the relationship between family variables and mathematics learning performance ?	-SES	Multilevel analysis Main study (N=10,959)	Chapter 5
- Class	RQ5: How do teaching approaches in the classroom contribute to students' mathematics learning performance?	- Macro-analysis: interaction - Micro-analysis: questioning	Mixed-methods (NVivo) Video (N=9), Student (N=601)	Chapter 6
- Behavior	RQ6: How do teacher, parents and students related variables compensate for a disadvantaged learning environment; with an emphasis on different types of homework assignments?	- Homework assigned (teacher-parents-students)	Loglinear analysis Main study (N=10,959)	Chapter 7
Objective 3				
- General	RQ7: Which variables are significant predictors for learning difficulties in the Chinese context?	- School level variables - Class level variables - Individual level variables	Logistic regression analysis Main study (N=10,959)	Chapter 8
- Basic ability	RQ8: How do students – of different learning abilities levels - perform on a variety of mathematics tasks (e.g., fact retrieval, basic numerical exercises)?	- Mental calculation	MANOVA Chinese student (N=7,247) Flemish student (N=913)	Chapter 9

4.2 Overview of the dissertation

The entire dissertation can be split up into two parts. The first part focuses on “normal” performing students. The second part centers on students with learning difficulties.

The first chapter provides the general overview for this dissertation. Firstly, it present a review of models describing factors affecting learning and related performance. Second, based on an extensive review of the literature published during the last 50 years, a content analysis was carried out to map critical variables studied in national and international studies that link mathematics performance to predictors.

Chapter 2 discusses the development and implementation of a standardized instrument to diagnose mathematics performance in Chinese primary schools on the base of item response theory (IRT). The validity and reliability of this new mathematics scale are explained and illustrated in detail in this chapter.

Chapter 3 examines significant predictors of mathematics performance in Chinese primary schools on the base of multilevel model analysis. Three levels are considered: individual students level variables, family level variables and school level variables. Based on the result of this chapter, several variables are selected that are explored in more detail in subsequent chapters.

Chapter 4 focuses on the relationship between individual student variables and contextual variables. Structural equation modeling is used to test the direct and indirect relationships between predictors. Student variables are found to moderate between contextual variables and performance.

Chapter 5 explores the relationship between family variables and mathematics learning performance on the base of multilevel analysis. A U-shaped relationship is found between the socio-economic status of the family and learning performance. The aggregated SES scores at the school level seem to moderate between individual family variables and mathematics performance.

Chapter 6 investigates teacher-students interactions in the mathematics classroom and links this to mathematics performance. A multiple regression analysis reveals the importance of interactions between students and the relevance of teacher-student interaction in view of mathematics performance. Questioning approaches of teachers that stress evaluation and problem-solving also seem to improve student performance.

Chapter 7 studies the impact of homework on mathematics performance. Different types of homework are considered. A loglinear analysis of the impact of teacher, parents and student assigned homework reveals that parents and students with a disadvantaged background seem to assign homework to compensate for an at-risk situation. Nevertheless, parents and students approach towards homework has only a limited effect on academic achievement.

The next two chapters focus on students with learning difficulties.

Chapter 8 investigates the effect of biological, cognitive, affective, and context variables on the being at risk of developing learning difficulties. On the base of a hierarchical logistic regression, conclusions are presented that stress the importance of individual level and classroom level variables. School level variables – such as the region's development level – seem to have a lesser impact on the likelihood of developing mathematics difficulties.

Chapter 9 examines the development pattern of students' numerical facilities, and this in low versus high achievers, and comparing Chinese and Flemish learners. A MANOVA reveals that students perform better on addition as compared to subtraction or multiplication tasks; the latter being easier than division tasks or mixed exercises. Both in Flanders and China, high achievers perform highly and stable across school levels and the variety of numerical facility tasks. Low achievers do not reflect a stable level of achievement across school grades.

In the final chapter, a general discussion is presented and a conclusion that integrates the results of all chapters. Next, theoretical and practical implications are presented on the base of these conclusions. The chapter ends with a discussion of limitations and directions for future research. Figure 10 illustrates how the different studies are linked to one another.

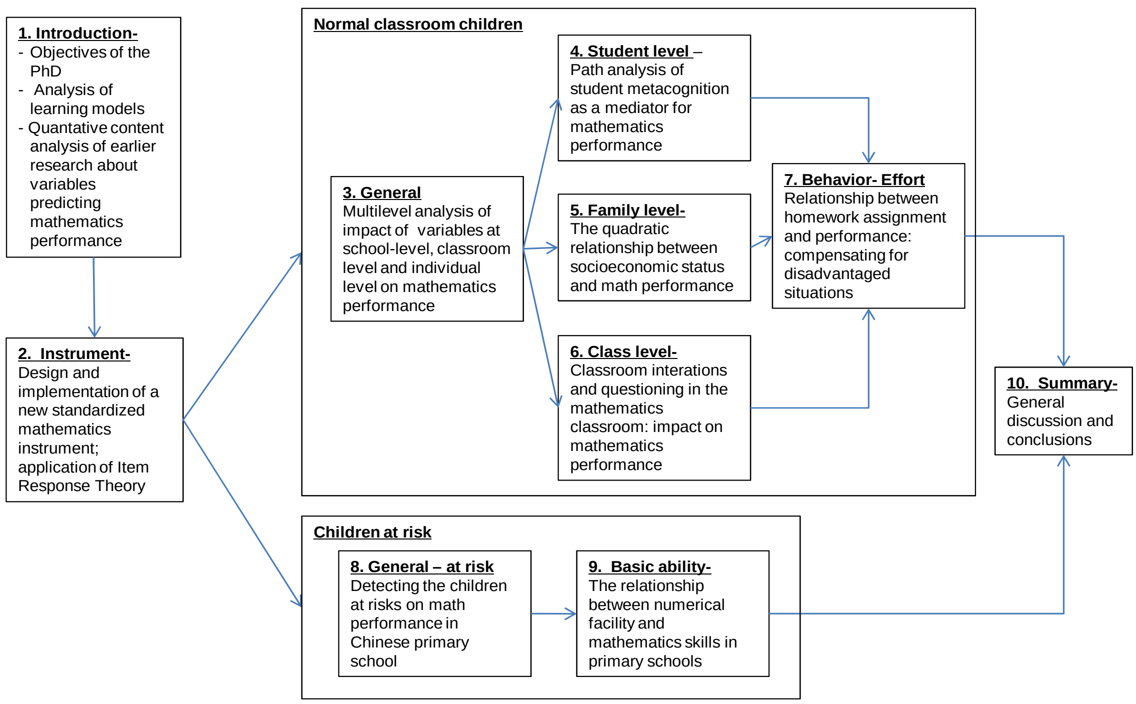


Figure 10 Schematic overview of the different chapters in the dissertation and how they are related.

Reference

Note: A comprehensive list of the 230 articles included in the content analysis study were added as an appendix at the end of the dissertation before the summary.

Altheide, D. (1996). *Qualitative media analysis*. Newbury Park, CA: Sage.

Baade, L. E., & Schoenberg, M. R. (2004). A proposed method to estimate premorbid intelligence utilizing group achievement measures from school records. *Archives of Clinical Neuropsychology, 19,* 227−243.

Beilock, S. L., Gunderson, E. A., Ramirez, G. & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *PNAS, 107*, 1860-1863.

Berelson, B. (1952). *Content Analysis in Communication Research*. Glencoe, Ill: Free Press.

Binet, A., & Simon, T. (1905). M6thodes nouvelles pour le diagnostique du niveau intellectuel des anormaux [New methods for the diagnosis of the intellectual levels of subnormals] (E. S. Kite, Trans.). In J. J. Jenkins & D. G. Paterson (Reprint Eds.), *Studies in individual differences: The search for intelligence* (pp. 90-96). New York: Appleton- Century-Crofts. (Reprinted in 1961)

Binet, A., & Simon, T. (1908). La d6veloppement de l'intelligence chez les enfants [The development of intelligence in children] (E. S. Kite, Trans.). In J. J. Jenkins & D. G. Paterson (Rprint Eds.), *Studies in individual differences: The search for intelligence (pp.* 96-111). New York: Appleton-Century-Crofts. (Reprinted in 1961)

Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist, 52*(10), 1046−1050.

Brownell, W. A. (1928). *The development of children's number ideas in the primary grades*. Chicago: University of Chicago.

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

Byrnes, J. P. & Wasik, B. A. (2009). Factors Predictive of Mathematics Achievement in Kindergarten, First and Third Grades: An Opportunity-Propensity Analysis. *Contemporary educational psychology, 34* (2), 167-183.

Carole, A. & Jennifer, A. (1988). Achievement goals in the classroom: students' learning strategies and motivation process. *Journal of Educational Psychology, 80*(3), 260-267.

Carpenter, W. A. (2000). Ten Years of Silver Bullets: Dissenting Thoughts on. Education Reform. *Phi Delta Kappan*, *81* (5): *383*-389.

Carroll, J.B. (1963). A model of school learning. *Teacher College Record, 64,* 723–733.

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, R. & York, R. (1966). *Equality of educational opportunity.* Washington DC: Government Printing Office.

Creemers, B. P. M. & Scheerens, J. (1994) Developments in the educational effectiveness research programme. *International Journal of Educational Research, 21* (2), 125-140.

Creemers, B.P.M. (2007) Educational effectiveness and improvement: the development of the field in mainland Europe. *In:* Townsend, T. (ed) *International Handbook of School Effectiveness and Improvement*, (pp. 223-242), New York, Springer.

Cronbach, L. J. (1949). *Essentials of psychological testing.* New York: Harper.

Cui, Y. (2010). Cujing xuexi: Xueye pingjia de xin fanshi [Improve the learning: new approaches for the performance assessment approaches]. *Jiaoyu kexue yanjiu, 3*,11-20.

De Jong, R. , Westerhof, K. J. and Kruiter, J. H.(2004). Empirical Evidence of a Comprehensive Model of School Effectiveness: A Multilevel Study in Mathematics in the 1st Year of Junior General Education in The Netherlands. *School Effectiveness and School Improvement, 15,* 1, 3-31.

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

Engle, P., Grantham-McGregor, S., Black, M., Walker, S., & Wachs, T. (2007). How to avoid the loss of potential in over 200 million young children in the developing world. *Child health and education, 1* (2), 67-87.

Fraser, B. J., Walberg, H. J., Welch, W. W. & Hattie, J. A. (1987). Synthesis of Educational Productivity Research. *International Journal of Educational Research, 11*, 145-252.

Frey, M.C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*, 373-378.

Geary, D. C. & Bjorklund, D. F. (2000). Evolutionary developmental psychology, Children development, 71 (1), 57-65.

Geary, D. C. (2006). Development of mathematical understanding. In D. Kuhl & R. S. Siegler (Vol. Eds.), *Cognition, perception, and language, Vol 2* (pp. 777–810). W. Damon (Gen. Ed.), *Handbook of child psychology (6 th Ed.)*. New York: John Wiley & Sons.

Geary, D. C. (2007). An integrative model of human brain, cognitive, and behavioral evolution. Acta Psychologica Sinica, 39 (3), 383-397.

Geary, D. C. (2008). Children's academic development: where evolution meets culture. Anuario de Psicologia, 39 (2), 209-223.

Geary, D. C., & Hoard, M. K. (2005). Learning disabilities in arithmetic and mathematics: theoretical and empirical perspectives. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 253-268). New York: Psychology Press.

Geary, D. C., C. C. Bowthomas, et al. (1993). Even before Formal Instruction, Chinese Children Outperform American Children in Mental Addition. *Cognitive Development,8*(4): 517-529. *

Geary, D. C., T. A. Salthouse, et al. (1996). Are East Asian versus American differences in arithmetical ability a recent phenomenon? *Developmental Psychology,32*(2): 254-262. *

Geary, D. C., T. A. Salthouse, et al. (1996). Are East Asian versus American differences in arithmetical ability a recent phenomenon? *Developmental Psychology, 32*(2), 254-262.

Geary, D.C. (2005). *The origin of mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.

Haertel, G. D., Walberg, H. J., & Weinstein, T. (1983). Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research, 53*(1), 75-91.

Hansen, A., Cottle, S., Negrine, R.,& Newbold, C. (1998) *Mass Communication Research Methods.* London: MacMillan Press.

House, J. D. (2002). Instructional Practices and Mathematics Achievement of Adolescent Students in Chinese Taipei: Results from the TIMSS 1999 Assessment. *Child Study Journal, 32*(3),157-178.

Huang, F. (2004). Curriculum reform in contemporary China: seven goals and six strategies. *Journal of Curriculum Studies*, 36 (1), 101-115.

Imbo, I. and J. A. LeFevre (2009). Cultural Differences in Complex Addition: Efficient Chinese Versus Adaptive Belgians and Canadians. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*(6), 1465-1476. *

Imbo, I., & LeFevre, J-A. (2009). Cultural differences in complex addition: Efficient Chinese versus adaptive Belgians and Canadians. *Journal of Experimental Psychology, 35* (6), 1465-1476.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.

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.

Kyriakides, L., Campbell, R.J., & Gagatsis, A. (2000). The significance of the classroom effect in primary schools: An application of Creemers' comprehensive model of educational effectiveness. *School Effectiveness and School Improvement, 11* (4), 501-529.

Li, X. H. (2006). Lun woguo yiwu jiaoyu ziyuan de gongping peizhi [Reflection on equal arrangement of educational resource for compulsory education in China]. Jianghaixuekan, 6, 117-120.

Liu, X. F. and Ke, L. (2008). Student performance and family socioeconomic status: Results from a survey of compulsory education in Western China. *Chinese Education and society, 41*(5), 70-83.

Lynn, R. (2008). IQ and mathematics ability of Tibetans and Han Chinese. *Mankind Quarterly, 48*(4), 505-510. *

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.

Marsh, H. W., K. T. Hau, et al. (2002). "Multilevel causal ordering of academic self-concept and achievement: Influence of language of instruction (English compared with chinese) for Hong Kong students. *American Educational Research Journal, 39*(3),727-76. *

Mau, W. C. (1997). Parental influences on the high school students' academic achievement: A comparison of Asian immigrants, Asian Americans, and white Americans. *Psychology in the Schools, 34*(3): 267-277. *

McGrew, K. S. (2007). Beyond IQ: A Model of Academic Competence and Motivation (Institute for Applied Psychometrics). Retrieved January 3, 2008, from http://www.iapsych.com/acmcewok/map.htm

McIlrath, D., & Huitt, W. (1995). The teaching-learning process: A discussion of models. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved [2001], from http://chiron.valdosta.edu/whuitt/papers/modeltch.html

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

OECD (2010a), PISA 2009 *Results: Learning Trends: Changes in Student Performance Since 2000* (Volume V). From http://dx.doi.org/10.1787/9789264091580-en.

OECD (2010b), *PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I). http://dx.doi.org/10.1787/9789264091450-en*

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.

Petrill, S. A., & Wilkerson, B. (2000). Intelligence and achievement: a behavioral genetic perspective. *Educational Psychology Review, 12*(2), 185−199.

Peverly, S. T. (2005). Moving past cultural homogeneity: Suggestions for comparisons of students' educational outcomes in the United States and China. *Psychology in the Schools 42*(3), 241-249.

Plowden Committee (1967). *Children and their primary schools.* Londen: HMSO.

Reynolds, A. J., & Walberg, H. J. (1992). A process model of mathematics achievement and attitude. *Journal for Research in Mathematics Education, 23*(4), 306-328.

Salili, F. and K. T. Hau (1994). The Effect of Teachers Evaluative Feedback on Chinese Students Perception of Ability - a Cultural and Situational Analysis. *Educational Studies 20*(2), 223-236. *

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

Siegler, R. S. (1996). Emerging minds: The process of change in children’s thinking. New York: Oxford University Press.

Siegler, R. S., & Mu, Y. (2008). Chinese children excel on novel mathematics problems even before elementary school. *Psychological Science, 19*, 759-763.

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

Stevens, T., Olivarez, A., Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *The Journal of Educational Research, 97*, 208-221.

Stevenson, H. W. and et al. (1990), Mathematics Achievement of Children in China and the United States. *Child Development, 61*,1053-1066. *

Stevenson, H. W., S. Y. Lee, et al. (1986). Mathematics Achievement of Chinese, Japanese, and American-Children. *Science, 231*(4739), 693-699. *

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.

Tang, Y., Zhang, W., Chen, K., Feng, S., Ji, Shen, Reiman E. M. & Liu, Y. (2006). Arithmetic processing in the brain shaped by cultures. *PNAS, 103*, 10775-10780.

Thorndike, E. L. (1922). *The psychology of arithmetic*. New York: Macmillan.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: Vol. 2. Estimation of magnitudes. *Psychological Review, 8*, 384–394.

Thorndike, R. L., Hagen, E., & France, N. (1986). Cognitive Abilities Test Second Edition: Administration Manual. Windsor: nferNelson.

Walberg, H. J., Haertel, GD, Pascarella, E., Junker, LK, & Boulanger, FD (1981). Probing a model of educational productivity in science with national assessment samples of early adolescents. *American Educational Research Journal*, *18*, 233-249.

Walberg, H. J., Pascarella, E., Haertel, G. D., Junker, L. K., & Boulanger, F. D. (1982). Probing a model of educational productivity in high-school science with national assessment samples. *Journal of educational psychology, 74*(3), 295-307.

Wang, C. (2004). Renzhen duidai qingshi zhishi de jiaoyu sichao: zaiping you yingshi jiaoyu xiang suzhi jiaoyu zhuangui tifa de taolun [A Critical Reflection on the Thought of "Despising Knowledge" in Chinese Basic Education]. Peking University Education Review, 2(3), 5-23.

Wang, J. (2006). An Empirical Study of Gender Difference in the Relationship between Self-Concept and Mathematics Achievement in a Cross-Cultural Context. *Educational Psychology, 26*(5), 689-706. *

Wang, J. and Y. Li (2009). Research on the Teaching Quality of Compulsory Education in China's West Rural Schools. *Frontiers of Education in China, 4*(1),66-93. *

Wang, J.J. (2007). A trend study of self-concept and mathematics achievement in a cross-cultural context. *Mathematics Education Research Journal,19*(3), 33-47.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1997). What do we know: widely implemented school improvement programs (Report No. MF01/PC01). Philadelphia: Mid-Atlantic Lab for Student Success.

Wechsler, D. (1949). The Wechsler Intelligence Scale for Children. New York: Psychological Corp.

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

Zang, Y. F., Z. X. Zang, et al. (2008). Effects of hand clasping and arm folding on academic performance. Neural Regeneration Research, 3(8): 914-917. *

Zhong, Q. (2006). Curriculum reform in China: Challenges and reflections. *Frontiers of Education in China. 3* (1), 370-382.

Chapter 2

A standardized instrument to diagnose mathematics performance in Chinese primary education: Application of item response theory*

Abstract

Monitoring mathematical performance requires instruments that fit the curriculum and are sensitive to track learner progress. The present study aims at developing a standardized instrument to diagnose primary mathematics performance in China. In a pilot study, test items were developed and tested by experts and teachers within all primary school grade. In a main study, the items were evaluated, involving 10,959 students from schools in five different Chinese regions. Confirmatory factor analysis presented uni-dimensionality of test forms. Prosperities of items were examined by 2-parameters Item Response Theory model of BiLog-MG3. A good model fit was achieved for a set of 386 items (G^2 = 1378.22, p = 1.000). Item information curves and test information curves revealed the items covered a satisfactory range of abilities and provided reliable information about students' abilities.

1. Introduction

The importance of mathematical literacy cannot be overemphasized in modern society (Grégoire & Desoete, 2009). The development of mathematical literacy is considered to be critical for all students during compulsory education. Currently, it is difficult to diagnose mathematics ability or monitor the quality of primary mathematics performance across China because few established mathematics scales are available. Monitoring student progress and educational quality of mathematics education requires reliable instruments with sustainable characteristics that fit the Chinese new curriculum requirements. In this paper, we present a new item pool constructed on the base of item response theory (IRT) that is expected to contribute to the measurement and assessment of mathematics performance and might be helpful to monitor educational quality.

1.1 Available mathematics performance instruments

* This chapter is based on the submitted paper of Zhao, N.N., Valcke, M., Verhaeghe, J., & Desoete, A. (submitted). Modeling a standardized instrument to diagnose mathematics performance in China. *Asia Pacific Journal of Education.*

In the Chinese context, we observe two popular approaches to monitor student performance and/or to diagnose the learners' performance and detect learners with mathematics learning difficulties: the use of criterion-referenced tests and norm-referenced tests (See Table 1).

A typical example of a Chinese criterion-referenced test is the *Learning difficulties academic performance test for grade 6 of primary school* (Liu, 2004). In Taiwan, regular use is made of the *Mathematics Diagnosis Test on Primary School Primary-grade Students* (Qin & Wu, 1996; Zhou, 1996) and the *Mathematics Diagnosis Test* (Yu, Lin, & Cai, 2001).

Secondly, some of the popular norm-referenced scales in mainland China were derived by Chinese scholars from international intelligence scales. Additional scales were developed by Chinese scholars (See Table 1) to assess mathematics learning difficulties. Most scholars adopt the definition of learning difficulties from the World Health Organization (WHO)'s International Classification of Diseases-10th Revision (ICD-10) or, the definition of National Association of Community Health Centre from America. During recent years, some scholars began to integrate the Das and Naglieri's Cognition Assessment System (1994, 1997) into their diagnosis of learning difficulties which was developed from Luria's theory (Deng, Zuo, Li, & Das 2007; Zuo & Xi, 2006).

As can be derived from Table 1, most of these instruments are useful to diagnose mathematics cognitive abilities or help to provide information about students difficulties in the relation to mathematics cognitive skills. But, we also observe some limitations: (1) most instruments test cognitive abilities rather than scholastic skills (Frey & Detterman, 2004). Therefore, they can hardly be used by primary school teachers due to a lack of fit to the math curriculum. In addition, test administration materials (e.g., manuals) are not freely available to school teachers. (2) Most instruments center of a sub-population of the primary school or only center on a sub-domain of the math curriculum. Most norm-referenced scales can't be used with school children whose age are outside the available norm-tables. Also, the norms also reflect the abilities of samples involved during scale development. Moreover, especially the criterion-reference instruments are not adapted to the new mathematics curriculum syllabus "*Mathematics Curriculum Standards in the Phase of Full-time Compulsory Education (Experimental Manuscript)*" in 2001. As a consequence, they are less relevant.

Table 1

Summary of the Instruments Used in China for mathematics performance diagnosis

	Researchers	Year	Revised Instruments	Original Instruments	Met-hod
Revised Scales	Lin and Zhang	1986	Wechsler Intelligence Scale for Children-Chinese Revised (WISC-CR)	WISC (Wechsler, 1949)	CTT
	Wang and Qian	1987	Combined Raven's Test (CRT)	Raven's Standard Progressive Matrices (SPM) (Raven, 1938; Raven, Count and Raven, 1983, 1988)	CTT
	Wang and Qian	1997	Combined Raven's Test-C2 (CRT-C2)		CTT
	Wang, Di and Qian	2007	Combined Raven's Test –C3(CRT-C3)		
	Fan	1988; 1989	Children Developmental Scale of China- Infant Mental Scale (CDCC)	Bayley Scales of Infant Development (BSID) (Bayley, 1969, 1993)	CTT
	Jin and Li	1991	Student Group Intelligent Quality Test(SGIT)	Test Your IQ (Munzert, 1991)	CTT
	Jing, Yu and Deng	1995	Pupil Rating Scale - revised screening for learning disabilities (PRS-R)	Pupil Rating Scale (PRS) (Myklebust 1971)	CTT
	Zeng	2002	Learning Disability Evaluation Scale (LDES)	Stephen and McCarney (1996)	CTT
	Researchers	**Year**	**Scale**	**Organization**	
Local Scales	Lv	1991	Diagnosing Scale of Cognitive Ability for Children (DSCAC)	Hangzhou University	CTT
	Zhang Zhou & Zhang	1992-1994 2005	Children Developmental Scale of China (CDCC)	Beijing Normal University	CTT
	Shanghai institution of educational research	1989-1994	Mathematics Learning Difficulties Test	Institution of educational research of Shanghai city	CTT
	Beijing institution of educational research	1997	The Middle School Mathematics Learning Difficulties Test	Institution of educational research of Beijing City	CTT
	Shao, Chen and Shan	2000	Learning Disability Diagnosis Scale (LDDS)	East China Normal University	CTT

1.2.The characteristics of a new diagnostic instrument

1.2.1 Focus on both the scholastic assessment and a cognitive ability test

For the above mentioned reasons, a new test was developed covering the new mathematics curriculum as well as reflecting a validated conceptual model with twelve mathematical building blocks. In the next section, we describe in more detail the development of this instrument. Firstly, item development took into account mathematics literacy conceptions and available resources about mathematics performance. In China, the definition of mathematical literacy has changed over time in parallel with social, economic and cultural changes. We therefore based the delineation of the mathematics test content on the "*Mathematics Curriculum Standards in the Phase of Full-time Compulsory Education - Experimental Manuscript*" (2001). This approach distinguishes between three mathematical literacy domains in the primary school: number and algebra, shape and space, statistics and probability.

Secondly, items were gathered and clustered in line with twelve mathematics building blocks as distinguished in other studies (see Desoete & Roeyers, 2005; Nunes et al, 2007; Zimmermann & Cunningham, 1991):

1. Firstly, mathematical problem solving depends on number-identification or number reading skills (NR). Numbers can be translated from one type of presentation (e.g., the Arabic presentation '9') to another type of representation (e.g., the verbal oral representation of the number word 'nine'). Children need to know that 'nine' is not written as '6' and that '47' is not read as 'seventy four'.
2. A second building block is related to the mathematical lexicon. To solve mathematical problems, children have to deal with operation symbols (S) (e.g., x, +, <, >) without making mistakes.
3. Furthermore, mathematical problem solving requires knowledge (K) of the number position system. This refers to the ability to establish base-ten structure relationships. K skills are required to understand that 47 is composed by 4 tens and 7 units and that 47 is 1 unit larger than 46.
4. In addition, mathematics depends on procedural (P) knowledge and skills to calculate and to solve mathematical tasks (e.g., 47-9=_). Children have to know how to make subtractions to solve 47-9 as 38 and not as 42. Those P-skills depend on the mastery of the position system to carry out multi-digit operations. Though, to succeed, a child also needs to have access to arithmetical facts and stored solutions for calculations.
5. Linguistic skills (L) are conceptual skills, enabling children to understand and to solve one-sentence mathematical word-problems (e.g., 9 less than 47 is_).

Some children have no problems with the formula format (47-9=_), but seem to have problems translating words (e.g., 'less') into calculation procedures (e.g., a ‘subtraction’).

6. A mental representation (M) is required to solve most word problems, since a simple 'translation' of keywords (e.g., ‘less’) to a calculation procedure (e.g., ‘addition’), without representations, leads to ‘blind calculation’ or ‘number crunching’. This superficial approach leads to errors; e.g., answering '38 when presented with the tasks '47 is 9 less than _', 29 is 9 more than _' and ‘76 is half of _’.
7. Contextual skills (C) are cognitive skills enabling the problem solving of complex word-problems. Some children can have problems with this type of task due to the limited capacity of their working memory (‘cognitive load’) and/or to an insufficient knowledge base (‘expertise’).
8. In addition, some children fall behind when selecting relevant information (R) in order to create an adequate mental representation of a problem. These children have difficulties ignoring irrelevant numbers or information in a task. They believe all numbers have to be ‘used’ in order to solve a mathematical problem. They answer ‘59’ (47+3+9) to the problem ‘Willy has 47 cards. Wanda has 3 books and owns 9 more cards as compared to Ann. How many books does Wanda have?’.
9. Number sense skills (N) are the ninth building block. N enables solving tasks without calculating the correct answer. Some children lack such estimation skills to orient their solution of formula-tasks; e.g., 250-49=_ is about 200.
10. G (Geheugen in Dutch and memorising in English) skills are the tenth building block referring to memory tasks or automated skills. These are stored in long term memory and applied in an often less conscious way. For example, when executing tasks related to multiplication tables.
11. Visualization skills (VS) are the building block for solving spatial problems. VS skills enable students to produce and use geometric or graphic representations of mathematical concepts, principles, or problems. (Fennema & Sherman, 1977; Zimmermann & Cunningham, 1991).
12. The last building block refers to ‘logical thinking’ (LT). Logical thinking skills refer to the understanding of logical relations between quantities in order to learn how to represent numbers and arithmetic (Piaget, 1952; Nunes et al, 2007) .

1.2.2 Item Response Theory as an adequate method for development of diagnostic instrument

The validity and reliability of measurement instruments are central to the diagnosis of mathematical performance. In the Chinese context, Classical Test Theory (CTT) used to be the most popular measurement approach to test the psychometric quality of diagnostic instruments. CTT assumes that the true values (T) can be measured by the observed variables (X) considering some measurement error (e). The basic equation underlying this assumption is: X = T + e. CTT is central to various techniques for reliability assessment (such as calculating Cronbach's alpha) and to factor analysis approaches to get the validity of the scale structures (Churchill, 1979; Gerbing & Anderson, 1988; Singh, 2004). It has the additional advantage to provide a difficulty and discrimination index which can be easily understood by teachers and practitioners in educational contexts.

However, this approach is limited by a number of shortcomings. Firstly, test respondent parameters (test score) depend on the sample of items used. Mathematics tests for different grades in primary school would typically involve different samples of items (taken from the "universe of all possible math items") and are therefore not comparable with each other when a CTT based approach has been adopted. Second, item parameters (e.g. difficulty level) depend on the sample of individuals that took the test. As a consequence, when different items are administered in different grades, with CTT there is no way to establish how these different items compare to each other with respect to difficulty level. As a consequence, with CTT it is impossible to develop an instrument that covers the whole range of math abilities from grade 1 to grade 6. Such a unique scale however is critical if one aims to estimate the growth in mathematics ability in the course of a school career.

As an alternative, item response theory (IRT) starts with the proposition that estimation of student's performance (also called trait or ability) depends on both the student's response to a test item, and the properties of this item (Andrich, 1988; Embretson & Reise, 2000). When applying IRT, logistic regression models are used to estimate the probability (P) for a student with a given ability level (θ) to give a correct answer to any item g. Formula 1 shows that this probability is a function of how much the student's ability level θ exceeds the "difficulty level" (δ) of item g under the condition that θ and δ are measured along the same scale. With $\theta = \delta$, this probability would be .5. Different IRT models can be chosen to test the fit between the theoretical model explaining item quality and the actual solving of the item by individuals. In three parameters logistic (3PL) IRT model, the probability P depends on the item difficult δ_g, discrimination value α_g on the item and the guess parameter γ_g. In

2PL model, the γ_g is set to zero. In 1PL model, the γ_g is set to zero and α_g is set to 1. In the current study, since the items build on an open question format, there is no need to set the guessing parameters γ_g. Thus, we can start from a 2PL model and aim at attaining a parsimonious model.

$$3PL: P\left(X_g = 1/\theta\right) = \gamma_g + \left(1-\gamma_g\right)\frac{\exp\left[\alpha_g\left(\theta-\delta_g\right)\right]}{1+\exp\left[\alpha_g\left(\theta-\delta_g\right)\right]}$$

$$2PL: P\left(X_g = 1/\theta\right) = \frac{\exp\left[\alpha_g\left(\theta-\delta_g\right)\right]}{1+\exp\left[\alpha_g\left(\theta-\delta_g\right)\right]}$$

$$1PL: P\left(X_g = 1/\theta\right) = \frac{\exp\left[\theta-\delta_g\right]}{1+\exp\left[\theta-\delta_g\right]} \qquad (1)$$

The IRT procedure helps to estimate –at the same time - item difficulty and the participants' ability (trait), and helps to position a student along an underlying latent trait continuum. To summarize, compared to CTT, IRT presents the following advantages: first, item parameters can be estimated independent of the specific sample being involved. Second, ability parameters of students can be estimated independent of the specific subset of items (from a large pool of calibrated items). All the participants (e.g., from Grade 1 to Grade 6) can be positioned along the same trait continuum, even if they belong to different samples or take a different test form. As a result, IRT helps to determine both the characteristics of the ability (trait) distribution in the sample and the characteristics of the items of the test. This is why this approach is called "item response" since the focus is not on the item itself but on the way a participants "respond" to the item. Moreover, IRT is a method supporting an instrument's sustained development by integrating earlier developed and newly developed items. These IRT characteristics make it possible to keep an item bank open to new requirements, new content domain developments and new curriculum ideas. However, IRT has also disadvantages: first, it is not easy to understand and explain. second, IRT models have a high risk of not fitting the data; thirdly, some IRT models require large samples to develop accurate and stable parameter estimates although Rasch measurement model can be used with small to moderate size samples.

1.3 Research questions

Obtaining an instrument for diagnostic testing and for determining mathematics performance levels can have a significant impact on both primary education quality control, and the improvement of curriculum, and instructional approaches. As to the latter, the item bank gives reliable feedback about students' development and can help to detect student at risk and coping with specific mathematics difficulties.

In the present article, we report the results of a study, building on IRT, during the development of a comprehensive mathematics assessment instrument for Chinese primary schools that fits the 2001 curriculum. Item response theory (IRT) measurement models have been applied to test the quality of the items and to develop the item pool.

The objective of this paper is therefore to answer the following questions:

(1) What is the reliability and validity of this item pool from the perspective of item response theory?

(2) How can different users make use of this dataset or item pool ?

2. Methods

In the present study, an item pool was developed covering the whole range of primary mathematics topics, consisting of sets of items to be used for the six different grades. The six grade level tests include unique items for the specific grade, but also overlapping items that function as anchors with the former or later school grade test. The construction of this large item bank followed a three-phase procedure: (first step) item construction and selection, including delineation of the mathematics content and mathematics abilities by four experts and eighteen primary school. (Second step) item calibration during a pilot study, focusing on the six grade level sets of items, calibration of the items and selection of items in view of developing six test forms. (Third step) item calibration across the grades for the main research, including a data collection phase, testing the fit of the IRT model and final selection of the item pool.

3.1 Phase One: item development and preliminary item selection

A variety of item formats was adopted to develop items in relation to the twelve building blocks and the three mathematical literacy domains: completion questions, error correction questions, matching question, simple-choice questions, multiple-choice questions, open-ended questions and closed questions. A two step procedure was followed to structure this item bank. During a first step, four mathematics experts from the Educational Bureau of Beijing city and Guangzhou city, and eighteen Chinese primary school teachers participated in an interview. The four experts were recommended by the Head of Institution of Educational Research in Beijing and Guangzhou city. These experts were asked to select the eighteen teachers. The latter have at least 5 years of experience. Also, the teachers represent low, medium, and high performing schools. The experts and teachers were asked to construct the items and then study, select, and to evaluate the items considering the six different primary school grade levels. The experts and teachers were also asked to code the items: (1) “How do you think this item fits the new curriculum?” (1= fit, 0=

weak or no fit); (2) “How difficult do you think this items is for your students? ”(1=difficult, 0=easy).

In a second step, the items were reviewed by mathematics experts in Belgium and China. Items were classified according to their difficulty and the Chinese curriculum sub-domains and the twelve mathematics building blocks. The selection process aimed at developing a balanced set of items for each grade level. This resulted in a structured item bank of 2,521 items to be used for the six primary school grades (See Table 2).

Table 2
Distribution of items in the pilot study

Grade	N	Mathematics sub-domain		
		Number and algebra	Shape and space	Statistics and probability
1	498	362	95	41
2	539	437	68	34
3	291	174	78	39
4	486	266	206	14
5	488	321	151	16
6	219	157	43	19
Total	2521	1717	641	163

3.2 Phase Two: Calibration within grades during a Pilot study

In May 2008, a pilot study was set up to study the items based on IRT analysis. In total, 3,002 children and 42 teachers from grade one to grade six participated in this study. Sampling considered the variables rural/urban, and GDP of the province. Student sample characteristics are as a result: 59.89% from urban and 40.11% from rural areas. Considering the provincial gross domestic product (GDP) level (1=Highest level), 13.6% of students were from level 1 GDP provinces, 39.8% from level 2 and 60.2% from level 4. The procedure required whole classroom test administration during 60 to 120 minutes a day. Students worked individually to solve the test items in their test booklet.

All data were entered by the researchers and data files were screened for missing data and errors. The properties of the mathematics items were studied by using a IRT model analysis. The software tool (BiLog-MG3) was adopted for this analysis. This procedure helped to define a final item pool, containing items for each grade level. Exploratory factor analysis, building on all items at six grade levels, points out that a single factor solution explains about 60% of the variance in mathematics performance. Excellent internal consistency ($\alpha = 0.90$ to 0.95) and acceptable item-total correlations were achieved. No significant ceiling effects, floor effects, or gaps were detected. To attain good model fits, it was necessary to remove 145 items. After calibration, 2,376 items could be retained in the item bank. This new item pool was the starting point for the main study, set up in November 2008.

3.3 Phase Three: Main study with calibration across grades

3.3.1 Research procedure and sample characteristics

In November 2008, a main study was set up to calibrate the six grade level tests, to examine the item characteristics, and the implementation of the mathematics learning performance diagnostic scale. A test design was developed based on six grade level mathematics tests, with overlapping items. In order to limit the test administration time, test forms were constructed with about 80 items for each grade. Each grade level test administration was estimated to require 120 minutes. In table 3, we report the number of test items in each grade level booklet. In addition, the table shows how each grade level test covers the relevant mathematics domain and buildings blocks. The numbers with Bold style show the number of items overlapping with item sets for the former or next grade level. These items are expected to define a continuous item difficulty scale, across the 6 grade levels.

Table 3
Structure of the six grade level tests

	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Sum
set 1	67						67
set 2	9	9					9
set 3		42					42
set 4		7	7				7
set 5			63				63
set 6			4	4			4
set 7				46			46
set 8				7	7		7
set 9					51		51
set 10					8	8	8
set 11						90	90
set 12	1		1				1
set 13	7	7	7				7
set 14	2	2	2	2	2	2	2
set 15		4	4	4	4	4	4
set 16			4	4	4	4	4
Sum	86	71	92	67	76	108	412

The mathematics tests were presented to a sample of 10,959 students from 20 Chinese primary schools (See Table 4). Sampling approaches were comparable to the pilot test. Firstly, thirty one provinces and cities (excluding the Special Administrative Regions) were ranged along six levels according to their Gross Development Product (GDP) (Level 1 = highest level), building on the 2005 classification of the Chinese Economic Bureau. Secondly, one province or city was selected from the first to the fifth GDP level province. Schools from sixth GDP level provinces were not included in the study due to difficulties getting access to these areas and schools. Thirdly, within each of the five provinces or cities, four schools were selected randomly: two urban and two rural area schools.

Table 4
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%
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%

3.3.2 Data analysis

All raw student answers were entered into an Excel file to be able to retain the original item responses. Next, the responses were corrected and recoded into new variables.

The first step in the analysis focused on a basis assumptions: the uni-dimensionality of the items. This analysis was based on the results of an exploratory factor analysis. If this resulted in a single dominant factor, explaining close to 20% of the variance, uni-dimensionality could be confirmed (Reeve & Masse, 2004).

In a next step, the BiLog-MG3 program was used to fit a 1PL and 2PL-IRT model. The BiLog-MG3 tool makes use of the maximized marginal log-likelihood (MML) estimation process. It builds on first- and second-order derivatives to estimate item parameters. By using the Bayes theorem, the MML estimation process within the expectation-maximization (EM) algorithm, comprises of three steps in BiLog-MG3, which are repeated until convergence of the item parameter estimates is achieved (Rupp, 2003). After achieving convergence, the model is considered to have a good fit if (a) the fit statistics for each individual items have a p-value > 0.1 and (b) satisfactory overall fit-statistics (with $p>0.1$) are attained. In the present case, starting with 412 items and without elimination of specific items, a number of 41 (=10%) items with $p<.10$ could be accepted with maximum 20 of them having $p<.05$, including a maximum of 4 items with $p < .01$. Achieving a good model fit necessitated a process of subsequently eliminating bad fitting items, while trying to keep as much anchor items as possible, and trying to keep a balance between the different mathematics sub-domains covered in the test (in order to maintain content validity).

When a larger number of items is being calibrated than needed, further selection of items (or construction of parallel tests) is carried out based on the item characteristic curves (ICC), the derived item information curves (IIC) and the underlying item parameters. For each test form (grade), a shorter test length can be obtained by selecting the items with higher discrimination parameters. Inspection of the empirical reliability estimates for each of the test forms, gave a first idea of the reliability of each test form (comparable with Cronbach's alpha). Further examination of item parameters and test characteristics was performed to corroborate the appropriateness of the items and test forms to the student population.

In a third step, the item-person continuum map is visually inspected. This mapping puts both item location ("difficulty") and respondents' estimated ability on the same logit scale. This provides critical information about the ability range covered by the tests (and the different test forms). This is useful to determine to what extent the items are appropriate for the target population by comparing the range of respondents and the items. This mapping was performed for the test as whole and for each grade (test form) separately.

In a last step, item characteristics curves (ICC) and item information curves (IIC) were examined. ICC show the probability of selecting the response at each logit. Due to space limitations, these plots are not incorporated in the present article, but can be obtained from the corresponding author. Test information curves (TIC) show the total test information for each test form and the associated standard errors at any point of the ability scale. Based on those curves, we can observe which test form provides us with the most information (=highest reliability) for any given ability range.

4. Results

4.1 Preliminary analysis: Exploratory Factor Analysis

Uni-dimensionality is an important criterion for IRT models (Hambleton, Swaminathan, & Rogers, 1991; Linden & Hambleton, 1997). In this study, the mathematics scales covered three main domains from the new Chinese mathematics curriculum and mirrored twelve mathematics abilities (Desoete & Roeyers, 2005). Consecutive exploratory factor analyses with Principal Axis Factor analysis (PAF) based on a Direct Oblimin rotation were carried out to examine the dimensionality of the items in each grade level test form. The results show that the first factor explains close to 20% of the observed variance; and this proportion of the explained variance is about three times the size of variance explained by the next factor (see table 6). Also the difference in the eigenvalues of the first two factors underpins the uni-dimensionality of the scale. Analysis of the screen plots presents a sharp drop from the first factor to the second factor, with a leveling off in the percentage of the variance explained after the second factor. This reconfirms opting for a one-factor solution as the best alternative (Scree plots can be obtained from the corresponding author).

Table 5
Percentage of explained variance of each factor at different grade levels

Subscale	% explained variance factor 1	% explained variance factor 2	ratio first vs. second factor
Grade 1	17.17	6.10	2.81
Grade 2	25.25	7.86	3.21
Grade 3	17.29	7.40	2.33
Grade 4	18.77	7.18	2.61
Grade 5	19.17	5.88	3.26
Grade 6	17.73	4.90	3.62

4.2 Model selection and Model fit

1PL and 2PL models were run by the BiLOG-MG3 to check for the most optimal model, starting from a 2PL model. To obtain a good fit, in total 86 items were removed from the item pool of 412 items. Eight items were detected by the

BiLOG-MG3 to have a biserial correlation value <-.15, and were therefore automatically removed. In addition, seventy eight items had to be deleted because of a bad fit. After deletion, a satisfactory fit index was obtained. The final model fit index (G^2) is 1378.22, SE = 236.0. The p-value for the chi-square is close to 1.000, reflecting a good fit between the data and the model.

In a next phase, the analysis focused on finding a parsimonious 1PL model. The relative fit of consecutive models is addressed by the analyses reported in Table 6. The results of the model comparison reveal that the -2 Log Likelihood indicators of 2PL model is the best model. The less complex 1PL model reflects a worse fit as compared to the 2PL model (Δx^2=2542.5, $\Delta df=326$, $p=.00$). In addition, the number of the bad fitting items increases from 10 items to 99 items in the former case, implying that even more items should be deleted to attain an acceptable fit of the 1PL model. This would result in a loss of information. To conclude, the analyses indicate that the 2PL model is the most optimal model.

Table 6
Results of model-fitting for 1PL, 2PL and 3PL models

	Number of bad fit items	x^2	***df***	***p***	**Deviance**	Δdf	***p***
1PL	99	3920.72	2^{326}-326*1-1	.00			
2PL	10	1378.22	2^{326}-326*2-1	1.00	2542.5	326	.00

4.3 Reliability and validity of the item bank

4.3.1 Empirical reliability of the test forms

For each of the six grade level test forms, empirical reliability (comparable to Cronbach's α) ranges from .94 to .96 (see table 7). According to Nunnally and Bernstein (1994), internal consistency should be at least .60 for a self-report instrument and at least .80 when used as a screening instrument.

The average proportion correct of the 326 items in the final test sets is .61 (*SD* = .20). The average mean value of Pearson item and test correlation is .39 (*SD* = .14). The average mean bi-serial correlation is .54 (*SD* = .19). The Pearson correlation value estimates the relationship between the dichotomously scored item *j* and the total score *x,* while the biserial correlation value estimates the relationship between the total score and the hypothetical score on the continuous scale underlying the (dichotomous) item.

Table 7
Test form characteristics

Test form (Grade)	Empirical reliability	Slope			Location		
		Min	Max	Mean	Min	Max	Mean
Grade 1	.94	.15	1.55	.72	-5.30	1.86	-1.24
Grade 2	.96	.33	1.24	.65	-4.83	1.65	-.89
Grade 3	.95	.13	1.18	.63	-2.26	9.27	.05
Grade 4	.94	.32	2.87	.74	-2.19	2.77	.18
Grade 5	.94	.33	1.19	.67	-2.05	2.82	.69
Grade 6	.93	.33	1.90	.88	-1.97	3.30	.83

4.3.2 Distribution of item parameters and continuum scaling

Calibration of the 326 items in the final item bank shows an intercept range from -3.84 to 2.83 (*M* = -.02, *SE* =1.09). The location of the items, showing the difficulties parameter, ranges from -5.30 to 9.26 (*M* = -.08, *SE* = .18). The slope of the items, showing the discrimination parameter, ranges from .13 to 2.87 (*M* = .75, *SE* = .34).

Item location parameters range from -5.30 to 9.26 (*M* = -.08, *SE* = .18).The student parameter (ability) ranges from -4.89 to 4.14 (*M* = .57, *SE* = .26). This implies that mathematics abilities are well covered by the items (-5.30 < -4.89 < 4.14 < 9.26). The distribution of ability scores is graphically represented at the left side in Figure 1. The graphs of the students (ability) and items distributions are clearly mirrored. One can observe that there is a large overlap in the area covered by both distributions. We observe no floor or ceiling effects. At the higher end of the item location distribution, we perceive three outliers, i.e., items with a very high difficulty level.

As can be derived from table 7, the mean of the item locations shows becomes larger, indicating an increase in difficulty from grade 1 to grade 6, as could be expected. A more or less similar increase is seen for the item with the lowest location within each test form. With respect to the maximum location a very high maximum location of 9.26 is observed in the grade 3 test form. This is due to an extremely difficult item. As can be observed in figure 2, there are three items in an extreme position. The fourth highest location in the grade 3 test form is 3.30; but again within the expected range.

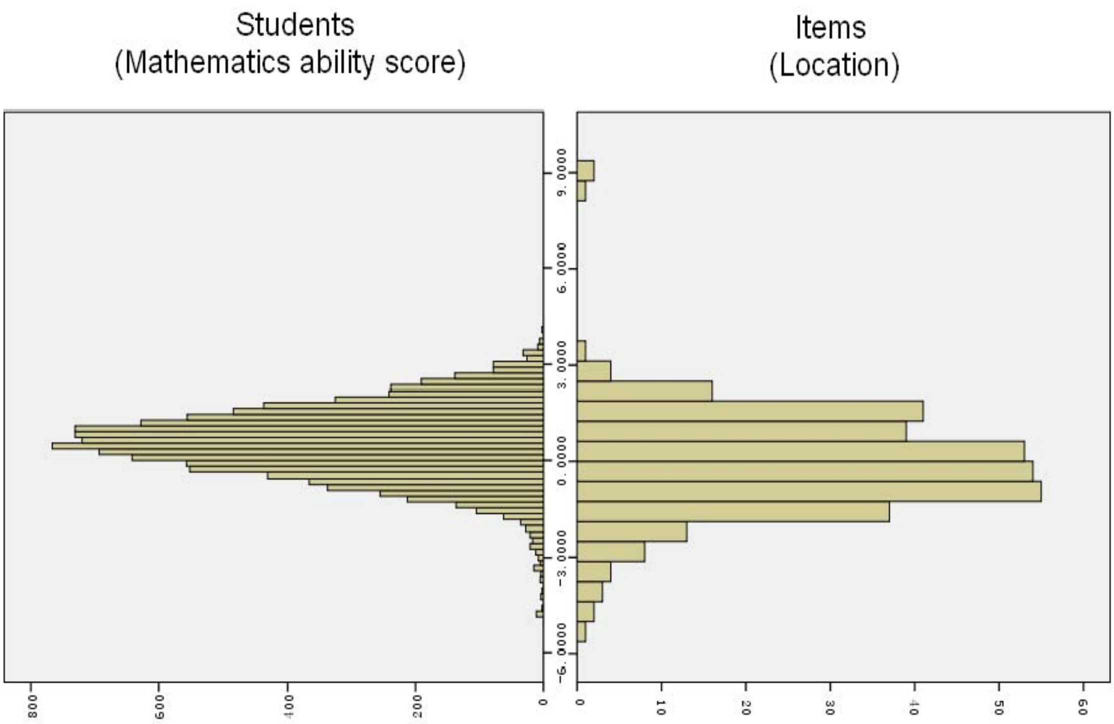


Figure 1
Distribution of students' ability estimations and item locations.

The minimal and the maximum for the item slope, show considerable differences between the items with respect to their discrimination value. Figure 2 illustrates this by representing the ICC of two items (item 168 and item 356) with very different slopes but a similar location. Detailed information about the item characteristics curves (ICC) of the deleted items can be obtained from the corresponding author.

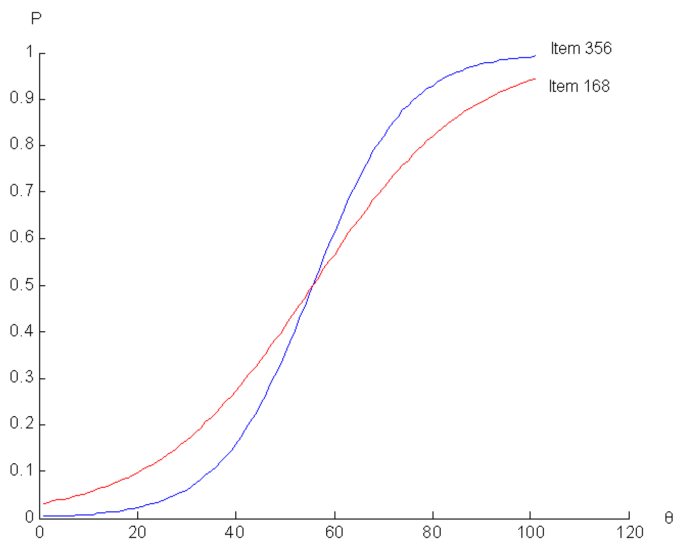


Figure 2
Item Characteristics Curves.

4.3.3 Information curve: a new index of the "reliability" in IRT

In an IRT measurement model, the item information and test information are interpreted as an index of the reliability of the test. The concept of "item information" tells us how much information about the person's ability is revealed by a particular item. The closer the item location matches a person's ability level, the more that item can tell about a person's ability. But also the discriminative value of the item plays a role, as is shown by the item information function given by Hambleton and Swaminathan (1985):

$$I_j(\theta) = \frac{a_j^2 e^{aj(\theta - bj)}}{\left[1 + e^{aj(\theta - bj)}\right]^2} \qquad (2)$$

The standard error (SE) of the ability score θ is inversely related to the amount of the information(I) provided by a set of test items (formula 3).

$$SE(\theta) = \frac{1}{\sqrt{\mathrm{I}(\theta)}} \qquad (3)$$

In formula (2), the maximum value is directly proportional to the square of the item discrimination parameter, *a*. A larger value of *a* is associated with a larger amount of information. The maximum information level is obtained at point b_j on the ability scale, i.e. for person with an ability $\theta = b_j$. Figure 3 depicts – as an example - the item information curves (ICC) of five different items, taken from five different grades. Item 66 (grade 1) and item 129 (grade 2) are clearly easier items and give more information about the lower range of abilities. For those ability ranges, their SE will be lower than for higher ability ranges. In contrast, item 291 (grade 5) and item 366 (grade 6) are more difficult items, giving information (and thus showing lower SE) about a higher ability range. For the lower ability range, the SE will be higher. Both item 252 (grade 4) and item 202 (grade 3) represent items of average difficulty, when the whole item pool from grade 1 up to grade 6 is considered. In cases where the slope of the ICC is steeper, the item is more sensitive to ability differences and provides us with more information about the person's ability (implying that the SE will be lower).

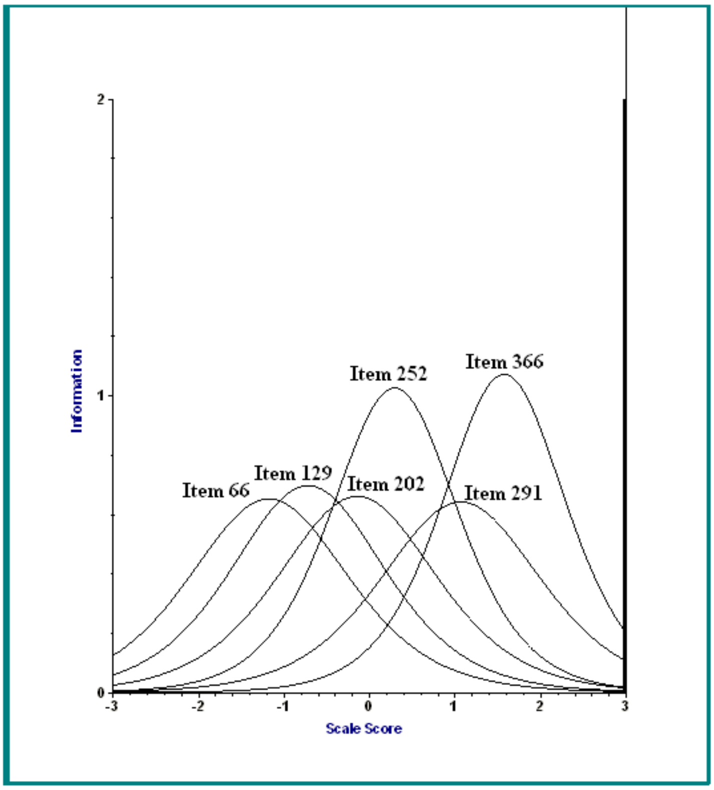
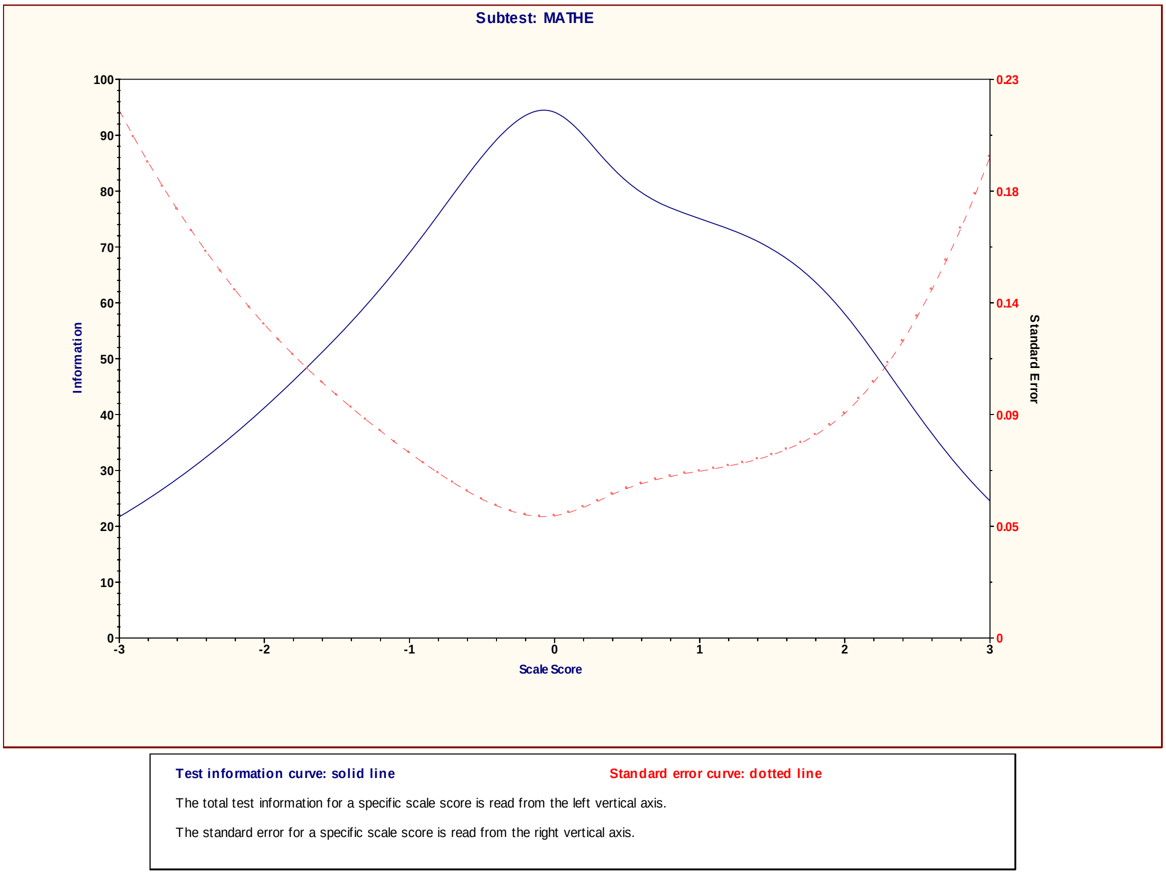


Figure 3
Item information curves of 6 typical items taken from 6 different grade level scales.

The "test information" curve for the 326 items, included in the cross-grade calibration, is shown in figure 4. It depicts the amount of information provided by the total set of items at each point on the ability scale. The efficiency of a particular proficiency estimate depends upon its standard error. The highest test information function value is 94 (corresponding *SE* = .05). Smaller standard errors of measurement imply better interpretable results as compared to items with larger standard errors. As can be derived from the figure, the complete test provides us with information for the range of the mathematics abilities between theta -1.65 and 2.22. Beyond this range, the information curve drops off and the standard error increases. Only about 9.1 % of the persons (998 students), are located in the area where test information is lower than the standard error. This group consists of 748 students (6.8%) who are located at the higher abilities end, and 250 students (2.3%) that are located at the lower abilities end of the curve.



Figuare 4
Test information and measurement error curve.

Test information curves synthesize how much information is provided by a set of items at any point of the ability scale. The test information curves for each test form (one per grade) are shown in figure 5. They look very similar to item information curves and can be interpreted in a similar way. As shown in figure 5, the six test forms differ in their maximum information value (the height of their peak), which means they differ in the accuracy of their scores. For the students falling within the ability range covered by the specific test form, the test forms for grade 1, grade 4 and grade 6 provide more information as compared to the other three grade test forms. But The ability range that is well-covered by test form 4 is much smaller as compared to the other test forms.

The test forms for grade 2, grade 3 and grade 5 cover a wide range in abilities, but their peaks are lower, meaning that even for students falling right in the middle of the ability range covered, the information is lower than is the case for the other test forms. There is some overlap between the ability ranges covered by the different test forms.

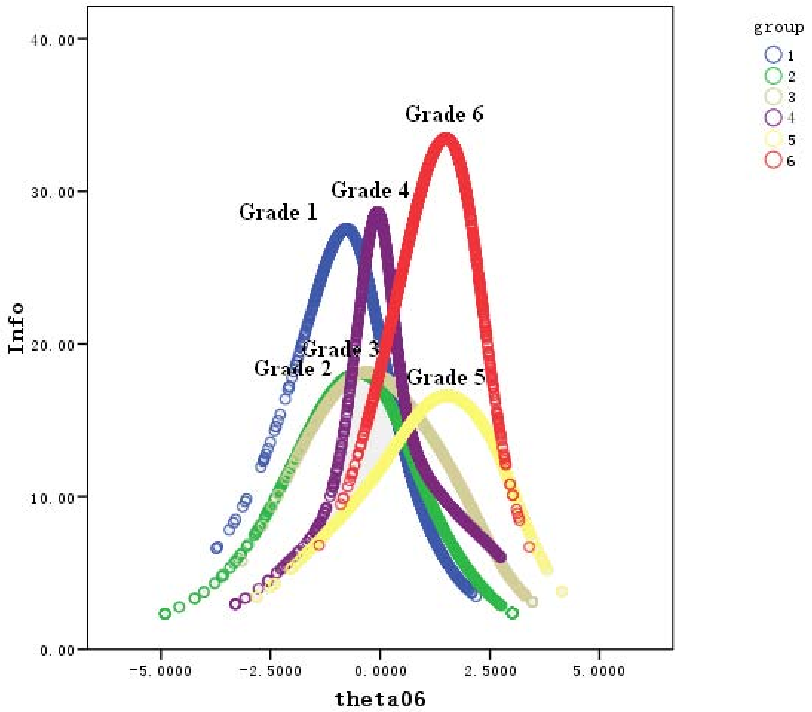


Figure 5
Information Curve for Each Test Form.

The empirical reliabilities reported in table 7, suggest that for each test form the test length can still be reduced without falling below the generally accepted lower boundary of .80. In order to get as much information as possible, when constructing new test forms from the pool of calibrated items, items could be selected in such a way that the test forms differ less in their peaks and more in the ability range they cover.

4.3.4 Content validity of the item bank

The development of the item bank described in the methods section also provided information about the content validity. After the calibration by IRT, the remain of the 326 items cover the the three mathematics curriculum fields and the twelve mathematics building blocks (See Table 8).

We can also compare the difficulty estimation based on experts' expectations with available IRT based empirical evidence to check internal content validity (Wilson, 2005). The related correlation coefficient is .82. This means that from the perspective of mathematics educational practicioners, the theoretical results and the practical results are almost identical.

Table 8 Distribution of the math knowledge and the building blocks

	Number and algebra	Shape and space	statistics	Total
NR	234	56	6	296
S	167	13	3	183
K	144	43	10	197
T	234	62	14	310
P	187	29	5	221
V	56	24	2	82
C	138	56	15	209
R	16	18	5	39
N	11	0	0	11
G	40	1	0	41
Vis	22	69	12	103
Log	9	4	0	13
Total	239	72	15	326

Note: The total of the building blocks is not equal to the sum of the items because the building block is Concurrent.

4.3.5 Construct validity of the items

The construct validity of items can be checked in relation to each specific test item (Wilson, 2005) (See Table 9 in Appendix I). As we can derive from Table 9 and taking item 1 as an example, the group of students who achieve the specific item (scored “1”) will be expected to have the higher mean location compared to the group of student who fail (scored “0”). The information for each of the remaining items can be asked from the correspondent author.

4.4 Utilization potential of this first version of the item bank

4.4.1 Predicting students’ ability levels: Relation between CTT and IRT

The number-correct score is an important measure to determine a students’ ability level in the context of a Classical Test Theory approach (CTT). However, this can be misleading. Wright and Stone (1979) list examples demonstrating the flaws in this approach, e.g., a person who guesses the correct answer to a multiple-choice items and raises in this way his/her test score. An Item Response Theory approach (IRT) can detect such misfitting item-score patterns. According to IRT, the probability of correctly answering item g is a function of θ and item characteristics. But for primary school teachers, getting the CTT score is the easiest way. Actually, CTT and IRT results are correlated to each other. The value of an IRT score can be predicted on the base of the percentage of the number-correct by using the curve estimation regression. Since the relationship between the percentage of items correct and IRT-based ability score can differ according to test form, this relationship has to be estimated for each test form separately. Figure 6 shows that for each of the math test forms in this study the IRT-score can be predicted from the percentage of items correct with a high degree of accuracy.

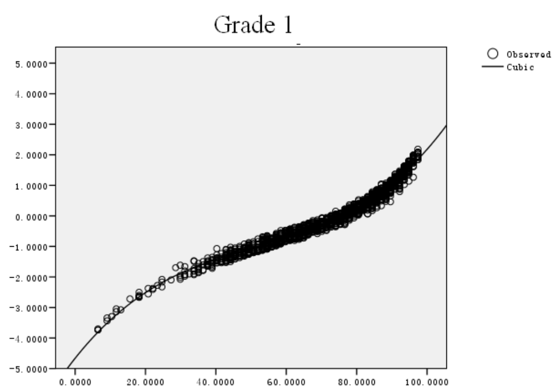


Figure 6A

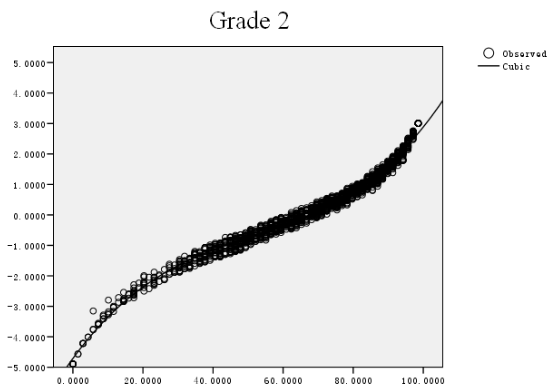


Figure 6B

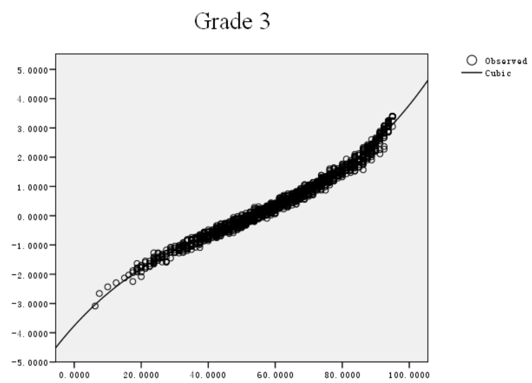


Figure 6C

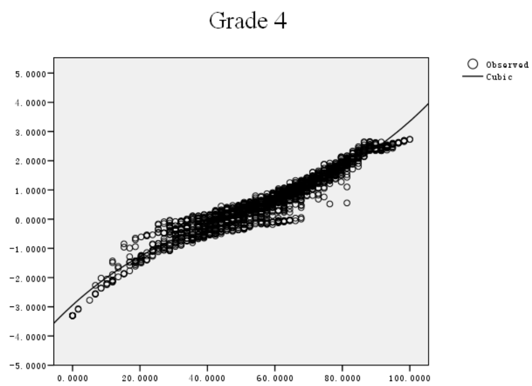


Figure 6D

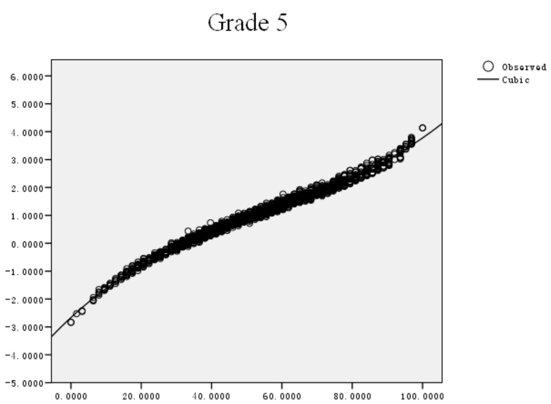


Figure 6E

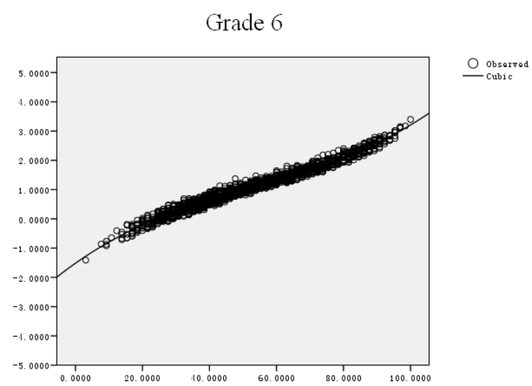


Figure 6F

Figure 6 Prediction of the IRT-based math ablity by percentage-of-items-correct score

The percentage of explained variance for each grade is 98.3%, 98.8%, 99%, 93.4%, 99.1%, 97.5% respectively. In addition, all coefficients are significant and can be included in the model. For example, the regression equation for grade 5 is:

$$\text{Mathematics ability} = 2.207*\text{number-correct score} - 2.623*\text{number-correct score}^2 + 1.448*\text{number-correct score}^3$$

These results are important because they show that reliable and quite accurate estimations of math ability scores can be made based on the number or percentage of items correct for a given test form, without having to perform an IRT-analysis on every new dataset. This allows the construction and publication of simple conversion tables from which teachers can easily read what math ability level corresponds to a number-correct or percentage-correct score for each of the six test forms. Note that these results also indicate that with respect to the 326 items that subsisted the IRT-analysis, we can claim that more competent students will be able to solve both easier and difficult items, whereas less able students will only succeed in solving easier items. Both the graph for the grade 4 test form and the corresponding figure 6D indicate that the grade 4 test form could still be improved.

4.4.2 General information about the impact of Chinese math education

Using the data of the more than 10,000 students that participated in the main study, a first tentative analysis can be made about the average math growth in Chinese students across primary school grades. Of course, one should take into account that taking unweighted means does not account for the fact that some strata were oversampled. Also, a correct estimation of “growth curves” based on cross-sectional data requires more sophisticated analysis techniques that are not possible within the context of the present study. However, on the base of the data, a first idea of growth differences between grades can be developed (See figure 7).

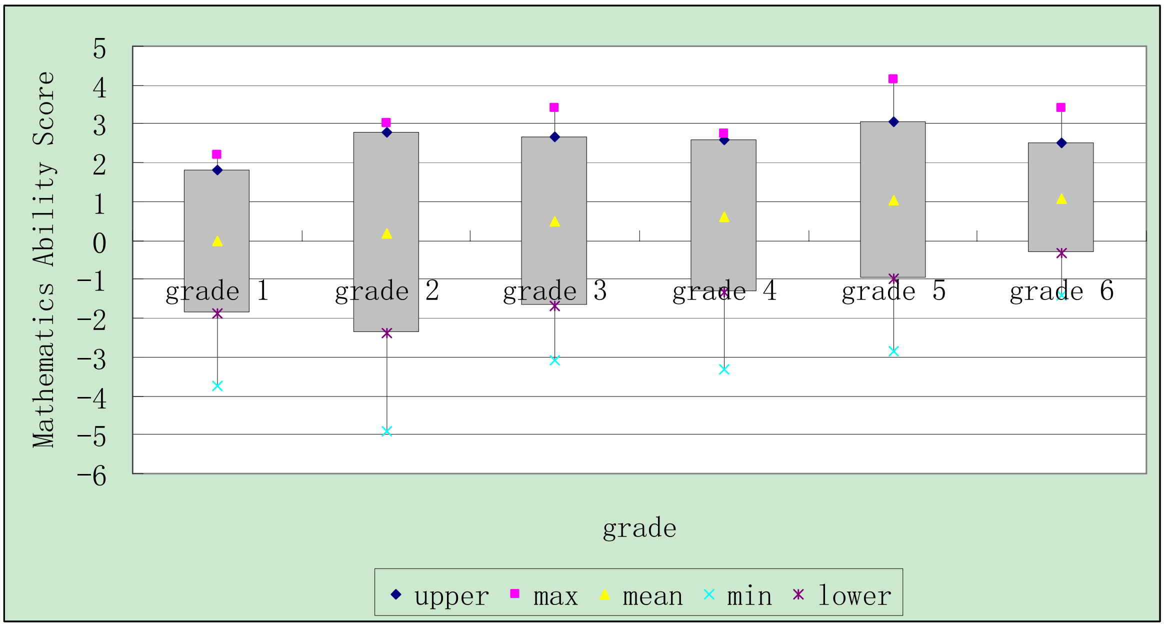


Figure 7
Boxplot of comparison of six grades.

As can be derived from figure 7, there is a smooth increase in the development of mathematics abilities according to primary school grades. The upper and lower edge of the box indicates the 95% of the data set. Each of the boxplots illustrates a different pattern. These results help to reconsider a number of observations about current mathematics performance in Chinese primary schools. On the base of the IRT results, Chinese mathematics performance in grade 4 and grade 6 seemed to be not very high, as compared to the mastery attained in grade 3 and grade 5. This is in line with the findings of other researchers. Liu (2007) reported that students in grade 4 and grade 6 hardly evolve in their geometry performance as compared to learners of grade 3 and grade 5. Basang (2006) reported similar results in the statistics sub-domain. This suggests that questions can be asked about the structure of the current curriculum. The fluctuations in mathematics performance can also reflect limitations in the way the curriculum is implemented via current Chinese mathematics textbooks. Grade 4 and grade 6 seem to be grades in which pervious textbook content is further exercised, resulting in less average growth. Another interpretation suggests a natural levelling off ability after grade 3 and after grade 5 that is independent of the curriculum content, and due to students needing more time to practice and assimilate the new material presented in grade 3 before they are able to process new mathematical content.

5. Discussion, limitations and conclusion

5.1 Discussion

5.1.1 An attempt to combine a domain-specific achievement test and a cognitive ability test

In the present study, test items were developed to give information about both mathematics achievement as well as about the underlying cognitive abilities. In the literature, research linked to international student assessment studies (like PISA, TIMSS) discusses whether domain-specific student achievement test can be distinguished from cognitive abilities tests (Baumert, Ludtke, Tautwein, & Brunner, 2009). This discussion centres on the different interests of educational psychologist and educational researchers. Previous studies tried to prove that domain-specific abilities tests present a particular challenge for the nested-factor model and the higher-order factor model (Burner, 2008). This implies that the scaling is to be based on both scales. In the present item bank, we tried to combine both perspectives by coding the data from both perspectives.

Nevertheless, the present item bank is to be considered as a first version. Future studies are expected to help to include more items and by involving other student samples to expand the coverage of a full range of domain-specific and cognitive abilities.

Also, considering the aims of a mathematics diagnostic test, a new test can be constructed by selecting items with particular characteristics from the present item bank. This test could help to link ranges in abilities and grade levels when controlling for background characteristics of learners in different Chinese provinces, gender and/or school performance levels.

5.1.2 Information about the quality of Chinese mathematics education

In 2001, a curriculum reform experiment was piloted in China. No evaluation was set up to ascertain whether the curriculum reform goals had been met (Beijing Report, 2006; Marton, 2006). The practical implication of the findings of the present study is that critical progress has been made with respect to the construction of an item pool that is in line with the new curriculum standards for Chinese primary schools and that covers a wide range of mathematics abilities. This new item bank can help to provide comprehensive information about mathematics, considering a wide variety of school performance levels in China. Such a comprehensive assessment approach is needed to ground remedial approaches that consider the strengths and weaknesses of individual children (Grégoire, 1997). This information can feed the development of remedial approaches of children with special needs (For example, See Appendix II).

5.2 Limitations and directions for future research

Despite the strengths of the item bank, described above, a number of limitations and directions for future research have to be put stressed.

Firstly, some limitations are due to the timing of the test administration. The pilot test was administered at the end of May, while the main test was administered at the end of November. Both test administrations were planned at the end of a semester and prior to an examination period. Some differences in mastery level can be influenced by the timing of the test administration cycles. Further validation of the IRT-model is needed building on new research data gathered at more appropriate times. The latter depends on the purpose of the test and is to be discussed with particular stakeholders: educational authorities, principals and teachers.

A second problem is related to between-school differences in the extent to which the new national curriculum reform launched since 2001 has been implemented. Between and within regions, schools adopted this curriculum innovation at a different speed. This can result – within the same school – in a mixture of old curriculum and new curriculum implementation levels. This also might have affected differences in mastery levels of particular mathematics content. Thus, during the IRT calibration process, some difficult - but curriculum relevant items - might have been unjustly rejected. This is in particular the case with items that reflect statistics content and content dealt with at the end of a textbook.

Third, the structure of the current item bank can be improved by focusing on a still broader conception of "mathematics ability". In the present study, a single-factor structure was pursued. Future research can focus on mathematics sub-domains, and/or related mathematics abilities. This implies that a multiple factor structure of mathematical abilities is to be adopted. In the future, the Attribute Hierarchy Method might be used to explore the relations between the twelve building blocks (Gierl, Zheng, & Cui, 2005).

Fourth, the validity of the item bank should be tested against external criteria. The present item bank can be contrasted with – when available - alternative Chinese mathematics performance tests. In addition, interviewing students who participated can help to verify the validity of test items which is not be done in the present study.

Lastly, future studies could focus on determinants of within-group and between-group differences. In this context, the sample stratification variables can direct the study of differences between schools in different provinces, urban/rural schools, high and low performing schools, as well as gender differences. Using data collected through means of background and teaching behaviour questionnaires, also the effectiveness of teaching strategies, the impact of social economic background of the students, etc. could be linked to current mathematics performance. This is critical in view of developing an in-depth understanding of the nature of mathematics performance in China and to direct policies aiming at higher mathematics achievement.

5.3 Conclusions

This paper presented an overview of available mathematics performance and diagnosis instruments in the Chinese context. Though some of these instruments try to link Chinese instruments to international developments, clear shortages can be observed in the availability of diagnostic instruments that fit the Chinese curriculum. In the current paper, a new mathematics item bank was developed to cover a wide range of scholastic knowledge and mathematics cognitive abilities. Moreover, the paper adopted the use of IRT methods to construct this new item bank to pursue sustainable characteristics. For example, new test items can easily be integrated into the item bank, and also for students who didn't participate in the current test, a score can be predicted. Also, data from new student samples can be linked to available data, to further study the evolution in the quality of Chinese mathematics education. Lastly, the present study resulted in a dataset with rich information about the impact of the curriculum reform and the current level in children's development in Chinese primary school. The findings from the research sample, comprising of over 10,000 students, from five provinces with different development levels, present a clear picture of the state-of-the-art in Chinese mathematics education. This information can be considered as a starting point for further studies.

References

Andrich, D. (1988). *Rasch Models for Measurement*. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-068. Beverly Hills, Sage Pubns.

Basang, Z. (2006). Zhongxiao xuesheng dui tongji de renzhi shuiping yanjiu. [*Research on middle and primary school student's cognition level to statistics*.] (Doctoral Dissertation, Northeast Normal University). Retrieved from http://www.cnki.net/

Baumert, J. Ludtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition:

Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review, 4*, 165-176.

Bayley, N. (1933). *The California First-Year Mental Scale*. Berkeley: University of California Press.

Bayley, N. (1969). *Bayley Scales of Infant Development. Manual*. New York: Psychological Corporation.

Bayley, N. (1993). *Bayley Scales of Infant Development (2nd ed.; Bayley-II*). San Antonio, TX: Psychological Corporation.

Beijing Report (2006). Beijingshi Jichu Jiaoyu Kecheng Gaige Shiyan Xueke Jiaoxue Jiankong yu Pingjia Baogao. [*Report on curriculum reform, teaching and evaluation on elementary education in Beijing City*..] *Jiaoyu Yanjiu* [*Educational Research*] *3,* 77-82. Burner, M. (2008). No *g* in education? *Learning and Individual Differences. 18*, 152-165.

Churchill Jr., G.A., 1979. A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research, 16* , 64–73.

Das J. P., Naglieri J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Boston: Allyn & Bacon.

Deng, C.P., Zuo, Z.H.,Li, Q.W., & Das, J.P (2007). Shuxue Xuexikunnan ertong de bianma jiagong tedian: jiyu PASS lilun de yanjiu [Coding and Processing Difficulties of Children with Mathematics Learning Disabilities: A Study Based on the PASS Theory]. Xinli Kexue [*Psycholgoical Science*]*, 30,* 830-833

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

Embretson, S. E., & Reise, S.P. (2000). *Item Response Theory for Psychologists*. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Fan, C. R. (1988). *CDCC Yingyouer zhineng fayu ceyan shouce*[*Manual of CDCC infant mental development scale*]. Beijing: Tuanjie chubanshe.

Fan, C. R. (1989). Yingyouer CDCC zhineng fayu liangbiao de bianzhi [The establishment of CDCC infant mental development scale]. *Xinli Xuebao[Acta Psycholgoica Sinica]. 21,*130-140.

Fennema, E., & Sherman, J. (1977). Sex related differences in mathematics achievement, spatial visualization, and affective factors. *American Education Research Journal, 14,* 51–71.

Frey, M.C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*, 373-378.

Gerbing, D.W., & Anderson, J.C. (1988), "An updated paradigm for scale development incorporating unidimensionality and its assessment", *Journal of Marketing Research*, *25*, 186-92.

Gierl, M. J., Zheng, Y.G., & Cui, Y. (2005). Using the attribute hierarchy method to identify and interpret cognitive skills that produce group differences. *Journal of Educational Measurement, 45*, 65-89.

Grégoire, J. (1997). Diagnostic assessment of learning disabilities. From assessment of performance to assessment of competence. *European Journal of Psychological Assessment*, 13, 10-20.

Grégoire, J., & Desoete, A. (2009). Mathematical Disabilities–An Underestimated Topic? *Journal of Psychoeducational Assessment, 27,* 171-174.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer-Nijhoff.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Jing, J., Yu, M., & Deng, G. F. (1995). Xuexinengli Zhangai shaicha liangbiao de xiuding he zai xiaoxuesheng zhong de shiyong [The Revise on learning disabilities screening scale and its use in primary school]. Xinlifazhan yu jiaoyu [*Psychology Development and Education*]*, 2,* 24-29.

Lin, C. D., & Zhang, H, C. (1988). Weishi Ertong zhili liangbiao Celiangshu [*Manual for Wechsler Intelligence Scale for Children*]. Beijing: Beijing Normal University Publisher

Linden, W., & Hambleton R. K.(1997). *Handbook of modern item response theory*. Springer – Verlag New York Inc.

Liu, C.W.(2004). Xiaoxue Liunianji Xuexi Kunnan Xueye Chengjiu Shaicha ceyan de chubu bianzhi [Primary construction of learning difficulties academic achievement test for grade 6th of primary school.]. (Doctoral Dissertation, Hunan Normal University). Retrieved from http://www.cnki.net/

Liu, X.M. (2007). Xiaoxuesheng Kongjian guannian de fazhan guilv ji tedian yanjiu[Study on the characteristics and laws of the development of the pupil's spatial sense]. (Doctoral Dissertation, Northeast Normal University). Retrieved from http://www.cnki.net/.

Lv, J.(1991). Ertong renzhi nengli zhenduan liangbiao de tedian ji fenxi[The character of children cognitive ability scale and the analysis on it]. Xinli Xuebao[*Psychology Journal*]*, 3*, 297-303.

Marton, A.M., (2006). The cultural politics of curricular reform in China: a case study of geographical education in Shanghai. *Journal of Contemporary China 15* , 244–245.

Myklebust, H. 1971). *The Pupil Rating Scale: Screening for Leaming Disabilities*. New York: Grune 6- Stratton, Inc.

Munzert, A. W. (1991). *Test Your IQ*. New York: Simon & Schuster.

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*, 147-166.

Nunnally, J.C., & Bernstein, I. H. (1994). *Psychometric Theory (3rd*). New York: McGraw-Hill.

Piaget, J. (1952). *The child's conception of number*. London: Routledge & Kegan Paul.

Qin, L. H & Wu, Y. Y. (1996). Guoxiao dinianji shuxue zhenduan ceyan[*Mathematics Diagnosis Test on Primary School Primary-grade Students*]. Taiwan: Psychology Publisher

Raven, J. C. (1938). *Progressive Matrices: A perceptual test of intelligence Individual Form.* London: H. K. Lewis & Co.

Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales, Part 3, Section 7, Research.* London: H. K. Lewis & Co.

Raven, J. C., Court, J. H., & Raven, J. (1988). *Manual for Raven's Progressive Matrices and Vocabulary Scales: Section 3. Standard Progressive Matrices*. London: H. K. Lewis & Co.

Reeve, B.B., & Masse, L.C. (2004). *Item response theory modeling for questionnaire evaluation.*(pp. 247-273) In Presser S., Rothgeb, J.M., Couper M. P., Lessler, J.T., Martin E., Martin, J.,Singer, E. (eds). Methods for testing and evaluating survey questionnaires. Hoboken, NJ: John Wiley & Sons.

Shao, Z.F., Chen, G. P., &Shan, Y. (2000). Xuexi Kunnan ertong zhenduan liangbiao changmo zhiding[Establishment of Norms for Diagnosis scale on Learning Difficulties]. *Xinli Kexue[Psychological Science]. 23,* 169-252.

Singh, J. (2004). Tackling measurement problems with Item Response Theory: Principles, characteristics, and assessment, with an illustrative example. *Journal of Business Research*, *57,*184-208.

Stephen, B., & McCarney, P. (1996). *Learning disabilities evaluation scale*. Hawthorne: Education Servieces, Inc.

Wang, D., & Qian, M., (1987). Lianhe xing ruiwen ceyan shouce[*Combined Raven's Test. Manual*].

Wang, D., Di, M., & Qian, M.(2007). Lianhe xing ruiwen ceyan zhongguo ertong changmo disanci xiuding (CRT-C3) [A Report on the Third Revision of Combined Raven's Test (CRT-C3) for Children in China]. Zhongguo linchuang xinlixue[*Chinese Journal of Clinical Psychology*], *12*, 559-561.

Wechsler, D. (1949). *The Wechsler Intelligence Scale for Children*. New York: Psychological Corp.

Wilson, M. (2005). *Constructing measures: an item response modelling approach*. London, Lawrence Erlbaum Associates Publishers.

Wright , B. D., & Stone, M. H. (1979). *Best test design. Rasch measurement.* Chicago: Mesa Press.

Yu, M. N., Lin, X. F., & Cai, J. Y. (2001). Guoxiaoxuesheng shuxue zhishi jiegou renzhi zhenduan pingliang zhi yanjiu [Cognitively Diagnostic Assessment of Mathematical Knowledge Structures of Elementary School Children]. *Jiaoyu yu xinli yanjiu[Psychology and Education Research], 24,* 263-302.

Zeng, S. C. (2002).*Shanghaishi erzhiwu nianji xuexi zhangai pingjia liangbiao xiuding de yanjiu baogao[The research report on the revision on learning disabilities scale on 2-5 grade students in Shang Hai*]. (Master Dissertation, East China Normal University). Retrieved from http://www.cnki.net/

Zhang, H. C, Zhou, R., &Chen, G. M. (1992). Zhongguo ertong fazhan liangbiao (CDCC)[Children Developmental Scale of China (CDCC)]. ISBN:026.

Zhang, H. C, Zhou, R., &Chen, G. M. (2005). Zhongguo ertong fazhan liangbiao (CDCC)[Children Developmental Scale of China (CDCC)]. ISBN:026.

Zhou, R., & Zhang, H.C. (1994). CDCC zhongguo ertong fazhan liangbiao (3-6 sui) de bianzhi. Xinli Kexue[Introduction to CDCC chilren's development scale of China (Age 3-6)]*, Xinli kexue[Pychological Science] 3*, 137-140, 132.

Zhou, T. J. (1996). Guomin xiaoxue shuxue chengjiu ceyan[*Mathematics Diagnosis Test on Primary School Primary-grade Students(Grade 1-6)*]. Taiwaishida tejiao zhongxin[Publisher of special education in Taiwai Normal University].

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

Appendix I

Table 9 Item statistics for the construct validity (Selected item)

	Response	***category***
	0	1
Item 1-Grade 1		
Domain-Number and algebra		
Count	38	1819
Mean location	-.66	-.01
Std. Dev. Of Locations	.87	.94
Item 87-Grade 2		
Domain-shape and space		
Count	52	1803
Mean location	-1.68	.25
Std. Dev. Of Locations	2.08	1.25

Appendix II

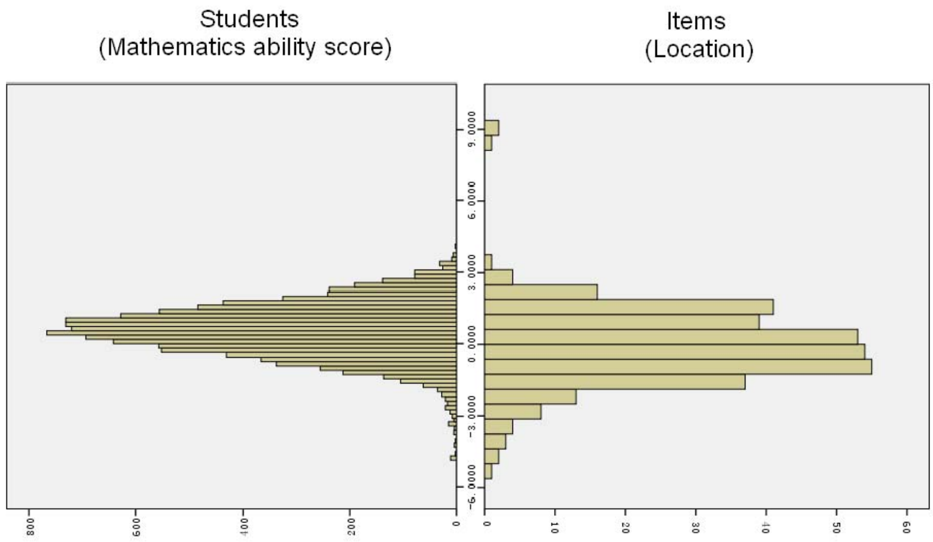
Profile of the students on mathematics performance and cognitive skills.

School: 135 Student: 135101001 Gender: Male Grade: 1

Master of the mathematic curriculum

Scaling score: .91

Description: The student have the 50% of the probability to achieve the items which difficulty is .91. See the whole picture:



Development of the cognitive skills.

NR	S	K	T	P	V	C	R	N	G	Vis	Log
1.17	1	0.81	0.88	1.05	1.01	0.8	0.24	0.9	-	0.47	0.02

Chapter 3

A multilevel analysis on predicting mathematics performance in Chinese primary schools*

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 within five provinces with different developmental levels. At the school level, the aggregated socioeconomic status of school was a significant predictor (x^2=4.3, df=1, p<.05) until individual reading level is included. At the class level, grade is a significant predictor. And teacher's level of graduation did predict 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) plays a large and significant impact. Socioeconomic status of family is a weak and polynomial predictor. The results reveal that individual variables are important predictors and explain 46.67% of the total variance. After controlling for student characteristics, school and class level variables disappeared which implies a interaction between the contextual and individual variables. Also, there are some policy implications for mathematics education in China: firstly, the education quality between regions is balance but the schools' quality within region is not balance; secondly, there is 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.

1. Introduction

Mathematics is a key component of the primary school curriculum. But, the implementation of mathematics curricula does not automatically lead to a specific

---

* This chapter is based on the accepted paper : Zhao, N.N., Valcke, M., Desoete, A, & Verhaeghe, J. (in press). A multilevel analysis on predicting mathematics performance in Chinese primary schools: Implications for practice. *Asia-Pacific Education Researcher*.

increasing 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 literatures about mathematics performance describe a range of factors potentially affecting mathematics performance. A primary set is related to the basic capabilities of the learners (Russel & Ginsburg, 1984; Silver, Pennett, Black, Fair, & Balise, 1999). These factors are difficult to be influenced, and resistant to educational interventions. Secondary factors affecting mathematics learning, 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), for example, home reading and homework support; (3) instructional environment variables, e.g., nature of the mathematics method, quality of educational interventions, teacher professional status, time investment, use of didactical tools, … (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). In the context of this article we focus on these secondary variables since – from an educational development point of view – the primary variables are difficulty to influence or change.

A vast body of empirical researches are already available that studied the single and combined impact of sets of secondary variables. Unfortunately, limited empirical evidence is available as to these variables explaining mathematics learning performance in 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 these at the school, classroom and student level. In 2001, the *Basic Education Curriculum Reform Programme (Draft) was* issued by the Chinese Ministry of Education (MOE) which aimed 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 secondary variables as discussed above. The study builds on mathematics performance data from 10,959 Chinese pupils, from primary schools in five different Chinese provinces with five developmental level. In this article, the effects of schools and classes on mathematics performance were analyzed by the multilevel. The purpose of the article is: (1) to examine the students performance across regions, school and classes; (2) to find which variables are the important predictors for the mathematics performance.

2. Theoretical Background

Several theoretical models 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 which play a role at very different levels. In the next paragraphs we structure these studies along in the following three levels: the school level, the class level, and the student level.

2.1 School level variables

As for school level variables, there are plenty of evidences that "schools" matter in terms of mathematics performance. Researches point 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). 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 region with a certain Gross Domestic Product (GDP), family SES can be strongly dependent on the former variable. Students are therefore not randomly distributed between/within regions and schools (Hanushek, Kain & Rivkin, 2004).

2.2 Class level variables

Class level variables have consistently been associated with factors affecting academic performance (Teddlie, 1994). Within the set of class variables, a subgroup is labelled as an organizational property; e.g., the quality of teachers, 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). Additionally, though teacher professional development is accepted as a significant predictor of student performance, some studies point at the interaction effect of student type (Kukla-Acevedo, 2009).

2.3 Student level variables

Besides for the demographic variables (e.g., age, gender, language proficiency, ethnicity), a large variety of personal characteristics; e.g., student's beliefs or self-efficacy, metacognition, and family variables (e.g., parental involvement, socioeconomic status) have been studied (Spelke & Ellison 2008; Tate, 1997).

Family variables. A particular set of background variables is closely 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 background are more at risk in 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 between 1990 and 2000 consistently observe a medium to strong SES–achievement relationship, which 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 by a number of authors 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 mean of family SES levels should also be considered as an aggregated variable at different levels in a research model (See Table 1).

Table 1
Varying perspectives on the composition of 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)	√	√	√	√	√	√	√	

SES is not a uni-dimensional variable. An example of such a composite variable is the educational level of the parents. Research of Alwin and Thornton (1984) has shown how both fathers' and mothers' educational level is associated with student performance. In other studies, it is especially mothers' educational level that is considered as the most critical variable (OECD, 2009). Educational level is one example of SES related variables. As becomes clear from Table 1, different authors and studies put forward a different set of variables to compose the SES value.

Demographic variables. Individual background variables, such as age, gender, language proficiency, ethnicity, … have been popular issues when studying mathematics performance (Scarr, 1988; Secada, 1992). These studies show that schooling grade is one of the most significant predictor for performance (Kyriakides & Luyten, 2009). Gender is central to the meta-analyses of Hyde, Fennema, and Lamon, (1990). Some researchers point in this context at gender stereotypes (Guimond & Roussel 2001; Meece, Parsons, Kaczala, Goff & Futterman, 1982; Spelke & Ellison 2008), or motivation (Martin, 2004). Ethnicity is also identified as a relevant but rather indirect predictor of mathematics performance, considering the mediating impact of e.g., mental style (Ginsburg, Posner, & Russell, 1981), mathematics self-efficacy (Stevens, Olivarez, Lan & Tallent-Runnels, 2004).

Individual characteristics. A group of studies centre on the relationship between a varied set of motivational beliefs and performance (e.g., motivation, attributions, regulation, participatory behaviours and engagement, self-concept) and mathematics performance (Elliott, DiPerna, Mroch, & Lang, 2004). Other researchers focus on the mediating 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 is found to play a critical role (Stevens, Olivarez, Lan , & Tallent-Runnels, 2004). And the researchers find that metacognition influences reading, spelling, mathematics and reading comprehension (Desoete, Roeyers, & Buysse, 2001).

It is also observed that reading comprehension performance is an important predictor for mathematics in primary school which suggesting that student with a higher level of reading comprehension in primary school achieve better compared to other students (Grimm, 2008).

In summary, the available theoretical and empirical research clearly describes the need to include in state-of-the-art models to describe and explain mathematics performance on the base of multiple sets of variables, such as opportunity-propensity framework (Byrnes & Miller, 2007). In contrast to the latter study in which hierarchical regression and structural equation modeling were used, the relationship between the levels in the predictor variables requires the adoption of multilevel analysis techniques (See Figure 1). 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).

In multilevel research, the data structure in the population is hierarchical, and the sample data is viewed as a multistage sample from this hierarchical population (Hox, 2002). Multilevel models assume that students and teachers are not randomly distributed to the classroom and schools while they are clustered in the classroom with different teachers (Lee & Bryk, 1989). The additional advantage is that the impact of variables - at 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 impact of the social context, and the influence of mediating intervening (e.g., instructional) processes (Hox, 2002). As explained earlier, though multilevel studies about 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.

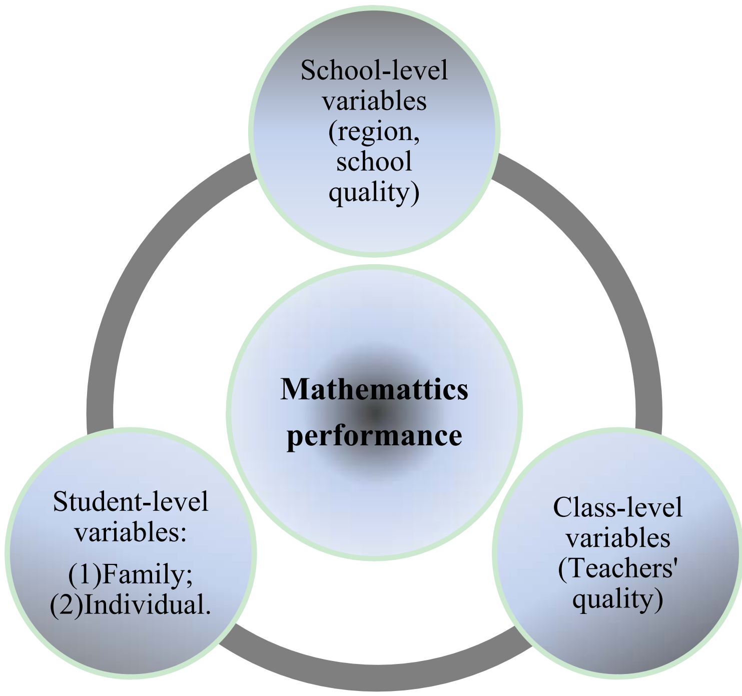


Figure 1
Model of the individual and contextual variables on mathematics performance.

3. Methods

3.1 Objectives

The present study aims at (a) examining student mathematics performance across regions, school and grades and (b) identifying student level variables, class variables and/or school variables that contribute in a significant way to the variation in performance.

3.2 Research sample

Data for this study were obtained from 10,959 primary school pupils. This is the multi-stage stratification sampling: the primary sampling units are provinces with different level of the gross development product (GDP) GDP, the secondary sampling units are administrative region within the provinces (called “city”, which include the urban and rural places at the same time) and tertiary sampling units are schools (N=20). Firstly, thirty one provinces and cities (excluding the Special Administrative Regions) were ordered into six level according to GDP(Level 1 = highest level) from report Year 2005 of the Chinese Economic Bureau. Secondly, one province or city was selected from the first to fifth level (excluding the sixth level). Thirdly, within each of the five provinces or cities, four schools were selected by random supported by the educational bureau in each city: two from urban area and two from rural area. All pupils from a specific classroom grade in a school participated in the study. The size of the selected schools ranged from 318 to 897 students (*M*=547.95, *SD*=140.19). Next to the pupils, also their classroom teachers were involved in the study. Of the 197 teachers, 73.4% were female.

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%
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%

3.4 Research variables

3.4.1 Dependent variable : Mathematics performance (MATH).

The test covers the three general mathematics domains according to recent Chinese curriculum standards: number and algebra, shape and space, statistics and probability (MOE, 2001). The test items cover an established 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). The test was designed by the mathematics teachers and experts in China for students in different schooling year (the first year students in primary school = grade one). The validity of the pool of test items was assessed by curriculum experts again.

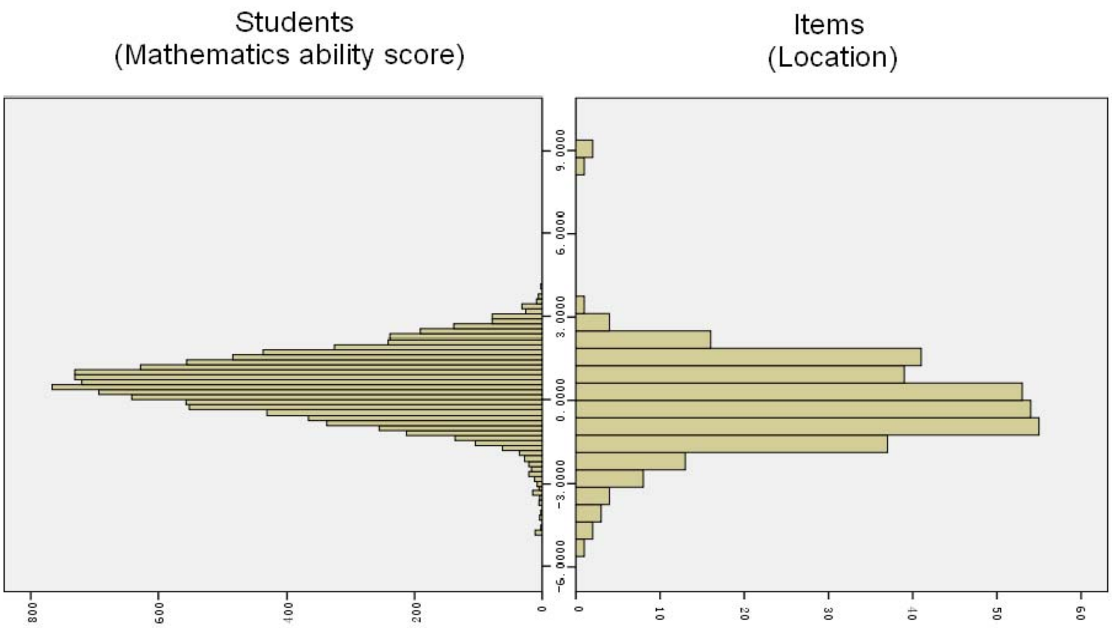


Figure 2
Distribution of students' ability estimations and item locations.

Students from each grade got different test forms with certain anchor items which aim to combine the calibration of different test forms together. A pilot study and a main study were set up to test and calibrate the mathematics performance test during May-November 2008. All the items and cases (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 .93 to .96. The test 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 Chapter 2).

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)

3.4.2 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 Product (GDP) classification scheme of the Chinese Economic Bureau. This resulted in five GDP levels ranging from 1 (highest GDP level) to 5 (lowest GDP level).

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

Some 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 wealth level.

3.4.3 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.**

Teacher's Graduation School (TGRA): This variable checks the type of teacher education institute of each individual teacher. Two categories are considered: 0 for lower level school (such as open university or self-learning system), 1 for higher level school (such as normal university or college).

3.4.4 Predictor variables at the student level

The students were asked to finish a survey including following variables supported by the teachers and parents. The family background information of the students were attained from the schools supported by the Educational Bureau in each city.

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 and mother, job level of father and mother based on the previous studies (Li, 2005a; 2005b), four variables related to wealth, income, and cultural possessions. Exploratory factor analysis (EFA) on the six FSES items suggested a two factor solution excluding the father and mother educational levels. This was 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 for the SES variable:

- *Family SES - Equipment (FSES_E):* This index reflects the number of wealth related goods is present in the family (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).

Parents' educational level: Parent's 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 as compared 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)

Birth order of the child in the family (ORCH): This variable is sensitive in the context of the Chinese family planning policy and birth order of the child are stated to impact performance (Zajonc & Markus, 1975). A first born child is often burdened by exceedingly high parental expectations.

Student mathematics self-efficacy (MSS): In order to control the possible influence of self-efficacy, a Likert-5-point mathematics self-efficacy scale (MSS; Marat, 2005) was administered. On the base of a pilot test on May 2008, 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). Since the chi-square is sensitive for the sample, we will omit it and use other coefficient to evaluate the scale.

*Metacogniton (META)***:** Following the 'post diction paradigm' - subjects 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). In this article, equation of the META is in the following: $\text{META} = (\text{Expected score} - \text{Actual score})^2$. This equation implies that if the META score is higher then the 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. Since it is difficult to compare these scores across different classes within the same school, the test scores were recorded by the percent of the correct in order to compare between different schools. Furthermore, differences between schools were cancelled out by calculating on the base of the performance scores five achievement levels: 1 for highest, 2 for high, 3 for average, 4 for low, 5 for the lowest achievement level.

3.5 Data Analysis

A multilevel analysis (programme MLWin 2.15) was used to evaluate to what degree 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 learning 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 learner 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 can 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 pupils, classes and schools

v_{0k} is random effect at the schools level, an allowed-to-vary departure from the grand mean;

u_{0jk} is random effect at the classes level, an allowed-to-vary departure from the school effect;

e_{0ijk} is randeom effect at the pupil level, an allowed-to-vary departure from classes 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 a test of 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 school and students within class. Finally, testing the full model implies the full set of predictor variables being entered in the random-coefficient regression analysis.

4. Results

4.1 Descriptive results

The descriptive analysis centers on testing bi-variate correlations by Kendall's tau between the 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. We can 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.

3.5 Multilevel Analyses

3.5.1 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 any 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 claim that school effects play a role in learner performance.

3.5.2 Hierarchical model testing

Step-by-step variables at the school level, class level and student level were added to the null model. 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 retained.

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, as a first class level variable (see model 2 in Table 5), grade is added as a predictor. The average estimate of the overall school on mathematics performance is about .000 when the value of the predictor is grade one. 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 (such as gender, graduation school and beliefs) were added to the model. This did not result in a significant improvement of the model. The model with inclusion of teacher's graduation school level (TGRA) was found to have a significant improvement in model 3 (x^2=4.84, *df*=1, *p*=.03). As will be explained later, the impact of the variable teacher's initial preparation is overruled by student level variables (See model 8a).

Table 5.
Multilevel regression of Null model and models with school, class level variables

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_{v0}	.23 (.09)	.18 (.07)	.19 (.07)	.19 (.07)
Level: Class				
σ^2_{u0}	.52 (.05)	.52 (.05)	.33 (.03)	.33 (.03)
Level: Student				
σ^2_{e0}	.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).

At a first step, specific student family background characteristics were entered in the model 4 and 5. we also 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), which implies that polynomial relationship between the family SES of job and the mathematics performance. But at the meantime, the SCFSES_J become not significant anymore (See model 4b). Adding FSES_E (SES based on wealth indicators), also results in a significant improvement of the model (x^2=12.91, *df*=1, *p*<.001). These findings suggest to be in line with Willms's (2003) hypothesis that there might be a nonlinear relationship between SES and academic performance. But these results need to be further explored. The analysis results 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, this does not result in a significant improvement of the model. 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 Models with family background variables of individual students

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 variables were added to the model. Only a few variables were found to have a significant impact. First, as can be derived from Table 5, the inclusion of gender result in a significant improvement in model 5 (x^2=15.58, *df*=1, *p*<.001). The results shows that girls significantly outperform boys on the mathematics performance scores. 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, the mathematics performance will decrease by .03 units. We can explain this by pointing at additional responsibilities that have to be adopted by older children, and higher expectations that result in higher school performance.

As a third step, other student characteristics were entered in the model. The difference in deviance between the consecutive models is statistically significant. When self-efficacy is added to the model (model 7), this again 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. It is critical to point out that by adding this variable, the coefficient of the teachers' graduation school and gender is no longer significant. Therefore, teacher's graduation school level is excluded in the model 8a (x^2=1.68, *df*=1, *p*=.19) and gender is also 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 the forth step, the performance of the Chinese language (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 drop to a high extent. This result implies that a higher level in Chinese language mastery, is related to higher mathematics performance scores.

Table 7

Multilevel regression Models with variables from individual students

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

Predictor	M5	M6	M7	M8a	M8b	M8c	M8d	M8e	M9	ES
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)	.29 (.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.

3.5.3 Full model

In the full model, 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.

Compared with the null model, a lower coefficient is observed related to school variance (.23 vs. .18), which helps to explain 21.74% in explained variance in mathematics performance between schools (See Figure 2). However, big change between classes variance (.52 vs. .33) in means can be explained by class level variables, which represent a 36.54% decreasing for classes (See Figure 2). When individual variables were entered into the model, additional within-school variance can be explained (.49 vs. .28 in Table 6). This represents a 42.85% of explained variance between individuals in mathematics performance. This is largely in line with the findings in other developing countries (Ma, 1997). In summary, we repeat that variables at the individual learner level, explain the largest proportion of variance in mathematics performance; even after the school level and classroom level variables have been controlled for. Remarkably, students' characteristics do not only explain differences between students within classes, but also more than 16.14% of the remaining variance at the school level and 28.52% at classroom level (See Figure 3). This reconfirms that Chinese primary schools still reflect the consequences of the abolished Key School policy. Certain schools seem to recruit higher performing students; at the same time these students are fostered to attain higher grades in these schools.

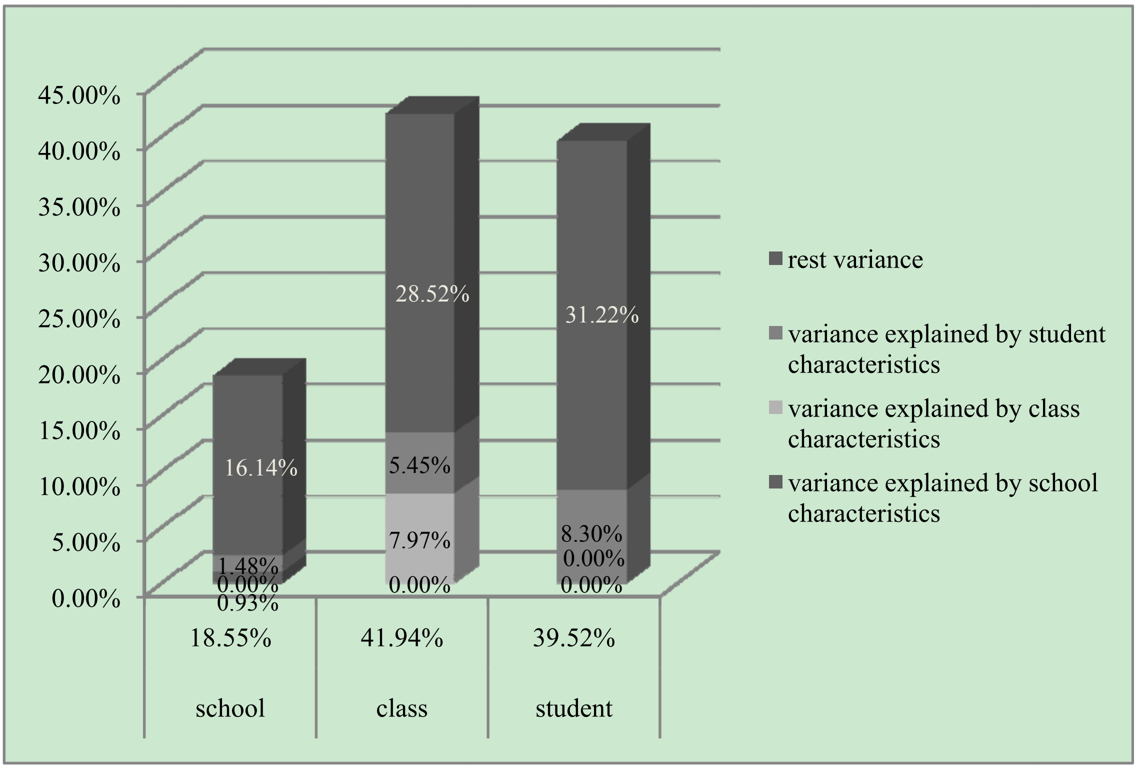


Figure 3
Contribution to the variance in mathematics performance proportioned by significant school, class and individual student level characteristics.

5. Discussion, implications and conclusions

5.1 School level effects

During recent years, studies from developing countries stress the controversy between the Coleman report and the Heyneman-Loxley effect (Huang, 2010). Coleman stressed that SES accounts for academic performance. But Heyneman and Loxley (1983) extended the argument 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 large proportion of the variance in academic performance. Differences in GDP_P and MUR are linked to large regional socio-economical differences. Gross Domestic Product (GDP_P) of the province and the urban or rural location of the schools (MUR) was considered to explain differences in mathematics performance. But both of the GDP_P and MUR variables do not appear to be significant predictors. This implies that wherever the school is located, individual students still are able to attain high performance levels in mathematics. Or, it implies that the education quality between regions is balance but the schools quality within region is not balance. The result in the next step proved this situation.

Adding – in an aggregated way – socio-economic variables at the school level shows that the job of parents plays a certain role. Nevertheless, when we allow socio-economical variables across students in classes, mathematics performance hardly varies. Initially, our results in 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, it is known that the school composition reflects specific SES-levels and how this affects the performance level of schools in a district (see comparable observations in India; Venkatanarayana, 2005). It is known that primary school performance in China differs widely within the same region not between regions. In the Chinese context, strong between-school differences can be explained on the base of *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 with high performance in each region. Students who want to enroll in these schools should obtain high performance in the entrance examination. This resulted in a situation where next to high performing school, there are middle and low performing schools. Though the Ministry of Education issued a ban on "key schools" in the mid-1990s, and the related unfair distribution of educational resources, unequal school performance is still a reality. At each province with different developmental level, there are key schools which keep the performance of the province are balance. The difference are significant between schools within one region.

5.2 Class level effects

The analysis results show that grade is a relevant predictor for student mathematics performance. This is to be expected given the systematic impact of school curricula on performance. The fact that “age” is not a significant predictor of mathematics performance, implying that the nurture effect is stronger than the nature effect in the primary school context (Morrison, Griffith & Alberts, 1997). It is also interesting to mention the differences in effect size; e.g., between grade five and grade one (the reference grade level) versus between grade six and grade one. This implies that there might be an imbalance of the difficulties of curriculum content for different grades. This result is in line with the 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 able to help the students to attain higher achievement levels, as compared to teachers from open teacher training system (e.g., open university teacher education). Teacher training has been a weak component of the Chinese educational system from the late 1970s. By 1985, still two-fifth of 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 in rural areas. This resulted in the late 1980s in a program delivered through the China Television Teachers' College and other types of open universities. The present research results reveal that although more new teachers got teacher training before they worked, teachers graduated from these institution attain lower performance levels. This suggests that a specific in-service teacher training program should be put in place to close the gap between teacher 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 extend more general university to do the teacher education (Zeng, 2008).

5.3 Student level effects

At student level, the Chinese language attainment level (CHI) 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 mathematics and language performance (e.g., reading comprehension). This result suggests, that remedial or intervention programs have to be put in place to be proactive as to difficulties of students with different language backgrounds. This is regularly mentioned in relation to multilingual contexts (Pretorius & Currin, 2010).

Metacognitive experiences and self-efficacy predict mathematics performance in a significant way. In the present study, metacognition is the most important predictor. In model 8a, comparing with model 7, there are 36.37% variance at the student level was explained by META. 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 have helped 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).

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

6. Summary, limitations and directions for further research

This study contributes to the limited understanding of the predictors for the mathematics performance in primary school in China. Firstly, among 10,959 students from west and east area, south and north area, rural and urban was covered by the study. Secondly, item response theory was used to calibrate items for reliable and valid performance assessment instrument which integrate the new curriculum content and mathematics abilities. Thirdly, the multilevel models was used to analyze the different level predictors and explore the relationship between the variables.

Though the present study was set up by involving a large sample of primary school children and focusing on a very broad area of student, class and school related variables, some limitations have to be stressed . Firstly, the specific variables considered at the school level and class level might have been insufficient to reflect the full impact of the real school situation. For example, it is possible that the variables school culture or school atmosphere have played a role (Heck, 2007; Philips, 1997; Mackenzie, 1983). Other interesting variables can be related to school policies (Lashway, 2002), educational beliefs of principals and teachers beliefs (Ross & Gray, 2006), etc. Secondly, the multilevel analysis approach only focused on mathematics performance as the dependent variable. Although this analysis technique considers the nested nature of the variables in our sample, the technique has some limitations. It is less clear along what path the variables have a direct or indirect impact on mathematics performance. Thirdly, additional research is needed to study whether the impact of determinants in the full model is the same when we consider different levels of mathematics performance or when we focus on specific developmental level of the geographical location of a school.

In future research, more detailed attention could be paid to variables at the school and classroom level. The composition effect of aggregated variables could be explored more deeply. Additionally, next to the study of mathematics as measured with the performance test, alternative research designs could be adopted to find information that corroborates the findings in the quantitative study. For instance, video-based analysis could help to analyze teacher and teaching related variables. Lastly, the results of the multi-level analysis only present a basic picture of what variables affect at different levels 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 having a causal relationship with mathematics performance.

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.

Ginsburg, H. P., Posner, J. K. & Russell, R. L.. (1981). The Development of Mental Addition as a Function of Schooling and Culture. *Journal of Cross-Cultural Psychology, 12(2)*,163-78.

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.

Guimond, S., & Roussel, L. (2001). Bragging about one's school grades: Gender stereotyping and students' perception of their abilities in science, mathematics, and language. *Social Psychology of Education, 4*(3–4), 275–293.

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.

Martin, A. J. (2004). School motivation of boys and girls: differences of degree, differences of kind, or both? *Australian Journal of Psychology, 56* (3), 133-146.

Meece, J. L., Parsons, J. E., Kaczala, C. M., Goff, S. B., & Futterman, R. (1982). Sex differences in math achievement: Toward a model of academic choice. *Psychological Bulletin, 91(2),* 324–348.

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.

Morrison, F. J., Griffith, E. M. & Albert, D. M. (1997). Nature-Nurture in the classroom: Entrance age, school readiness, and learning in children. *Developmental Psychology, 33*(2),254-262.

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.

Stevens, T., Olivarez, A., Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *The Journal of Educational Research, 97*, 208-221.

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.

Chapter 4

The mediator of the individual variables between the contextual variables and the mathematics performance*

Abstract

The present study aims at exploring predictors influencing mathematics performance. More in particular, the study focuses on internal students' characteristics (gender, age, metacognitive experience, mathematics self-efficacy) and external contextual factors (GDP of school location, parents' educational level, teachers' educational level, and teacher beliefs). A sample of 1 749 students and 91 teachers from Chinese primary schools were involved in the study. Path analysis was used to test the direct and indirect relations between the predictors and mathematics performance. Results reveal that a large proportion of mathematics performance can directly be predicted from students' metacognitive experiences. In addition, other student characteristics and contextual variables influence mathematics performance in direct or indirect ways.

1. Introduction

Students' mathematics literacy is essential for their further schooling and their success in the future work place. Therefore, exploring and understanding the factors that influence mathematics learning is an important topic. Available researches present a variety of views concerning the factors influencing mathematics performance. Those factors can be clustered into two groups: (1) internal student characteristics, such as gender (Hyde, Feenema, & Lamon, 1990), metacognition (Desoete & Roeyers, 2001), and math self-efficacy (Pajares & Graham, 1999), and (2) external or contextual variables, such as GDP (Gross Domestic Product) of the geographical school location (Young, 1998), parents' educational level (Sirin, 2005), teachers' educational level (Goldhaber & Brewer, 2000) and teacher beliefs (Mandeville & Liu, 1997).

A large body of the available studies focus on the impact of internal variables. Such studies ignore the specific and interaction effect of external variables, such as

* This chapter is based on the submitted paper of Zhao, N.N., Valcke, M., Desoete, A. Zhu, C. & Sang, G. Y.(submitted). A holistic model to predict mathematics performance: the interrelated impact of student, family and school context variables. *Scandinavian Journal of Educational Research*

family and school context. Nevertheless, contextual variables are also considered as important educational factors that are related to math performance (see e.g., Reusser, 2000). But, studies seldom focus on a holistic educational model that brings together both internal and external variables that influence mathematics learning and performance. The central aim of the present study is to study such a holistic approach when studying mathematics performance.

2. Conceptual model

2.1 Internal variables

2.1.1 Demographic variables: Age and Gender

Studies show that age is a highly significant predictor of mathematics performance (Kyriakides & Luyten, 2009). Gender - as proposed in the meta-analysis research of Hyde, Feenema and Lamon (1990) and Else-Quest, Hyde and Linn (2010) – clearly predicts learning performance. However, only a limited amount of studies have explored gender differences in mathematics performance at primary school level (Fennema, 1974; Hyde, et al., 1990). In addition, the available empirical evidences show that gender difference tends to decrease (e.g., Eisenberg, Martin, & Fabes, 1996; Hyde & Mertz, 2009) or even disappear with age (e.g., Frost, Hyde, & Fennema, 1994; Pajares & Graham, 1999).

2.1.2 Internal variables and mathematics performance

2.1.2.1 Metacognition and mathemactics performance

Recent studies suggest that metacognition is a significant predictor of learning performance in general and mathematical performance in particular (e.g., Veenman, Van Hout-Wolters, & Afflerbach, 2006). Flavell (1979) defined the concept of "metacognition" as "thinking about thinking". Furthermore, metacognition can be defined in terms of metacognitive knowledge and metacognitive experiences (Efklides, 2001, 2008; Flavell, 1981).

In the one hand, metacognitive knowledge includes information about tasks, strategies, and goals (Flavell, 1979). Research points out that when learners are sufficiently aware of their metacognitive knowledge and therefore the way their own mind works, the lower mathematics achiever can learn better after intervention by the metacognitive knowledge (Maqsud, 1998).

In the another hand, metacognitive experience is "what the person is aware of and what she or he feels when coming across a task and processing the information related to it" (Efklides, 2008, p.279). They take the form of metacognitive feelings, and metacognitive judgments/estimates. Metacognitive experiences make the learner aware of his/her cognition and trigger control processes that serve the pursued goal of the self-regulation process (Efklides, 2006; Efklides, 2008). When students have metacognitive experiences and know how to capitalize on these experiences, they have more chance to be a successful mathematics problem solving endeavour (Foong, 1993).

As Flavell(1987) reveals, the young children have more trouble than older children in metacognitive experience, such as comprehending their won feelings of incomprehension. Students with good metacognitive experience will achieve higher on mathematics than the their peers (Desoete, & Roeyers, 2001). Then in this paper, the focus will be on the metacognitive experience of young students in primary school.

2.1.1.2. Mathematics self-efficacy and mathematics performance

Self-efficacy can be defined as the belief in one's capacity to organize and execute actions required to attain a level of performance (Bandura, 1993; 1997). Previous studies reveal that the mathematics performance is correlated to the math self-efficacy (Hackett & Betz, 1989) and the mathematics problem solving is affected by the math self-efficacy (Pajares & Graham, 1999). Students with a higher level of self-efficacy adopt a wider variety of cognitive strategies and reflect a higher level of cognitive engagement (Pintrich & DeGroot, 1990).

There are also some interrelations between the internal variables themselves. Gender is assumed to affect the self-efficacy (Betz & Hackett, 1981; Hackett & Betz, 1989). But the relationship between gender and mathematics self-efficacy is still unclear (Post-Kammer & Smith, 1985; Skaalvik & Rankin, 1994). Self-efficacy mediates the effect of gender and prior experience on math problem-solving performance (Pajares & Graham, 1999). Also, studies show that metacogntive experiences control the impact of self-efficacy on performance (Akama, 2006; Panaoura, 2007).

2.2. External contextual factors

2.2.1. Family related variables and mathematics performance

SES is a complex variable that comprises – depending on the author or study – a different set of variables. parental educational level, parental occupation and home resources or wealth (Sirin, 2005). The link between the socioeconomic status (SES) of parents and mathematics performance has been subject of numerous studies; see e.g., the TIMSS and PISA research (Marks, 2006; Ming & Zeng, 2008; Webster & Fisher, 2000). Students with highly educated fathers and mothers perform considerably better than the other student which parents hold a medium schooling degree (Fertig, 2003). Of the predictors of mathematics performance at age 10, the effect size of mother education level is higher than the father education level (Melhuish, Sylva, Sammons, Siraj-Blatchford, Taggart, Phan, & Malin, 2008).

Besides the mathematics performance, parents' SES also influence the development of students' internal variables. For example, Vygotsky (1978) and Wertsch (1985) state that metacognition is affected by family social interactions. Previous research reveals explicitly how metacognition is related to environmental factors, such as the socioeconomic status (Pappas, Ginsburg, & Jiang, 2003), collaboration styles of mothers with their preschoolers during problem solving (Moss, 1990) and family culture (Eills, 1997). In this context, Schommer (1990) shows that higher educated parents expect to a larger extent that their children take up responsibilities at home and expect their children to think more independently. In addition, a supportive parenting style has proven to lead to higher levels of self-efficacy and subsequent school achievement (Whitbeck, Simons, Conger, Wickrama, Ackley, & Elder, 1997).

2.2.2. Teacher quality and mathematics performance

It is widely accepted that learning is influenced by a variety of academic contextual elements (e.g., Salomon & Perkins, 1997), such as teacher quality defined by their educational level, (Mandeville & Liu, 1997; Smith, Desimone, & Ueno, 2005). Teachers who have a standard certification have a statistically significant positive impact on student math test scores while teachers hold other certification or are not certified do not have the impact (Goldhaber & Brewer, 2000).

Teachers' beliefs about mathematics teaching have been revealed to influence mathematics performance in general (Van Steenbrugge, Valcke, & Desoete, 2009), and mental calculation in particular (Stigler, 1984). Previous studies show that teachers with cognitive constructivist orientation was associated with their students' larger achievement gains in mathematical word problems (Staub & Stern, 2002).

Also, teachers have a direct and indirect impact on math score and on mediating internal variables. The teacher impact on internal variables is found in studies about metacognition. It has been shown that different teaching methods might hinder or improve metacognitive processing (Nist, Holschuh, & Sharman, 1995; Van Keer & Verhaeghe, 2005). When it comes to the impact on self-efficacy, Siegle and McCoach (2007) reveals that teaching methods improve the students' mathematics self-efficacy.

2.2.3. Contextual variables and mathematics performance

The Chinese educational context is different from other countries. In addition, also within China, regions differ widely as to their economical activity. This is reflected in a large differences in the regional Gross Domestic Product (GDP). The regional GDP will affect mediating variable that impact mathematics performance, such as the investments in schools, instructional media, teacher professional development, etc. (Perry & McConney, 2010).

2.3. Towards a holistic conceptual model

Given the fact that most previous studies focus either on the relationship between internal or external variables that affect mathematics learning and performance, the present study adopts a holistic model approach that includes all these variables in a model to study mathematics performance. From a theoretical point of view, this is meant to be an important addition to the existing literature about mathematics education.

Figure 1 represents our conceptual model in a graphical way. Elementary mathematics can be seen as a broad domain, comprising various subdomains such as arithmetics and numerical facility skills (Desoete, Stock, Schepens, Baeyens, & Roeyers, 2009; Dowker, 2005). In our model, mathematics performance (MP) and mental calculations (MC) are regarded as dependent variables. MP represents complex mathematics performance while MC represents basic number retrieval processes.

We further distinguish internal variables such as grade, gender, mathematics self-efficacy (MSS), and metacognitive experiences measured by metacognition calibration score (MCS). As external variables, the Gross Domestic Product (GDP) of the region are positioned as contextual variables. The father's (FEL) and mother's educational level (MEL) are represent variables in the family context. Further, a teacher's educational level (TEL) and teacher beliefs are positioned in the school context. The following beliefs arte included in the study: Teacher beliefs about Student Learning (SL), Teacher belief about Stage of Learning (L), and Teacher beliefs about Teaching Practices (TP).

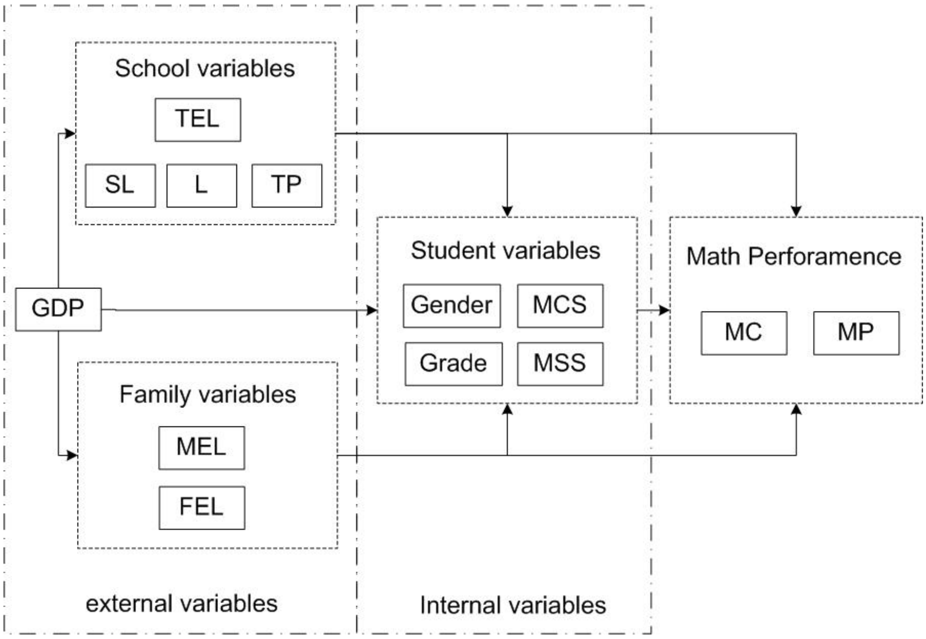


Fig.1 Integrated model of the impact of various internal and external variables on mathematics performance.

Note: GDP-Gross domestic product; MEL-Mother Educational Level; FEL-Father Educational Level; MP-Mathematics Performance; MSS-Mathematics Self-efficacy Score; MC- Veracity of Mental Calculation; Gender-Student's Gender; TEL-Teachers' Educational Level; SL-Teacher's belief on Student Learning; L-Teacher's belief on Stage of Learning; TP-Teacher's belief on Teacher Practice; MCS-Metacognition Calibration Score.

Considering the proposed conceptual model, three research questions are put forward:

(1) Which internal variables contribute to mathematics performance in elementary schools in the Chinese context?

(2) Which external variables influence mathematics performance in elementary schools in the Chinese context?

(3) What is the interaction between internal factors and external factors? How do internal factors mediate the relationship between external factors and mathematics performance?

3. Research Design

3.1 Sample

A sample of 1749 pupils (female = 49%) was involved in the study. In addition, the teachers of these pupils and information about their school was included in the study. The sampling was based on the following stratification variables: pupils are enrolled in grade two to grade six in 18 different schools, from five provinces in China, reflecting different levels in gross domestic product (see Table 1). The GDP distribution shows that 58.66% of pupils originate from a high GDP province and 41.34% from a low GDP province. Within each GDP level, there were equal numbers of boys and girls. Research data also includes information from 91 teachers, of which 3.30% got a senior school degree, 36.26% obtained a pre-Bachelor degree, and 60.44% of the teachers got a Bachelor degree.

Table 1
Stratification variables in the research sample.

Sample		Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Total
Students							
GDP	High	125	229	270	237	165	1026
of province	Low	129	210	128	148	108	723
Gender	Boys	129	221	219	199	122	890
	Girls	125	218	179	186	151	859
Total		254	439	398	385	273	1749
Teachers							
TEL	Senior school of teacher education	1	0	1	1	0	3
	Pre-Bachelor	8	8	3	9	5	33
	Bachelor	6	10	14	13	12	55
Total		15	18	18	23	17	91

3.2 Research instruments

Questions about background variables such as gender and grade were included in the students' questionnaire. Information about parents' educational level was obtained from the teacher. The gross domestics product index (GDP) was derived from the 2005 report of the Chinese Economic Ministry. The teachers were asked to fill out the teacher questionnaire that comprises beliefs related research instruments and questions about other background variables; e.g., their educational training level (TEL).

3.2.1 Mathematics performance

A mathematics test was designed for this study, with different forms for each grade. The test covers the three general elementary mathematics domains: number and algebra, shape and space, statistics and probability (MOE, 2001). In each test form, anchor items were defined in each form in order to be able to calibrate all the different test forms. This comprehensive mathematics performance test was analyzed by Item Response Theory (IRT). Mathematics Performance (MP) was calibrated with the BiLog-MG3 programme. The internal consistency (Cronbach's α) of each grade level test ranges from 0.93 to 0.96.

Research points at a positive relation between mathematics performance and mental calculations (Adams & Hitch, 1997; Ashcraft & Kirk, 2001; Hitch, 1978). Therefore, next to a mathematics performance test, we also administer the Arithmetic Number Fact Test (Tempo Test Rekenen, TTR; De Vos, 1992). The TTR is a mental calculation test, presenting pupils with 200 arithmetic number-fact problems (e.g. 5 x 9 =…). Subjects have to solve as many number-fact problems as possible in 5 minutes time. The subjects were presented with a Chinese version of the test. The test helped to determine a mental calculation scores (MC) that build on an effective and efficient basic number fact retrieval.

3.2.2. Metacognitive experiences

There are different methods of assessing metacognition (Desoete, 2008; Veenman et al., 2006). Self-ratings are usual measures to determine metacognitive experiences. Calibration studies - in the context of primary education - have been proven to result in a reliable measurement of metacognitive experiences (Desoete & Roeyers, 2006). The studies where a comparison is made between the predicted success or failure in carrying out a task and the actual performance quality after the task has been carried out (Desoete & Roeyers, 2006; Grimes, 2002). In higher education, Grimes (2002) revised the calibration approach introduced by Lichtenstein and Fischhoff (1977), again resulting in a reliable measurement of metacognitive experiences.

After administration of the mathematics performance test - following the post diction paradigm' - subjects were invited to predict the level of their test performance (e.g., 'I think I will obtain 70/100 on this test'). In line with Lichtenstein and Fischhoff (1977) and Grimes (2002), a metacognition calibration score (MCS) was calculated in the following way:

$$\text{Metacognition Calibration Score} = \frac{(\text{Actual score} - \text{Expected score})^2}{\text{Expected Score}}. (1)$$

3.2.3. Mathematics self-efficacy

A Mathematics self-efficacy scale (MSC) was developed on the base of the instrument of Marat (2005). The original scale is based on twelve items and have 85 items. For example, the first item ask "How well do you believe you can calculate accurately numerical problems mentally?" Respondents have to indicate their reaction to each item on a Likert scale, ranging from *Not well at all* (coded 1) to *very well* (coded 5). The instrument was presented to Chinese primary school learners in a pilot study prior to the present study. Items were deleted with a item-total correlation <.30. Cronbach's alpha reliability of the final version was .97. A one-factor model was confirmed on 77 items by Exploratory Factor Analysis (EFA, principle component analysis with orthogonal-varimax-rotation). This single component accounted for 34.41% of the item variance. The eigenvalue of this single factor was 22.92.

3.2.4. Teacher beliefs

The teachers completed the Mathematics Beliefs Scales (MBS) developed by Fennema, Carpenter, and Loef (1990). The Chinese version of the scale consists of 16 of the original 18 items. It is structured into three subscales: (1) teacher beliefs about how children learn, labelled as the student learning factor (6 items); (2) beliefs about the teacher role to teach computational and application skills, labelled as the stages of learning factor (4 items); (3) teacher beliefs about teacher practices (6 items). Item Likert scale categories ranged from *Not agree at all* (= 1) to *agree very well* (= 5). The survey was completed by 83.33% of the teachers; some teachers could not attend the administration session due to unforeseen timing problems. The reliability of the whole scale is .81 (Cronbach's alpha). The reliability of the subscale are .68, .65 and .62, respectively.

3.3 Data analysis

A variety of statistical procedures was applied in line with the research questions. Firstly, correlation analysis was applied to test associations between the variables in the model. Secondly, in order to test the complete model, structural equation modelling (AMOS 6.0) was applied to test direct and/or indirect relationships (Arbuckle, 2005).

4. Results

4.1 Description and correlation analysis

Table 2 summarizes the description of the variables in our study according to the endogenous and exogenous student, family and teacher characteristics. At the general level, the means of the MCS is 10.64 (*SD*=11.88) and the means of MP is .83 (*SD*=.96).

Table 2
Description of the Characteristic of demographic variables

Item	Options	Number	MCS Means (*SD*)	MP Means (*SD*)
GDP	High	1026	9.07 (11.48)	1.05 (.96)
of province	Low	723	12.86 (12.10)	.52 (.86)
TEL	Senior school of teacher education	53	15.37 (13.67)	.38 (.65)
	Pre-Bachelor	596	11.63 (12.45)	.57 (1.07)
	Bachelor	110	9.87 (11.39)	1.00 (.86)
Grade	Grade 2	254	9.09 (13.86)	.48 (1.15)
	Grade 3	439	10.62 (10.55)	.62 (.85)
	Grade 4	398	11.64 (12.33)	.73 (.94)
	Grade 5	385	12.27 (12.08)	1.22 (.94)
	Grade 6	273	8.33 (10.44)	1.10 (.71)
Gender	Boys	890	10.66 (11.88)	.84 (.95)
	Girls	859	10.61 (11.89)	.82 (.97)
FEL	No experience in school	13	12.71 (9.58)	.51 (.55)
	Primary school	190	10.70 (10.92)	.58 (.90)
	Junior school	484	11.69 (13.20)	.73 (.92)
	Senior school	553	10.74 (12.24)	.87 (.99)
	High school	409	9.21 (10.27)	1.01 (.94)
	Postgraduate or higher	100	10.45 (10.98)	.91 (1.02)
MEL	No experience in school	23	10.25 (10.22)	.54 (.65)
	Primary school	268	11.07 (10.49)	.64 (.90)
	Junior school	498	12.00 (13.57)	.74 (.97)
	Senior school	490	10.32 (12.09)	.91 (.97)
	High school	386	9.10 (10.44)	1.01 (.94)
	Postgraduate or higher	84	10.23 (10.07)	.77 (.96)
Total		1749	10.64 (11.88)	.83 (.96)

Note: GDP-Gross domestic product; TEL-Teachers' Educational Level; FEL- Father Educational Level; MEL-Mother Educational Level.

Table 3 gives an overview of the bivariate correlation between the research variables in our model. The results reflect significant interrelationships between all variables. Higher levels of mathematics performance was correlated with higher metacognitive experience (=smaller difference between expected score and real score) on MCS (r=-.66, $p<.00$). This result is in line with the result of previous study that there is significant correlation between metacognitive experiences and mathematics performance from grade 3 through grade 5 (Sperling et al., 2002). The mathematics performance decreased from lower deviance between actual score and predicted score to higher deviance of metacognitive experiences.

Table 3
Bivariate correlation between research variables in the conceptual model (*n=1,749*).

	MCS	**MP**	**MC**	**MSS**	**MEL**	**TEL**	**L**
MP	-.68***	-					
MC	-.12***	.21***	-				
MSS	-.16**	.33**	.13**	-			
MEL	-.07**	.12**	.01	.19**	-		
TEL	-.10**	.22**	.15**	.10**	.04	-	
L	-.07**	.15**	.03	.15**	.09**	-.08*	-
GDP	.16**	-.28**	-.09**	-.26**	.28**	-.17**	-.15**
Means	10.64	.83	.97	3.91	2.69	2.60	4.11
SD	11.88	.96	.09	.57	1.15	.55	.58

*p<.05, **p<.01,***p<.00

Note: GDP-Gross domestic product; L-Teacher's belief on Stage of Learning; MCS-Metacognitive experiences as Calibration Score; MEL-Mother Educational Level; MP-Mathematics Performance; MSS-Mathematics Self-efficacy Score; TEL-Teachers' Educational Level; MC- Mental Calculation.

4.2 Path analysis models

Three consecutive models were tested in this analysis approach. In a first model (Model A), internal characteristics were included and linked to the dependent variables. In a second model (Model B), the effects of the external family contextual variables were added. In the third final model, the additional effect of external school variables was explored. Also, the structural integrity of the model was tested. For reasons of parsimony, variables with non significant regression weights are not reported in Table 4. In view of decisions about the number, type and cut-off values for Goodness-of-fit criteria, we built on the work of a variety of authors (e.g., Schermelleh-Engel, Moosbrugger & Müller, 2003; Shulruf, Hattie & Dixon, 2007). The following "goodness-of-fit" indices were adopted: relative chi-square (χ^2/df) index, Goodness-of-fit index (GFI); adjusted GFI (AGFI) and Normed Fit Index (NFI) that makes the calculations independent of degrees of freedom (cut-off value $\geq$ 0.95), the Root Mean Square Error of Approximation (RMSEA, cut-off value 0.08).

Table 4
Overview of the direct effects on MP: Standardised regression coefficients (β) and fit indices (n=1,749)

		MP		
		Model A	Model B	Model C
Basic numerical facility	Mental Calculation (MC)	.06***	.10***	.08***
Internal system				
	Grade	.23***	n.s.	n.s.
	Metacognition Calibration Score (MCS)	-.65***	-.62***	-.61***
	Mathematics Self-efficacy Score (MSS)	.19***	.18***	.17***
External system				
Non-academic				
	GDP	-	-.12***	.09***
	Mother's educational Level (MEL)	-	n.s.	n.s.
Academic				
	Teacher Educational Level (TEL)	-	-	.13***
	Belief on Stage of Learning (L)	-	-	.08***
Adjusted R^2		.58	.54	.56
Chi-squre		.81	5.47	8.43
df (p-value)		1 (.06)	4 (.24)	7 (.30)
GFI		1.00	1.00	1.00
AGFI		1.00.	1.00	.99
NFI		1.00	1.00	.99
RMSEA		.00	.01	.01
AIC		1707.75	39.466	66.43

Note. – not included in model, n.s. not significant; *p<.05, **p<.01,***p<.005

Note: GDP-Gross domestic product; L-Teacher's belief on Stage of Learning; MCS-Metacognition Calibration Score; MEL-Mother Educational Level; MP-Mathematics Ability; MSS-Mathematics Self-efficacy Score; TEL-Teachers' Educational Level; MC- Mental Calculation.

In the first model 60% of the variance in mathematics performance (MP) can be attributed to the ability of mental calibration of basic number retrieval (MC), metacognitive experience (MCS), Mathematics self-efficacy (MSS) and grade. The Grade (β = .23), MCS (β = -.65), MSS (β = .19), MC(β = .06) are found to be predictors of mathematics performance. In the second model – after adding the family variables – the coefficient of direct effect from the mathematics performance changes. GDP of the school level affects directly mathematics performance (β = -.12). Father educational level do not play an important role on the mathematics performance while mother educational level have an indirectly influence through mathematics self-efficacy. And the effect of influences of Grade on mathematics performance disappeared. In the third model, when academic variables are added to the model, 55% of the variance in mathematics performance can be attributed to the complex interplay of the variables. Both direct and indirect effects on MP can be observed.

The final path model is presented in Fig.2, reporting the standardized path coefficients.

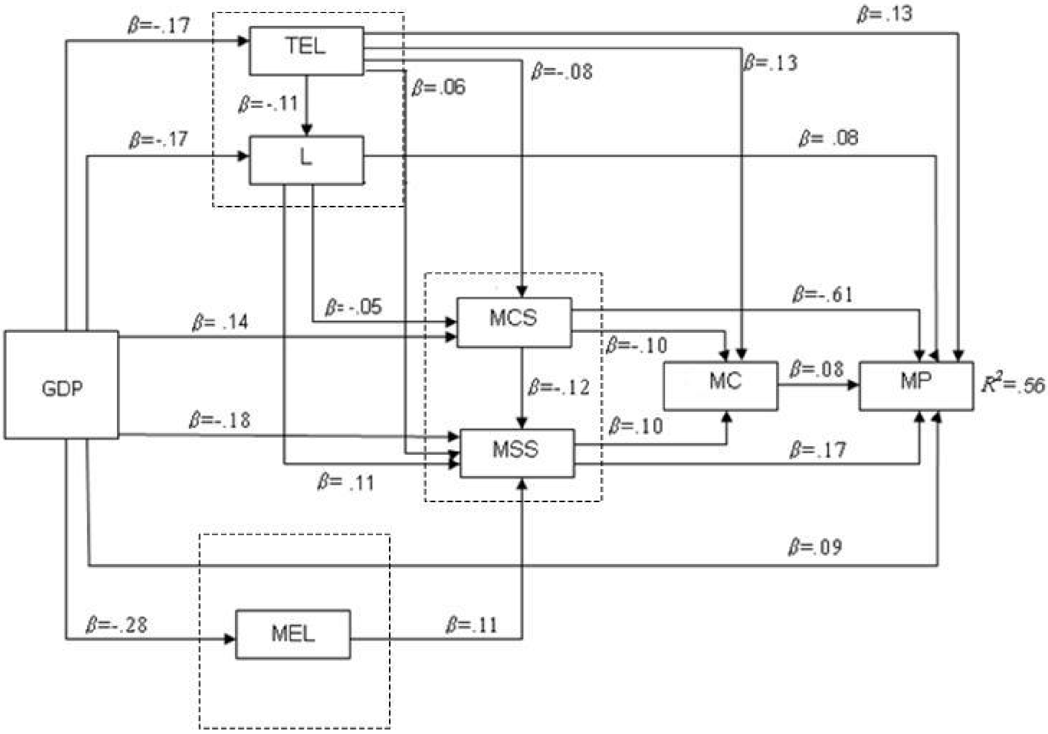


Fig. 2
Result of the path analysis

Note: GDP-Gross domestic product; L-Teacher's belief on Stage of Learning; MCS-Metacognition Calibration Score; MEL-Mother Educational Level; MP-Mathematics Ability; MSS-Mathematics Self-efficacy Score; TEL-Teachers' Educational Level; MC- Mental Calculation.

In the third model, the impact of MC on MP is not that important (β= .08). Here, arithmetical mental calculation can be assessed to control for the children with deficits in semantic memory (Ashcraft, 1992; Dehaene, 1992; Logie, Gilhooly, & Wynn, 1994). Students with learning difficulties often have problem with basic mental calculation tasks, especially due to deficits in semantic memory (Wilson & Swanson, 2001). This suggests that once children reach a baseline level, complex mathematics problem solving does no longer depend largely on the basic fact number retrieval system.

The results reveal that the metacognitive experiences (MCS) clearly affects mathematics performance (β = -.61). Mathematics performance is negatively associated with the metacognition calibration score (MCS). Students with smaller difference between expected score and real score (=higher metacoginitive experiences) reflect a higher mathematics performance. This means that the students, who are able to predict their score more accurately, end up with a higher mathematics performance. In addition, student with a higher mathematics self-efficacy(MSS) reflect a higher mathematics performance (β = .17). Considering the different aspects in mathematics performance, MC is clearly linked to MCS (β= -.10) and MSS (β= .10).

When considering the external variables, significant findings can be reported in the way that specific school and family context are related to mathematics performance. Comparing to the internal variables, the coefficients of external variables (GDP, β= -.09; TEL, β= .13; L, β= .08) are small but nevertheless statistically significant. Students with a higher educational level in a province with a higher GDP level obtain higher mathematics performance scores. The educational level of the mother (MEL) does not seem to have an direct impact on mathematics performance by the mediating of students' mathematics self-efficacy (β= .11).

Regarding the school variables, it is interesting to consider how they are intertwined with metacognitive experiences and mathematics self-efficacy as mediating variable. Teacher educational level (TEL, β = -.08) and teacher beliefs about the stage of learning (L, β = -.05) are related to the metacoginitive experiences (MCS) and mathematics self-efficacy of pupils. Teacher educational level (TEL, β = .06) and teacher beliefs about stage of learning (L, β = .11) are positively related to mathematics self-efficacy (MSS). It is interesting – in this context – to see that teachers who have students with higher self-efficacy and higher metacognitive experiences, tend to reflect higher belief levels about the need to sequence the teaching of computational skills in the classroom. This also implies that, although Chinese teachers strictly sequence the stage of teaching and learning for mathematics curriculum, does not restrict the development of students' metacognitive experiences and mathematics self-efficacy. This seems to be in conflict with common conceptions about student-centred learning. But this has to be understood from the Chinese context. Also other authors referred in this context to the Paradox of Chinese Learner, which means that the seemingly unfavorable learning environment (focusing on rote learning and highly structured) yet produces students who outperform their counterparts in the West (Biggs & Watkins, 1996; Marton, Dall'Alba & Lai, 1993).

Focusing on the impact of family context, the GDP level of a school affects the level of metacognitive experiences (MCS, β= .14), and mathematics self-efficacy (MSS, β= -.18). Students enrolled in schools that are located in provinces with a higher GDP tend to predict the mathematics score more accurately and reflect a higher level of metacognitive experience. Mother educational level is related to MSS (MEL, β=.11) but not to MCS. We can assume that a higher mother's educational level implies that she expects her children to take more responsibilities at home and in relation to their thinking and learning (Schommer, 1990). Also, it will be more likely that children will develop a higher level of self-efficacy.

In summary, the results of path analysis indicate that the external variables clearly affect internal variables and play as such an direct and indirect role in mathematics performance. A number of external variables add more explanatory power to the model. Compared to the other variables in the model, metacognitive experiences and mathematics self-efficacy are clearly dominant predictors for mathematics performance. This result is in line with previous studies.

5. Discussion, Limitations, and Conclusion

The aim of the present study was to re-examine the impact of students' characteristics, family and school context on mathematics performance of primary students in China from bio-ecological and transactional perspectives. Compared to earlier studies, additional variables were added to a conceptual model to study the direct and indirect impact of internal and external variables. Firstly, the structural equation modelling (SEM) confirmed that internal variables such as students' metacognitive experiences and self-efficacy play an important role on the mathematics performance. And also, students' mother's educational level, teachers' educational level and teachers' beliefs on the stage of learning were also related to mathematics performance of primary school children. Secondly, the study explore the interaction between the variables and provide an overview of the relationship between the variables and between the variables and math performance. This study seems to be more close to real educational context. Thirdly, the study provides the different cultural results to the existing studies. The sample of this study covers the students from the grade 2 to grade 6 in rural and urban areas.

In answer to research question 1, the results suggest that the largest proportion in mathematics performance variance could be explained by the internal variables; especially metacognitive experiences and self-efficacy. These results are in line with Kruger and Dunning (1999) and Kruger (2002). The knowledge that underlies mathematics ability is also the knowledge that underlies the ability to solve mathematical problems. Students with poor mathematics performance scores tend to overestimate their performance. As stated earlier, this shows how underachievers are presented with a dual burden: poor performance and poor metacognitive experiences.

As to research question 2, the path analysis results show that external factors such as teacher quality and teacher beliefs, mother educational level are important to be included in the model. Although mother education level does not have a direct impact on the mathematics performance, it has a influence on the mathematics self-efficacy and indirectly affects the mathematics performance. In school, teacher's quality and beliefs affect the internal variables and mathematics performance.

For the research question 3, our data reveals that metacognitive experience control the impact of mathematics self-efficacy as we can see in a previous study (Akama, 2006). And the external factors have direct influence on mathematics performance and indirect influence through the internal factors. Teachers, who adopt the belief of stage of learning that strictly sequenced mathematics teaching is important, are linked to higher metacognitive experiences and mathematical self-efficacy in their students. These results are in line with the findings of An, Kulm and Wu (2004) who compared the mathematics teachers' knowledge in U.S. and China and concluded that Chinese teachers especially emphasized the acquisition of both procedural and conceptual knowledge, which might explain the higher results of Chinese children on mathematical tasks.

Another interesting finding is that mother's educational level (but not father's educational level) is indirectly related to the mathematics performance of their children. Mothers seem to influence mathematical self-efficacy, but not the metacognitive experiences of their children. This is in line with prior research demonstrating that mothers are more involved in their children's education than fathers (Epstein, 1986; Princiotta, Flanagan, & Germino Hausken, 2006 and with the results of Davis-Kean (2005) revealing that SES is indirectly related to children's performance via parents' beliefs and behaviour. Students with low-SES backgrounds were exposed to greater risks in mathematics performance (Borman & Overman, 2004; Coleman et al., 1966). The level of mother's education improves the mathematics self-efficacy of their children.

These results should be interpreted with care, since there are clearly limitations to the present study. Firstly, the findings of this study only refer to Chinese children and need to be replicated in other countries. Moreover, this study only included metacognitive experiences and did not take into accounts more complex metacognitive skills and knowledge. Additional information in relation to classroom variables (teaching approach, textbooks used, homework …) and family context (extra schooling activities at home, impact of brother or sisters …) can be added to our model. Some external variables have been measured via the teacher (experience, educational level, gender). This can be criticized since our model wants to study the interplay between internal and external variables at the level of individual learners. In addition, metacognitive experiences with the expected score have some relation with student self-efficacy. More studies should be done in this area. In addition, in the present study, learners are approached as individuals. This can be criticized since the learners are nested within classes, within schools and within regions. This reflects a multilevel structure that should be respected when analyzing the impact of the variables on mathematics performance. Future studies should adopt a multilevel approach in the analysis of the data. Lastly, the SEM analysis approach was helpful to study the direct and indirect relationship between models, but it remains yet unclear whether all the relationships should be interpreted as causal relationships. More theoretical and empirical research is needed to underpin the nature of these relationships.

Despite these shortcomings, the study was helpful to illustrate the internal and external factors that are related to student mathematics performance at the primary school level. From a bio-ecological and transactional perspective (Kaiser, Hester, & Mc Duffie, 2001), a person's cultural worldview constitutes a social and cultural difference and it causes differences in learning performance (Aleven, McLaren, Roll, & Koedinger, 2006). It might therefore be important to add these variables to the assessment approach and intervention strategies for students at risk in view of mathematical learning difficulties. In addition, our data suggests that in future research about internal variables and mathematics performance, other external variables should be incorporated, such as mothers' educational level and family and teacher related variables.

References

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology*, 67, 21-38.

Akama, K. (2006). Relations among self-efficacy, goal setting, and metacoginitve experiences in problem-solving. *Psychological Report, 93*, 895-907.

Aleven, V., McLaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence and Education*, *16*, 101-128.

An, S.H., Kulm, G., & Wu, Z. H. (2004). The pedagogical content knowledge of middle school, mathematics teachers in China and The U.S. *Journal of Mathematics Teacher Education, 7,* 145-172.

Arbuckle, J. (2005). Amos 6.0 *update to the Amos user's guide*, SmallWaters Corporation, Chicago.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44,* 75-106.

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology*, *130*, 224-237.

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

Bandura, A. (1997). *Self-efficacy : The exercise of control*. New York: W.H. Freeman.

Betz, N. E., & Hackett, G. (1981). The relationship of career-related self-efficacy expectations to perceived career options in college women and men. *Journal of counseling psychology, 28,* 399-410.

Biggs, J. B., & Watkins, D. A. (2001). The Chinese learner in retrospect. In D. A. Watkins & J. B. Biggs (Eds), *The Chinese Learner: Cultural, psychological, and contextual influences* (pp.269-285). Hongkong: Comparative Education Research Center, The University of Hong Kong; Melbourne, Australia: Australian Council for Education Research.

Borman, G., & L. Overman.(2004). Academic Resilience in Mathematics Among Poor and Minority Students. *The Elementary School Journal, 104*, 177-196.

Coleman, J., E. Campbell, C. Hobson, J. McPartland, A. Mood, F. Weinfeld, & York., R. (1966). *Equality of educational opportunity*. Washington, D. C. :U.S. Government Printing Office.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: the indirect role of parental expectations and the home environment. *Journal of Family Psychology, 19,* 294-304.

De Vos, T. (1992). *Tempo Test Rekenen. Handleiding [Arithmetic Number Facts test. Manual]*. Berkhout: Nijmegen.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1-42.

Desoete, A. (2008). Multi-method assessment of metacognitive skills in elementary school children: How you test is what you get. *Metacognition Learning, 3,* 189-206.

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

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

Desoete, A., Stock, P., Schepens, A., Baeyens, D., & Roeyers, H. (2009). Classification, seriation, and counting in Grades 1, 2 , and 3 as two year longitudinal predictors for low achieving in numerical facility and arithmetical achievement. *Journal of Psychoeducational Assessment*, 27, 252-264.

Dowker, A. (2005). *Individual Differences in Arithmetic. Implications for Psychology, Neuroscience and Education.* New York: Psychology Press.

Efklides, A. (2001). Metacognitive experiences in problem solving: Metacognition, motivation, and self-regulation. In: A. Efklides, J. Kuhl and R.M. Sorrentino, Editors, *Trends and prospects in motivation research* (pp. 297–323), The Netherlands: Kluwer, Dordrecht.

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

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*, 277-287.

Eills, S. (1997). Strategy choice in socio-cultural context. *Developmental Review, 17*, 490-524.

Eisenberg, N., Martin, C. L., & Fabes, R. A. (1996). Gender development and gender effects. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 358-396). New York: Simon & Schuster Macmillan.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103-127.

Epstein, J. L. (1986). Parents' reactions to teacher practices of parent involvement. *Elementary school Journal, 86*, 277-294.

Fennema, E. (1974). Mathematics learning and the sexes: A review. *Journal of Research in Mathematics Education. 5,* 126-139.

Fennema, E., Carpenter, T., & Loeg, M.(1990). Mathematics beliefs scales. Madison, WI: University of Wisconsin, Madison,

Fertig, M. (2003). Who's to blame? The determinants of German students' achievement in the PISA 2000 study. Discussion Paper No. 739 for the Institute for the Study of Labor (IZA). Retrieved from http://ftp.iza.org/dp739.pdf.

Flavell, J. H. (1979). Metacognition and cognition monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34,* 906-911.

Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills.* (pp. 35–60). New York: Academic Press.

Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In F. E. Weinert & R. H. Kluwe (Eds.), *Metacognition, Motivation and Understanding* (pp. 21-29). Hillside, New Jersey: Lawrence Erlbaum Associates.

Foong, P. Y. (1993). Development of a framework for analyzing mathematical problem solving behaviours. *Singapore Journal of Education, 13*(1), 61-75.

Frost, L. A., Hyde, J. S., & Fennema, E. (1994). Gender, mathematics performance, and mathematics-related attitudes and affect: A meta-analytic synthesis. International *Journal of Educational Research, 21*, 373-385.

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

Grimes, P. W., (2002). The overconfident principles of economics student: An examination of a metacognitive skill. *Journal of Economic Education. 33,* 15-30.

Hackett, G.;, & Betz, N. E. (1989). A self-efficacay of the mathematics self-efficacy/mathematics performance correspondence. *Journal for research in mathematics education, 20,* 261-273.

Hitch G. (1978). The role of short-term working memory in Mental Arithmetic. *Cognition Psychology 10, 302– 323.*Hyde, J. S., Feenema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139-155.

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *PNAS, 22*, 8801-8807.

Kaiser, A. P., Hester, P.P., & MC Duffie, A.S. (2001). Supporting communication in young children with developmental disabilities: Mental retardation and developmental. *Disabilities Research Reviews, 7,* 143-150.

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

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77,* 1121-1134.

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

Lichtenstein, S., & B. Fischhoff. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance. 20*, 159-183.

Logie R, Gilhooly KJ, & Wynn V. (1994). Counting on working memory in arithmetic problem solving. *Memory Cognition, 22, 395–410.*

Mandeville, G. K., & Liu, Q. D. (1997). The effect of teacher certification and task level on mathematics achievement. *Teaching and Teacher Education, 4*, 397-407.

Maqsud, M. (1998). Effects of metacognitive instruction on mathematics achievement and attitude towards mathematics of low mathematics achievers. *Educational Research, 2,* 237- 243.

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.

Marks, G. N. (2006). Are between- and within-school differences in student performance largely due to socio-economic background? Evidence from 30 countries. *Educational Research, 48,* 21-40.

Marton, F., Dall'Alba, G., & TSE, L.T. (1993). The paradox of the Chinese learner (Occasional Paper 93.1). *Educational Research and Development Unit*, Royal Melbourne Institute of Technology.

Melbhuish, E. C., Sylva K, Sammons P, Siraj-Blatchford I., Taggart B., Phan M. B., & Malin A., (2008). The early years. Preschool influences on mathematics achievement. *Science, 321*, 1161-1162.

Ming, M.C., & Zeng, X. H. (2008). Family and motivation effects on mathematics achievement: Analysis of students in 41 countries. *Learning and Instruction, 18,* 321-336.

Moss, E. (1990). Social Interaction and Metacognitive Development in Gifted Preschoolers. *Gifted Child Quarterly, 34*, 16-20.

Nist, S. L., Holschuh, S. J., & Sharman, S. L. (1995). *Making the grade in undergraduate biology courses: Factors that distinguish high and low achievers.* Mimeograph. Paper presented at the Annual Meeting of the American Educational Research Assoication in San Franciso, Calif.

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

Panaoura, A. (2007). The impact of recent metacognitive experiences on preservice teachers' self-representation in mathematics and its teaching. *CERME,5,* 329-338.

Pappas, S., Ginsburg, H.P., & Jiang, M. (2003). SES differences in young children's metacognition in the context of mathematical problem solving. *Cognitive Development. 18*, 431-450.

Perry, L. B., & McConney A. (2010). Does the SES of the School Matter? An Examination of Socioeconomic Status and Student Achievement Using PISA 2003. *Teachers College Record, 112(4)*: 1137-1162. Retrieved from http://www.tcrecord.org ID Number: 15662, Date Accessed: 6/1/2010 11:49:34 AM.

Pintrich, P.R., & De Groot E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82,* 33-50.

Post-Kammer, P., & Smith, P. L. (1985). Sex differences in career self-efficacy, consideration, and interest of eight and ninth graders. *Journal of Counseling Psychology*, 32, 63-81.

Princiotta, D, Flanagan, KD, & Germino Hausken, E. (2006). *Fifth Grade: Findings From the Fifth Grade Follow-up of the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99. NCES 2006-038.* Washington, DC: National Center for Education Statistics.

Reusser, K. (2000). Success and failure in school mathematics: effects of instruction and school environment. *European Child & Adolescent Psychiatry, 9,* 17-26.

Salomon, G., & Perkins, D. N. (1997). Individual and social aspects of learning. *Review of Research in Education, 23,* 1–24.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online, 8,* 23-74.

Schommer, M. (1990). The effects of belief about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498-504.

Shulruf, B., Hattie, J., & Dixon, R. (2007). Development of a new measurement tool for individualism and collectivism. *Journal of Psychoeducational Assessment, 25*, 385-401.

Siegle, D., & McCoach, D. B. (2007). Increasing student mathematics self-efficacy through teacher training. *Journal of Advanced Academics, 18,* 278–312.

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

Skaalvik, E. M., & Rankin, R. J. (1994). Gender differences in mathematics and verbal achievement, self-perception and motivation. *British Journal of Educational Psychology, 64,* 419-428.

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*, 75-109.

Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*, 51-79.

Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology, 2,* 344-355.

Stigler, J.W. (1984). ''Mental abacus.'' The effects of abacus training on Chinese children's mental calculation. *Cognitive Psychology, 16*, 145-176.

Van Keer, H., & Verhaeghe, J. P.(2005). Effects of explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy perceptions. *Journal of Experimental Education, 73*, 291-329.

Van Steenbrugge, H., Valcke, M., & Desoete, A. (2009). Mathematics learning difficulties in primary education: teachers' professional knowledge and the use of commercially available learning packages. *Educational Studies*, *36,* 59 – 71.

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.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Webster, B. J., & Fisher, D. L. (2000). Accounting for variation in science and mathematics achievement: A multilevel analysis of Australian data Third International Mathematics and Science Study (TIMSS). *School Effectiveness and School Improvement, 11,* 339-360.

Wertsch, J. V. (1985). Adult child interaction as a source of self-regulation in children. In R. Yussen (Ed.), *The growth of reflection in children* (pp. 69-97). New York: Academic Press.

Whitbeck, L. B., Simons, R. L., Conger, R. D., Wickrama, K. A. S., Ackley, K. A., & Elder, G. H. (1997). The effects of parents' working conditions and family economic hardship on parenting behaviors and children's self-efficacy. *Social Psychology Quarterly*, *60,* 291-303.

Wilson, K. M., & Swanson, H. L. (2001). Are mathematics disabilities due to a domain-general or a domain-specific working memory deficit? *Journal of Learning Disabilities, 34,* 237-248.

Young, D. J. (1998). Rural and urban differences in student achievement in science and mathematics: a multilevel analysis. *School effectiveness and school improvement, 4*, 386-418.

Chapter 5

The quadratic relationship between socioeconomic status and learning performance in China by multilevel analysis*

Abstract

The purpose of the present study is to explore the relationship between family socioeconomic status and mathematics performance on the base of a multi-level analysis involving a large sample of Chinese primary school students. A weak relationship is found between socioeconomic status and performance in the Chinese context. The relationship does not follow a linear, but a quadratic curve, implying that students from a disadvantaged family and higher socioeconomic background have a higher probability to attain higher mathematics scores. This can be explained on the basis of Chinese cultural beliefs about education, exams and social class mobility. Moreover, the aggregated socioeconomic status at the school level seems to moderate in the relation between individual SES and academic performance. This suggests that individuals from a disadvantaged family will benefit more from the school with a higher family socioeconomic status than students who are enrolled in schools with a lower and average family socioeconomic status.

1. Introduction

The relationship between socioeconomic status (SES) background and academic performance has received ample attention since the publication of the "Coleman Report" in 1966 (Coleman, Campbell, Hobson, McParland, Mood, Weinfeld and York, 1966). Supported by 150,000 students sample, Coleman Report argues that student's family socioeconomic status is much more important in predicting educational performance than are measured differences in school resources. A large body of empirical evidence is available about the relationship between SES and student performance in the context of the critical school subject mathematics (Sirin, 2005; White, 1982). Findings in developed countries reveal that students with a high family SES perform better than students with a lower SES (Lee, and Burkam, 2002; OECD,

---

* This chapter is based on the accepted paper of Zhao, N.N., Valcke, M., Desoete, A., & Verhaeghe, J. (in press). The quadratic relationship between socioeconomic status and learning performance in China by multilevel analysis: Implications for policies to foster education equity. *International Journal of Educational Development*.

2004; Wößmann, 2003). But, whether the relationship between SES and mathematics performance is different in other countries (developed or developing, various cultural value) is still an open question when we consider contradictory research results (Baker, Goesling and Letendre, 2002; Heyneman and Loxley, 1982, 1983). As shown in the Third International Mathematics and Science Study (TIMSS), the influence of family socioeconomic status on educational performance is complex depending on the economical development level of a region (Schiller, Khmelkov and Wang, 2002; Lockheed & Zhao, 1993). For example, it is argued that increasing economic situation in the region may change the role of families and the educational stratification and its effect on performance (Baker, Goesling & LeTendre, 2002). This brings us to the particular geographical context of the present study: China.

The question about the relationship between academic performance and students' SES is central in discussions about educational quality improvement in China as a developing country. As a country with 9,596,961 sq.km area, the differences between developmental levels and the distribution of wealth varies heavily between different provinces (Brenner, 2001; Feng, 2003). Since the establishment of the P. R. China in Year 1949, governmental policies have resulted in a cultural context characterized by a high intergenerational mobility and a rapid urbanization (Deng and Treiman, 1997; Tamura, Menton, Lush, Tsui & Cohen, 1997). However, the P. R. China is a huge continent, with initially a poor population in both rural and urban areas (Chen, 2005). After the cultural revolution, the government adopted new policies that resulted in the development of a stronger economical and cultural autonomy in local regions from the late 1970s. In order to fight educational inequity, particular policies were installed to allow students from poor areas to have access to better educational opportunities. From 1978, the "Gaokao" policy (China's National Matriculation Tests Policies, NMTP) was reinstituted to regard a student's achievement as the criterion for entrance to higher education (Agelasto & Adamson, 1998). Also, the increased possibilities for geographical mobility helped to fight unequal access to education. However, inequity between different provinces continues to exist because of differences in resources, transport conditions, etc. The Chinese government continues to make efforts for balancing the developmental levels of all regions and adopted this as a long term goal (Ministry of Education, MOE, 2001; Zeng, Deng, Yang, Zuo, Chu and Li, 2007). For example, considering the gap between educational opportunities in rural and urban areas, the government adopted the "Decision on Reforming and Developing Basic Education by State Council" (MOE, 2001; 2008). This resulted in closing down small primary schools with only 3 grades and with weak resources (called "jiaoxue dian") and integrating them into bigger schools (called "zhongxin xiao"). These continuing efforts to fight educational inequity make it very interesting to explore the relationship between SES and student performance in mainland China and to focus on the impact of different development levels within the same culture.

In the Chinese context, available studies about the relationship between SES and performance mainly focused on Hong Kong and Macau; regions with a rather high level in economic development (Liu and Lu, 2008; Park and Hannum, 2001). The limited available Chinese empirical studies in mainland China reveal that – when controlled for other factors - a higher family background has a positive but lesser impact on performance (Liu and Lu, 2008; Xue, 2007; Zuo,1994). In their multiple regression analysis, Liu and Lu (2008) found that SES only explains .8% of student performance.”. Also, some studies point in particular at the impact of the educational level of the father having a significant positive relationship with mathematics performance (Park and Hannum, 2001). Other studies point at the decisive impact of the educational level of the mother (Park and Hannum, 2001). However, there are limitations to previous studies set in mainland China: samples were rather small or did not represent a variety of developmental regions/provinces; the SES indexes remained restricted, and the interaction between individual SES and school average SES was hardly considered.

In the present paper, we center on the mainland Chinese context to set up a comprehensive empirical study, while trying to further develop educational theory development. This brings us to the research aims of the present study: (a) to construct a comprehensive SES index based on input from previous studies; (b) to explore the general relationship between SES and mathematics in P. R. China, considering different developmental levels of the region; and (c) to analyze the extent to which aggregated SES at the school level influences student mathematics performance, regardless of students different individual SES levels. The study builds on data gathered from 10,959 Chinese pupils enrolled in schools that are geographically located in five Chinese provinces with different economic developmental levels. The data used in the present study were gathered during a project, funded by the BOF Project ‘Mathematics Education in China’ of Ghent University (Belgium). We collected the data from 20 schools in 5 Chinese regions in close collaboration with researchers from the Educational Bureau of Beijing Normal University and of South China Normal University.

2. Theoretical background

2.1 Measuring socioeconomic status

The way to define and measure socioeconomic status (SES) has changed a lot during recent years (Entwisle and Astone, 1994; McLoyd, 1998). As can be derived from Table 1, a variety of variables have been used in the literature to develop SES indexes. In the 1980s, SES indexes stressed family income, father’s educational level, mother’s educational level, and father’s occupational status or occupation type (White, 1982). In latter studies, additional variables were added; such as home resources (Sirin, 2005), home atmosphere or context, number of books in the household, and other resources related to the learning (Caldas and Bankston, 1997; OECD, 2004).

Table 1
Overview of the variables constituting recent SES indexes

Authors	Individual SES						School SES	Neighourhood SES	Home atmosphere
	Parental educatio	Parental occupati	Parental income	Home resource	Number of books	Reduced price			
White (1982)	√	√	√	√					
Duncan, Brooks-Gunn and Klebanov (1994)			√						
Brooks-Gunn, Duncan, and Aber (1997)							√	√	
Caldas, S. J. and Bankston, C. L. (1997).	√	√	√	√	√	√			√
Baer, J (1999)	√	√							
Louis, V. V. and Zhao, S. Y. (2002)	√		√						
Sirin, S. R. (2005)	√	√	√	√	√	√	√	√	
OECD (2004) for PISA report	√	√		√					√
Olson, J.F., Martin, M.O., and Mullis, I.V.S. (2007) for TIMSS report	√			√	√				

Since the 1990s, next to the individual students' SES, also aggregated SES measures were developed to consider the impact at the level of the school and related contexts. This resulted in additional indexes, such as school SES-level and neighborhood SES-level (Brooks-Gunn, Duncan, and Aber, 1997; Sirin, 2005). The adoption of aggregated indexes is not generally accepted. Some researchers report that when family SES is controlled for, neighborhood SES only plays a minor role (Duncan, Boisjoly, and Harris, 2001; Sanbonmatsu, Kling, Duncan and Brooks-Gunn, 2006).

Although available SES indexes seem to vary, most SES constructs seem to incorporate the following variables: parental educational level, parental occupation and home resources or wealth. Consequently, we also adopt this approach in the present study.

2.2. Varying impact of SES on mathematics performance

2.2.1 Family SES and academic performance

In the literature, family SES is consistently found to be a single strong predictor of educational outcomes (Fransoo, Ward, Wilson, Brownell, and Roos, 2005). To explain this, researchers argue that parents from families with a low SES are less involved in their children's schooling and give less support to the children as compared to parents from families with a higher SES, resulting in low academic achievement (Ho and Willms, 1996; Jeynes, 2003; Silinskas, Leppanen, Aunola, Parrila and Nurmi, 2010). Recently, the meta-analysis of both White (1982) and Sirin (2005) reveals that the direct relation between socioeconomic status and performance might be less strong as supposed. White's meta-analysis claimed that his meta-analysis reveals that the average correlation is somewhat .29 (studies set up between 1918 and 1975), while Sirin's meta-analysis claims a correlation value of .34 (studies set up between 1990 and 2000). They focus in their analysis on large within-group differences when studying the relationship between SES and performance. They also stress that the impact of family SES on performance differs largely depending on the economic development level of the region or country. Economical differences tend to result in different regional and school educational policies.

2.2.2 SES and performance: the moderating effect of school aggregated SES variables

In this context, little is known about how school level variables in a specific region moderate between family SES and mathematics performance (Peng and Hall, 1995). Previous studies show that higher levels of an aggregated school SES are related to an increase in student performance and in students with a different level of family SES (Perry and McConney, 2010). Researchers argue that the acquisition of values and goal-orientations within the schools and the combined effect of students' attributes contribute to the changes in performance of all students (Alexander, Fennessey, McDill and D'Amico, 1979; Haller and Woelfel, 1972). In addition, other researchers stress that specific classroom variables mediate between SES and performance (Aypay, Erdogan,and Ma, 2007). The latter explanation is adopted to refer to poorer student outcomes in rural schools as compared to schools in urban or suburban regions (Webster and Fisher, 2000). These authors explain that the geographical isolation and lower economical development restrict access to learning materials and other educational resources. Additionally, the developmental level of the region where schools are located can result in more or less advanced governmental educational policies (Marks, 2006). At present, hardly empirical evidence is available to test the interaction between school SES and learners' SES in the Chinese context.

3. Method

3.1 Sampling

In this study, mathematics performance data were obtained from 10,959 students, enrolled in grade one to grade six, from twenty schools. A multistage stratification sampling procedure was followed. These twenty schools are located in five Chinese regions reflecting different development levels; and are located in a rural or urban setting (See Table 2). Total school enrolment ranged from 318 to 897 students (*M*=547.95, *SD*=140.19). Sampling strata were based on the location of the school in a specific region. Within a school, grade level classes – in case parallel classes were present - were randomly chosen by the researchers and after negotiations with the school principals. In total, 51.88% of the learners is male, 53.14% are enrolled in urban schools. Five economical development levels are distinguished, based on data about the regional gross domestic product (GDP). As a result, distribution of pupils in these regions is as follows: 27.58% in level 1, 15.07% in level 2, 19.88% in level 3, 16.22% in level 4 and 21.24% in level 5 respectively.

Table 2.
Sample characteristics (N=10,959).

Grade[b]		Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Total	Percent
Gender	Female	903	883	872	854	845	917	5274	48.12%
	Male	954	972	933	977	906	943	5685	51.88%
Region[a]	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%

Note: Since the data have been collected from students and schools in urban and rural areas within five regions with different GDP levels, the indicators of urban/rural, and GDP were added from the start to the model. However, these variables at this additional level did not contribute in explaining variance in the model (Authors, in press).

Note: The data was collected from pupils enrolled in the six primary school grades. It can be hypothesized that the influence of SES might be different according to the grade (Chiu, 2010). Therefore, the interaction between grade and family SES was added to the model. This did not result in a significant improvement of the model. As a result, grade is not considered as a critical variable in the subsequent analyses.

3.2 Variables

3.2.1 Dependent variables

Mathematics performance level. The items constituting our mathematics performance test were taken from a previous study in which a new mathematics test was developed, aligned with the most recent 2001 Chinese mathematics curriculum. The test, covering the mathematics curriculum from grade 1 to grade 6, was calibrated on the base of item response theory. The design, development and calibration, involving 10,959 primary school learners, was carried out in May till November 2008. All the items and cases were calibrated on a continuum scale ranging from grade one to six. Reliability of the scales (Cronbach's α) was reported to be high: for grade one to grade six respectively .94, .96, .95, .94, .94 and .93. Reported means are: -1.24, -.89, .05, .18, .69, .83. Mathematics standardized performance scores range from -5.30 to 3.30 (M = .57, SE = .26).

3.2.2 Independent variables

3.2.2.1 Individual learner's level SES

Based on our literature review, eight items were developed to collect data about SES indicators: father educational level, mother educational level, father's job, mother's job, ownership of a TV, refrigerator, washing machine and computer. In order to create a generic SES index, factor analysis was carried out including the eight SES-related items by programme Mplus 5.1. Information about the parents' background was obtained from the schools with the support of the Educational Bureau. Information about wealth indicators was obtained via a questionnaire filled out by the parents.

In a first step, an exploratory factor analysis (EFA) was carried out by applying the WLSM method, able to deal with categorical data (Mplus5.1). The results suggest a two-factor structure in the SES variable. But, weak factor loadings were observed for both father and mother's educational level; moreover these SES items loaded on both factors. They were therefore excluded from the further analysis. In a second step, a confirmatory factor analysis was carried out building on the two-factor solution. This two-factor structure resulted in satisfactory goodness-of-fit indexes (x^2=158.45, *df*=8, *p*-value<.00, CFI=.98, TLI=.97, RMSEA=.04). A first factor grouped SES variables focusing on parents' occupation status; a second factor groups SES variables in relation to family wealth.

Finally, according to the results of the factor analysis, a generic SES index was calculated, combining the items of the two factor structure: (1) learners' SES based on the parents' job; (2) learners' SES based on the family's wealth. And two independent variables: (3) Parents' educational level, including father educational level and mother educational level.

Parents' educational level: Prior research about student achievement in primary schools has shown that father's (FEL) and mother's educational level (MEL) influence learners' achievement (Alwin and Thornton 1984). Parental educational level was coded as: no schooling experience (1), primary school graduate (2), junior school graduate (3), senior school graduate (4), Pre-high school (5), high school graduate (6), postgraduate or higher (7). This categorization fits educational levels resulting from the Chinese educational system.

***Family SES _Job* (FSES_J)**: With support of their parents or teachers, learners reported their parents' job. These answers were coded into 26 categories (26 = highest level ranking) that reflect the classification scheme of Li (2005a; 2005b; See Appendix 1). Her research was based on the Chinese context and provided a national valid estimate of job levels by using the calculation adopted in Lin's studies (Lin & Xie, 1989; Lin & Ye, 1997). This is a higher number of levels as compared to the study of Xu (2000) study. Compared to Lu's (2000) studies with 10 levels job, Li's study provides more detailed information about the jobs and as such fits better with our study. The Spearman correlation between the ranking of Li (2005) and Lu (2001) is .97 ($p<.00$), reflecting a high correlation between existing ranking/coding systems.

Family SES_J_Record: In view of some analyses (see section 4.3), the job levels have been recoded into three categories, building on the distribution in job levels: lower 25%, middle 50% and higher 25% with the codes of *FSES_J_low job group, FSES_J_middle job group and FSES_J_high job group, respectively*. This will make it possible to study the interaction between individual level SES variables and aggregated school level SES variables.

Family SES_Wealth **(FSES_W)**: This variable builds on the answers to four questions about wealth indicating property: ownership of a television, a refrigerator, a washing machine, and a computer.

Family SES_Wealth_Record (FSES_W_R): In view of analyses about interaction effects (see section 4.3), this variable was also recoded into three categories, building on the distribution in ownership of the four different wealth indicators: lower 25%, middle 50% and higher 25%) with the codes of *FSES_low wealth group, FSES_middle wealth group and FSES_high wealth group, respectively*.

3.2.2.2 School level SES

Two aggregated SES indexes were calculated at the school level: (1) the average parent's socioeconomic status of the learners attending this school, based on the father's and mother's occupational level (*SCFSES_J*) and (2) the average level of wealth of the children's family in the school (*SCFSES_W*).

3.3 Data analysis

First, as explained above, both an exploratory (EFA) and confirmatory factor analysis (CFA) were carried out to develop fitting SES indexes (Mplus by WLSM methods).

Second, a multilevel analysis was applied to study the impact of variables at the school, class and student level on mathematics performance. Multilevel linear modeling overcomes major shortcomings of single level regression analysis. Firstly, multilevel regression builds on iterative generalized least squares (IGLS) techniques to estimate the direct and cross-level effects for the hierarchical data. This is in sharp contrast to Ordinary Least Squares (OLS) regression techniques that overestimate the contextual and cross-level variables, and consequently are prone to Type I errors (Aitkin, Anderson, and Hinde, 1981; Rowe, 1992). Secondly, multilevel regression allows us to estimate the fixed part and random effects to explain the variance in a model. Thirdly, multilevel analysis helps – in the present context - to study the influence of SES on mathematics performance considering school SES and the regional economical level.

4. Results

4.1 Descriptive analysis

In Table 3, we summarize the bivariate correlation coefficients reflecting the association between all variables in the present study (Kendall's tau). The correlations between mathematics performance and parents' occupational level (FSES_J, SCFSES_J) and family's wealth (FSES_W, SCFSES_W) underscore the decision to include these variables in relation to the SES index. We observe strong correlations between the other variables in the study: SCFSES_J and SCSES_W (r = -.80), FEL and MEL (r = .63), FSES_J and FSES_W (r = -.59). Though the large correlation variables raise concerns about multi-collinearity, we decided to include these variables in the regression model. But, by entering and removing the variables one by one, we are nevertheless able to control the interaction between predictors.

Table 3
Zero-order correlations between variables in the study (N=10959)

	2	3	4	5	6	7	Mean (σ)
1.MATH	.041**	.032**	.164**	.157**	.097**	.084**	.567 (1.113)
2.FEL		.702**	.373**	.353**	.417**	.324**	4.094 (1.394)
3.MEL			.370**	.353**	.387**	.323**	3.963 (1.417)
4.SCFSES_J				.952**	.486**	.492**	-.077 (1.375)
5.SCFSES_W					.462**	.515**	-.016 (.165)
6.FSES_J						.624**	.070 (2.828)
7.FSES_W							-.014 (.319)

Note: ** $p < .001$

MATH refers to the mathematics performance; FEL and MEL refer to father and mother education level; SCFSES_J refers to the aggregated indicator of parents' occupation in school level; SCFSES_W refers to the aggregated indicator of parents' wealth in school level; FSES_J refers to the indicator of parents' occupation of family SES; FSES_W refers to the indicator of parents' wealth of family SES.

4.2 Multilevel analysis of the relationship between SES variables and mathematics performance

4.2.1 The weak relationship between SES variables and mathematics performance

In this section, we focus on the model 0 to model 6 in table 4. In table 4, model 0 points at the null model without any predictors in the multilevel analysis. In total, 18.55% of the overall variance is explained by the school level, 39.52% of variance is explained at classroom level and 41.94% of the variance is explained at the individual learner level. The analysis results also reveal that the variance at these three levels is different from zero; implying that a three level multilevel analysis should be applied. In a first step, father's educational level and mother' education level were entered into the model. Compared to the null model, there is no significant improvement between model 1 and 2 after adding father's and mother's educational level (See Table 4; x^2=.225, *df*=1, *p*=.635; x^2=.005, *df*=1, *p*=.944).

Secondly, the indicators of parents' job level (FSES_J) and indicators of wealth (FSES_W) were entered in a subsequent model. As shown in table 4, there is a significant improvement of the model 3 when the linear variable of indicator of parents' job (FSES_J) is added to the model. A second-order polynomial regression of parents' job (FSES_J) on mathematics performance reveals a highly significant U-shape relationship (See model 4; x^2=12.936, *df*=1, *p*<.001).

Similarly, when a linear and 2-order polynomial function of indicator of family wealth (FSES_W) is conducted, the model of 5 and 6 improves significantly. Considering the polynomial variable family SES of parents' job, the linear variable of family wealth (FSES_W) significantly improves the model in model 5. When the two-order polynomial of family wealth (FSES_W) is added to model 6, this results also in a significant improvement from model 5, although the coefficient of the linear variable of family wealth (FSES_W) is now no longer significantly different from zero.

Though - consistently - a significant improvement in consecutive models can be observed, these models only account for 0.41% of the variance in mathematics performance at the individual learner level (.486 vs. .484) (compare model 0 to model 6 in Table 4). This implies that the SES variables under study are not strong predictors of mathematics performance in primary school after controlling for school level variables.

4.2.2 The U-shaped relationship between SES variables and mathematics performance

Another interesting finding from our study is the U-shaped relationship between SES variables and performance in the primary school. The relationship between mathematics performance and SES was not studied as a linear function, but as a quadratic one.

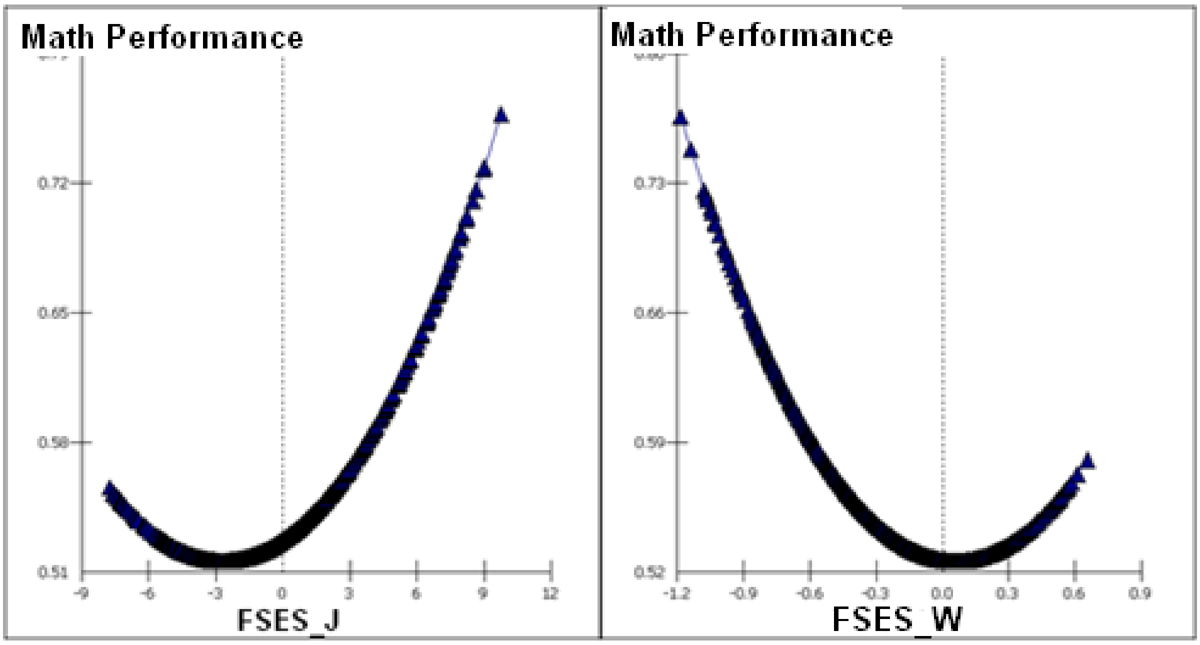


Figure 1
Prediction of mathematics performance as a function of two SES variables.

Figure 1 shows that the relationship between SES and mathematics performance is U-shaped. We observe that children with a higher mathematics performance can come from a higher family SES level or lower family SES level while the children with a very weak mathematics performance can also come from a middle SES family. Previous research in the Chinese setting stressed a linear relationship between family SES and mathematics performance (Chiu and Zeng, 2008). The present results are in conflict with these findings. However, the reason might be that the previous studies didn't explore polynomial relationships. As Figure 1 shows, the results with the Chinese students reveal a different pattern: students with higher and lower family SES achieve higher as compared to students from a middle family SES background.

It is interesting to develop a better understanding of what these SES levels represent. We centre first on the 2920 students who are expected to belong to the lowest 25% in math performance considering their parents' occupation. Their parents' job level (FSES_J) ranges from -3.729 to -1.656. About 45.34% of these parents are both workers (rank 19 in Appendix 1) while 28.42% of these parents are both peasants (rank 21). The latter families represent more than 80% of the families in China (Li, 2005a). The present results reiterate an urgent problem for the government, since the children of this large group face a risk to attain a lower math performance level.

Second, we focus on the performance of students from the families with the 25% lowest SES considering the parents' occupation (worker, peasant, bodily labor worker, servant and unemployed). As we can derive from Figure 1, this is a rather small sample in our study (n = 472). A surprising result is found. These students achieve higher than students from a family with an average family job level. Nevertheless, these students are expected to attain means of .53 in mathematics performance which is still very weak compared to the higher achievers.

When we focus on the other SES component of wealth (FSES_W), we hardly distinguish varying levels in ownership of wealth indicators. Considering students with the lowest 25% performance, around 95% of these students' family own a TV, refrigerator, washing machine and computer. For students attaining a score within the group of the highest 25% in mathematics performance, also these families own only one or two of these typical wealth equipments

Students from families with an higher level of ownership of wealth indicators, have a larger probability to attain lower mathematics scores as compared to other students. Students with a relatively high family SES have fairly good family situation. It is conjectured that families with a higher SES background provide sufficient support for learners; such as a richer learning environment or a higher level of parent involvement in school related activities. Although students from a disadvantaged family SES background might receive less learning support from their parents, they nevertheless struggle to achieve better in primary school in order to compensate for their disadvantaged family situation. While attaining a higher performance level, these students can counter their disadvantaged family background and potentially attain a higher status in society. This can motivate them to learn harder and achieve higher from the start of the primary school. However, the short tail in figure 1 also reveals that the final attainment level of learners at the lower side of the FSES_J axis will never be as high as the performance of learners belonging at the upper side of the FSES_J axis. This implies that the students with lowest family SES do not attain the same performance level as students with the highest family SES.

4.3 SES and mathematics performance: the moderator effect of school level aggregated SES indexes

4.3.1 The stronger effect of School level of FSES_J on mathematics performance

In this section, we focus on the model 7 to model 9 in Table 4 that centers on the interaction between school effects and family effects on mathematics performance. When comparing model 6 with model 0 in table 4, approximately 0.77% of the variance at the class level and about 1.76% variance at the school level is explained by SES related variables. Family SES seems to play a different role in these two models.

This underpins the relevance to add an aggregated family SES variable to the models. In the models 7 to 9 in table 4, we added the mean family SES level of pupils in the same school to the equation.

The average SES at the school level can impact student performance in a variety of ways. For instance, school administrators can make different efforts to improve conditions for learners with a disadvantaged SES background. Or, learners affect each other by bringing a richer cultural capital into the classroom, because of the language they speak, because the experiences they share, etc. In the multilevel analysis, both school level SES variables and learner level SES indicators will therefore be considered as predictors in the model. As can be derived from Table 4 (model 7), adding the school level SES variable - based on wealth indicators - does not result in a significant model improvement (x^2=3.369, *df*=1, *p*=.066). But entering the school aggregated SES variable - based on the students' parents' job - does significantly improve the model (x^2=4.301, *df*=1, *p=.038*). In model 8, about 18.06% variance at the school level is explained by the average SES of parents' job at the school level. More concretely, in model 8, mathematics performance increases by .160 units when the FSES_J at school level increases with one unit. To summarize, when means of SES at school level are entered into the models, about 22.02% of the total variance in mathematics performance can be explained (.227 vs. .177 see model 8 in Table 4) at the school level.

Table 4. Multilevel analysis results in relation to the subsequent models explaining mathematics performance

	Model 0	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
Fix part										
Intercept	.555 (.116)	.555 (.116)	.555 (.116)	.555 (.115)	.537 (.115)	.538 (.115)	.527 (.115)	.548 (.107)	.544 (.105)	.530 (.105)
FEL		.003 (.006)								
MEL			-.000 (.006)							
FSES_J^1				.010 (.003)	.007 (.003)	.012 (.003)	.008 (.004)			
FSES_J^2					.002 (.001)	.002 (.001)	.002 (.001)			
FSES_W^1						-.086 (.029)	-.023 (.042)			
FSES_W^2							.154 (.075)			
SCFSES_W								1.215 (.636)		
SCFSES_J									.160 (.073)	.161 (.074)
SCFSES_J * *FSES_high job group*										-.043 (.017)
SCFSES_J * *FSES_low job group*										.022 (.019)
Random part										
$\sigma^2_{school.\ 0}$	.227 (.085)	.227 (.085)	.227 (.085)	.221 (.084)	.220 (.083)	.223 (.084)	.223 (.083)	.186 (.072)	.176 (.069)	.177 (.070)
$\sigma^2_{class.0}$	.517 (.049)	.517 (.049)	.517 (.049)	.517 (.050)	.514 (.049)	.514 (.049)	.513 (.049)	.516 (.049)	.517 (.049)	.515 (.049)
$\sigma^2_{learner.0}$	.486 (.007)	.486 (.007)	.486 (.007)	.485 (.007)	.485 (.007)	.484 (.007)	.484 (.007)	.486 (.007)	.486 (.007)	.485 (.007)
-2LL	24186.831	24186.606	24186.826	24174.951	24162.015	24153.492	24149.303	24183.462	24182.530	24170.357
x^2		.225	.005	11.880	12.936	8.523	4.189	3.369	4.301	12.173
Df		1	1	1	1	1	1	1	1	2
p-value		.635	.944	<.001	<.001	.004	.041	0.066	.038	.002
Reference		Model 0	Model 0	Model 0	Model 3	Model 4	Model 5	Model 0	Model 0	Model 8

Note: FEL and MEL refer to father and mother education level; FSES_J^1 refers to the indicator of parents' occupation of family SES, FSES_J^2 refers to the square of the FSES_J^1;

FSES_W^1 refers to the indicator of parents' wealth of family SES, FSES_W^2 refers to the square of the FSES_W^1;

SCFSES_W refers to the aggregated indicator of parents' wealth in school level; SCFSES_J refers to the aggregated indicator of parents' occupation in school level;

4.3.2 The moderator effect of school SES on mathematics performance

In this next step – while focusing on parents' job level - we consider the interaction between school level SES and learner level SES. As explained earlier, for this purpose the SES indicator based on job level was recoded into three job levels: 87.78% of the students in the FSES_J_low job group have parents with a job being lower than worker (ranking 19); and, 83.67% of the students in the FSES_J_high job group, have a father with a job higher than a less professional experts (ranking 13, see Appendix 1) and a mother with a job ranking from 4 to 26 (See Appendix 1).

In model 9, the interaction between school aggregated SES and the recoded learner level variable is added. This results in a further significant improvement in model fit (x^2=12.173, *df*=2, *p*=.002). As table 4 reveals, compared with learners from an middle SES family background, the interaction coefficient for learners from low SES family is -.043, being significantly different from zero in model 9. The interaction between school level SES and learner level SES based on parents' occupation can also be represented in a graphical way.

Figure 2 shows how the mathematics performance of learners with a different level of SES, based on their parents job level (recoded in three categories), varies according to the aggregated SES variable at school level.

The vertical axis represents a student's math performance while the horizontal axis refers to a schools' average SES. We focus in this figure on different students that study in a school with the same average school SES. These three students differ in their family SES (L, M, H representing a lower, middle or high family SES). The results are clear. Although the three students study in a school with the same average family SES index, the student L with a a lower family SES attains a lower mathematics performance as compared to student M with a middle high family SES. The later attains a lower mathematics performance as compared to student H. The reverse is true when we consider the results from students in a school with an average SES score that is lower than .00. Students with a lower family SES outperform both students with a middle and high family SES.

The differences in the slopes of learners with a low, middle or high SES, based on their parents' job level, show the moderating effect of schools on individual learner's family SES levels. The slope of the disadvantaged group with lower individual family SES is less steep as compared to the middle group and the advantaged group. This implies that the moderating effect is stronger for learners with a higher individual SES.

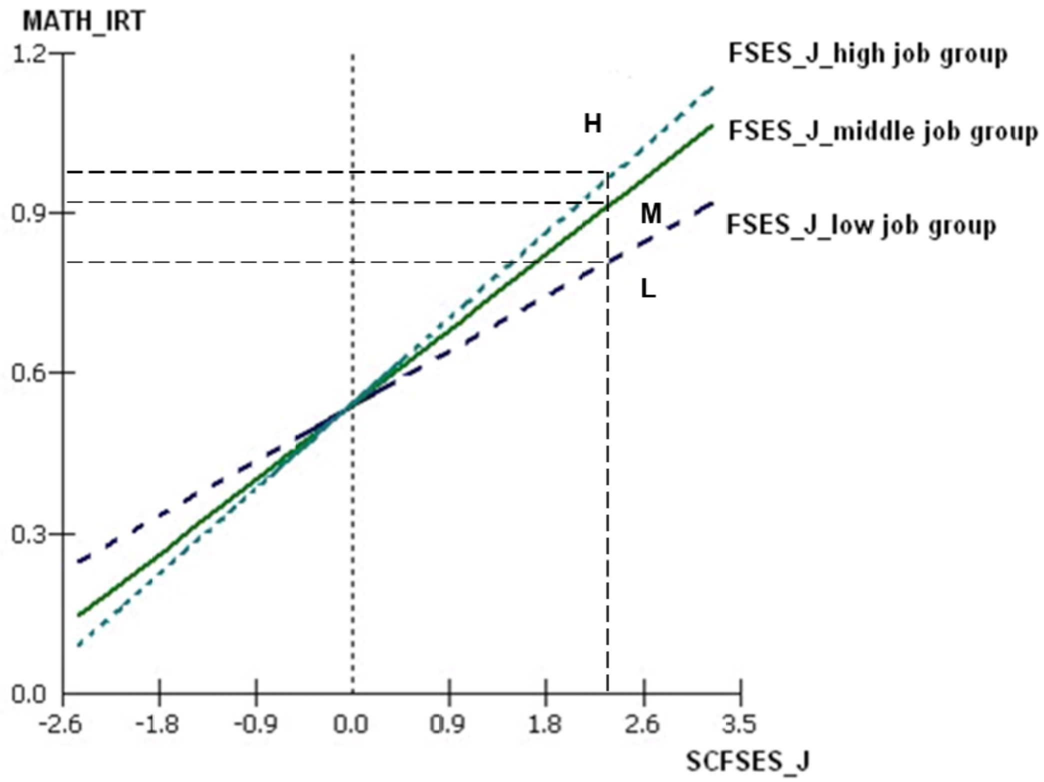


Figure 2.
Prediction of school SES on mathematics performance in primary school.
Note: MATH refers to the mathematics performance; SCFSES_J refers to the ; FSES_J_low job group refers to the group of students whose indicator of parents' job ranks at the lower 0-25%; FSES_J_ middle job group refers to the group of students whose indicator of parents' job ranks at the medium 25-75%; FSES_J_ high job group refers to the group of students whose indicator of parents' job ranks at the higher 75-100%.

However, figure 2 only shows an ideal situation. In reality, - since the school level aggregated SES is based on the means of the learners' family SES - schools with a lower SES will hold more students with lower family SES and vice verse. This implies that the analysis should be further refined to cater for the inequal representation of the different SES levels. This results in a revised version of the depiction as represented in Figure 3. It is clear that the proportions of learners with a specific family SES group level are different in schools.

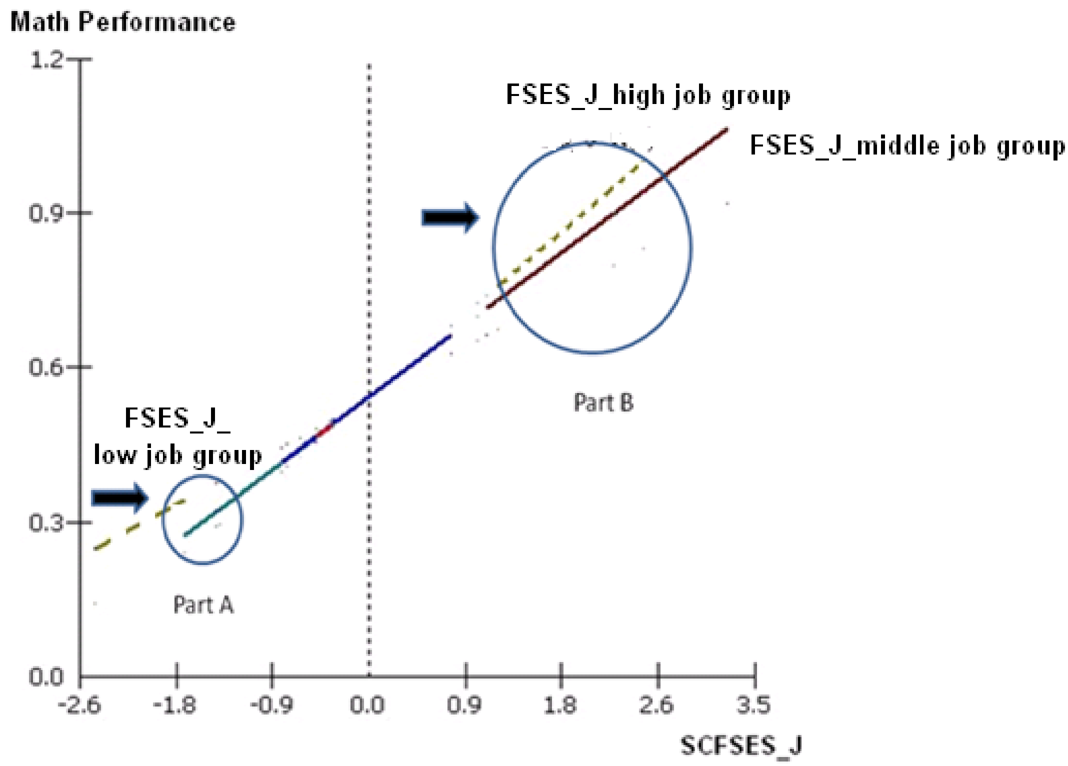


Figure 3.
Prediction of school SES on mathematics performance in primary school.
Note: MATH refers to the mathematics performance; SCFSES_J refers to the ; FSES_J_low job group refers to the group of students whose indicator of parents' job ranks at the lower 0-25%; FSES_J_ middle job group refers to the group of students whose indicator of parents' job ranks at the medium 25-75%; FSES_J_ high job group refers to the group of students whose indicator of parents' job ranks at the higher 75-100%.

In Figure 3, the slopes of the different groups of learners – considering their parents' job level – seem to partially overlap. When students with a lower SES family and students with an average SES family are enrolled at a school with a comparable aggregated SES, learners with a lower SES will attain higher mathematics score as compared to the learners with an average SES level (See Part A in Figure 3). In contrast, when learners with a higher SES family and learners with an average SES family enter the same school, learners with a higher SES family will attain higher mathematics scores than the learners with a average family (Part B in Figure 3). The higher the school aggregated SES, the higher performance of learners with a higher family SES. In general, these results replicate the quadratic relationship between SES and mathematics performance as discussed before.

5. Discussion and conclusions

5.1 Does the U-shape relationship between SES and mathematics performance result from governmental efforts or is it the artifact of cultural-historical variables? Implications for policies focusing on disadvantaged individuals

Building on the present research results, it is interesting to observe that SES of individual learners is not a strong predictor for mathematics performance in primary school. This is in line with other studies, set up in the Chinese context (Liu and Lu, 2008). But, this is clearly in contradiction to the findings in international studies (Fransoo, Ward, Wilson, Brownell, and Roos, 2005; Huang, 2010). This means that learners with a lower and higher SES achieve better as compared to learners with a middle SES level. It is to be stressed that the latter group especially comprises learners with parents who are workers and peasants.

A variety of rationales can be presented to explain these specific findings within the Chinese context. First, we can refer to Confucian cultural values that play a role. A basic value embedded in Confucian culture encourages children to learn hard and work hard in order to attain a better position in society. This can result in generational class mobility, and builds on a – hundred years old – tradition that students are selected on the base of their level of academic performance and not on their family background, also referred to as “KeJu” (Entrance Examination for higher education). Whatever a student’s family background, students will get an opportunity at a higher occupation/job pending a high performance in examinations.

In contemporary society, this situation is still clearly observed in the time and effort spent for e.g., National Matriculation Tests Policies (“Gaokao”). Sociological studies reveal that in China there is an unusually high degree of generational and occupational mobility, and “openness” of the society (Blau and Ruan, 1990; Kracke, 1947; Parish, 1981; Wu, 2007). After 1949, the government carried out equity promotion policies for farmers and workers to break the barriers in generation and class mobility These policies decreased the reproduction of the generation-locked occupational levels (Deng and Treiman, 1997; Lin and Bian, 1991). This openness in mobility encourages students with a disadvantaged family background to achieve better in schools in order to attain a higher occupational level. This particular motivational impact on learning has also been reported by other studies; e.g., researchers report that the achievement motivation of Chinese students is higher as compared to the motivation of Western students (Biggs, 1997; Ginsburg, and Bronstein, 1993). This cultural value and the subsequent adoption of compensatory policies seem to promote equity in society. Thus, the Chinese students seem to value schooling; though there are variation between the different type of students (Maslak, Kim & McLoughlin, 2010). Anyhow, while claiming that cultural-historical variables play a role in relation to the U-shape in our results, we also have to recognize that Chinese governmental policies fostered a relative openness in society that maintains this cultural spirit. For example, the "Gaokao" examination policy in relation to the the secondary education exams, continues to value learning performance of students with whatever family background (Li, 2009).

Although the cultural values and political policies can have a positive impact, it remains nevertheless clear that disadvantaged students still run a higher risk to encountering attainment difficulties. As we can derive from Figure 1, a group of students performing well are from families where both parents only have a low bodily labor job. On the one hand, without the academic support of the school and the parents, it is difficult to assure that these students continue to perform well during further education, such as middle schools, secondary schools and university. They might nevertheless meet learning difficulties or problems during a further phase in their school career. On the other hand, students who succeed to graduate from the university might also meet some problems in their "Quan Mian Fazhan" (Education should be concerned with the full development of the students, such as intellectual, moral, physical, aesthetics and labour development, not only develop intellectual dimension and ignore the others) or "Gao Fen Di Neng" (higher performance in school but low ability to live in society). Academic performance is only one part of being prepared for future life. Governmental policies should pay more attention to these students at risk and provide continuous support. This is reflected in the "Planning of Mid-Long Term Education Reform and Development Program (2010-2020)" (CPC & State Council, 2010). The government highlights educational equity as a key principle and it promises that by 2020 the quality of compulsory education will be the same in all regions, ensuring that all school-aged children and adolescents have equal access to high quality education. Also, no child shall be allowed to drop out due to family related financial difficulties.

5.2 The impact of the school background and family background: Implications for school development policies

The present study observed a very interesting moderating effect between school aggregated SES and individual family SES. But what level is predominant in this setting?T This is difficult to answer, since the interaction seems to be complex. Compared to students with an average family SES background (reference group), students with a disadvantaged family SES background benefit more from their school setting as compared to students with an average family SES background. In a comparable way, students with a higher family SES background also benefit more from their school as compared to students with an average family SES background. This finding is in line with the results of previous studies claiming that school-level variables could account for an important part of the students' achievement in primary schools (McEwan & Trowbridge, 2007). About 22.03% (.227 Vs. .177) of the variance at the school level can be explained by the school mean SES. This finding reconfirms that students with a disadvantaged family SES background can and should be supported with rich school resources and by their school peers. This is expected to result in a positive school climate that fosters learner motivation.

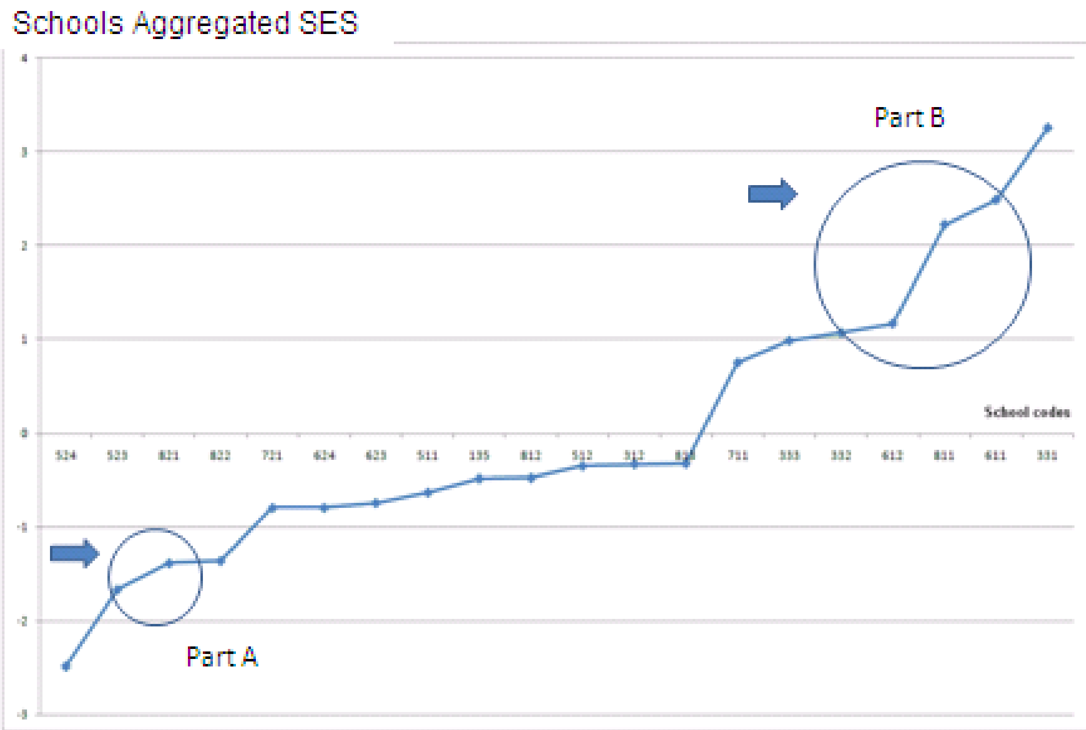


Figure 4.
Schools Aggregated SES scores.

How does the school play such an important moderator role? Building on a case, we can illustrate the above by commenting on mathematics performance of learners from school 524 where a strong improvement in mathematics performance is observed in disadvantaged students. School 524 has developed in the aftermath of the "Decision on Reforming and Developing Basic Education by State Council" (MOE, 2001; 2008). In order to narrow the developmental gap between learners from rural and urban areas, the Ministry of Education decided to reconstruct district schools (in Chinese called "Ce Dian Bing Xiao", that stops small primary schools and creates larger primary schools having more resources). In earlier days, in rural areas, small schools were set up that catered for rural students and only organized education for grade 1 to 3 (called

"Jiaoxue Dian"). Building on the reform mentioned above, rural schools were redevelop into large rural schools. School 524 is an example of this endeavor. The Chinese government additionally invested in these schools. Although this reform also resulted in some problems (e.g., a raise in the fee for living in a larger school), these new schools have resulted in the improvement of academic performance of students with a disadvantaged SES family background (See Liu, Zhang, Luo, Rozelle & Loyalka, 2010). Cases reflecting the impact of a higher average SES school level, are found in schools that have a stronger impact on performance of students with a higher individual SES background. Typical examples of these schools are school 333, 332, 612, 811 and 611. These schools have a longstanding history and are mostly positioned in an urban setting. The process of urbanization in China has accelerated during recent years. This resulted in a massive transfer of labor forces from rural to urban contexts; additionally resulting in an improvement of family conditions. But, this does also cause a wider heterogeneity within schools in urban settings. In the present study we did not control for this within-country migration process and did therefore not ask whether students were from the local area or had migrated with their parents.

5.3 Conclusions, limitations and directions for future research

The present attempt to explore the relationship between the family socioeconomic status and mathematics performance resulted in particular results about the Chinese setting. Although studying the interrelation between SES and academic performance in primary schools is not new, the results of the present study differ in a number of ways from the results of available research. First, the quadratic relationship between SES and mathematics performance is a particular finding that can partly be explained by particulars of the Chinese culture and educational policies. Moreover, our study explored the additional impact of school aggregated SES variables.

The present study also reflects some limitations. Firstly, the analysis approach was correlational in nature and thus cannot ground assumptions about causal directions between SES and mathematics performance. Also, the moderator impact of the school context cannot be explored in detail considering the correlation between school context and student family SES level. Another limitation is that the study did not focus on variables such as beliefs, attitudes or other motivational variables that are also mediators between family SES and mathematics performance. In the future, it is therefore interesting to study additional variables and to enter them in more complex multilevel models to explain mathematics performance.

Future research could also center more in detail on particular subgroups in the current large sample of learner: top performers versus weak performers in different school settings and how SES variables are impacted by contextual variables. The results bear clear policy implications in view of supporting students with disadvantaged family background. Firstly, we observe a complex combination of the ideological control approach resulting from communism, with the Confucian "guanxixue" and a tendency towards paternalism. This results in a new governance approach that protects the elites' interest (Yi, in press). This approach does not cater for the interest of the 80% of the population's children that only attain a floor performance level. Considering the U-shape relation between SES and performance, more compensatory policies centering on students from disadvantaged families should be developed within schools, especially in urban areas where a larger gap is observed in family gains. The same applies to children from migrating families who are relocated in urban areas and suffer due to lack of schooling whereas they have the potential to achieve higher performance. Considering the huge level of urbanization in China, students from migrating families or with peasant-labor parents should be supported by special policies. Secondly, based on our observations about the moderating school effect, more policies should be issued that foster the development of "quality education" throughout all schools. While the "Key school policy" - established in the late 1970s - was efficient in making good use of limited resources at the start of China as a developing and poor country (Organization of Educational Yearbook in China, 1984), it is now time to introduce policies that balance the distribution of educational resources.

Appendix 1 - Coding of career levels by Li (2005a; 2005b).

Actually, in Li's ranking, the first rank is the highest level. But in order to adopt a comparable vector direction as in other indexes, the reverse value in rank of a career level was computed (26 = highest to 26 = lowest) in the paper.

1: Official with high position in government
2: Professionals, such as professor or scientist
3: Official of the middle position in government, Dean or Head
4: Manager of enterprise, Rector of the hospital or rector of the newspaper, or headmaster
5: Cadre in government of enterprise
6: Professional, such as the reporter, lawyer or teacher
7: The manager of the company and the manager in middle position
8: The staffs in police, law-office, judiciary, business administration, tax administration, such as policeman
9: Average staffs in the government of enterprise
10: The professional in the middle of hospital, engineer, economy
11: The owner of the company
12: Managers with the middle position in the enterprise or company. such as the manager of the workshop or the head of the factory
13: The professional with middle or lower position: nurse, technical worker, primary teachers, teachers in kindergarten
14: The staffs such as the lower level secretary or accountant in the enterprise
15: The staffs such as the cleaner, managers, operation person
16: The technical person who in the rural, such as the veterinarian, doctor in rural
17: The owner of the little store or private company
18: Staffs for service of business and service company. Such as driver, barber, mail carrier
19: Worker, such as the workers in manufacture, include all the technical or not
20: The farmer: such as the fruit farmer or the fish farmer
21: The normal farmer and the fisher
22: Private worker, such as the butcher, packman, shoe-maker
23: The labor, porter (hammal), prospector, builder
24: Babysitter, servant
25: Loss of job
26: Unemployed

References

Agelasto, M., & Adamson, B. (1998). *Higher education in post-Mao China*. HongKong University Press.

Aitkin, M., Anderson, D. and Hinde, J. (1981). Statistical modeling of data on teaching style (with discussion). *Journal of the Royal Statistical Society: Series A, 144*, 419-461.

Alexander, K. L., Fennessey, J., McDill, E. L. & D'Amico, R. (1979). School SES Influences--Composition or Context? *Sociology of Education, 52* (4), 222-237.

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.

Aypay, A., Erdoğan, M., & Sözer, M.A. (2007). The variation among schools on classroom practices in science based on TIMSS-1999 in Turkey. *Journal of research in science teaching,* 44(10), 1417-1435.

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

Baker, D. P., Goesling, B., & Letendre, G. K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the "Heyneman-Loxley effect" on mathematics and science achievement. *Comparative education review, 46*(3), 291-312.

Brenner, M. (2001). Re-examining the Distribution of Wealth in Rural China. in Riskin, C., Zhao, R. W., & Li, S.. (eds), *China's Retreat from Equality income distribution and economic transition (Asia and the Pacific)*.(pp. 245-275) M.E. Sharpe: New York.

Biggs, J. B. (1997). Learning, Schooling, and Socialization: A Chinese solution to a western problem. In Lau, S. (Ed.) *Growing up the Chinese way: Chinese child and adolescent development*. (pp.147-168). Hong Kong: The Chinese university of Hong Kong.

Blau, P. M. and Ruan, D. C. (1990). In-equality of Opportunity in Urban China and America. in Kalleberg, A. L. (Ed.). *Research in Stratification and Mobility*, vol 9. (pp.3-32). CT: JAI Press.

Brooks-Gunn, J., Duncan, G. J. & Aber, J. L. (Eds). (1997). *Neighborhood poverty: Context and consequences for children: Vol. 1*. New York: Russell Sage Foundation Press.

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

Central Committee of the Communist Party of China (CPC) & State Council. (2010). *Planning of Mid-Long Term Education Reform and Development Programme*

(2010-2020). Retrieved Auguest, 15, 2009, from: http://www.gov.cn/jrzg/2010-07/29/content_1666937.htm.

Chen, J. Y-C. (2005). Guilty of indigence: the urban poor in China, 1900-1949. (Unpublished Doctor Dissertation). Yale University, United States. Retrieved from http://proquest.umi.com/pqdlink?did=1027483621&Fmt=7&clientId=27431&RQT=309&VName=PQD

Chiu, M. M., (2010). Effects of inequality, family and school on mathematics achievement: Country and student differences. *Social Forces, 4*, 1645-1676.

Chiu, M. M. & Zeng, X. H. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction, 18*(4) , 321-336.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J. M., Alexander, M., Weinfeld, F. D. and York, R. L. (Eds.) (1966). *Equality of Educational Opportunity*. Washington, D.C.: Government Printing Office.

Deng, Z. & Treiman, D. J. (1997). The Impact of the Cultural Revolution on Trends in Educational Attainment in the People's Republic of China. *American Journal of Sociology, 103* (2), 391-428.

Duncan, G. J., Boisjoly, J., & Harris, K. M. (2001). Sibling, peer, neighbor and schoolmate correlations as indicators of the importance of context for adolescent development. *Demography, 38*(3), 437-447.

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

Tamura, E. H., Menton, L. K., Lush, N. W., Tsui, F. & Cohen, W. (1997). *China*: *Understanding Its Past*. *Honolulu*: University of Hawaii Press.

Entwisle, D. R. & Astone, N. M. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development, 65*, 1521-1540.

Feng, W. (2003). Housing improvement and distribution in urban China: initial evidence from China's 2000 census. *The China Review, 3*(2), 121-143.

Fransoo, R., Ward, T., Wilson, E., Brownell, M. & Roos, N. (2005). The whole truth: Socioeconomic status and educational outcomes. *Education Canada*, *45*(3), 6-10.

Ginsburg, G. S., & Bronstein, P. (1993). Family factors related to children's intrinsic/extrinsic motivational orientation and academic performance. *Child Development, 64*, 1461-1474.

Haller, A. O. & Woelfel, J. (1972). Significant others and their expectations: concepts and instruments to measure interpersonal influence on status aspirations. *Rural Sociology*, 37, 591-621.

Heyneman, S. P., & Loxley, W. A. (1982). Influences on academic achievement across high and low income countries: A re-analysis of IEA data. *Sociology of Education, 55*(1), 13-21.

Heyneman, S. P., and 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.

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

Ho, S-C. E. & Willms, J. D. (1996) Effects of parental involvement on eighth-grade achievement. *Sociology of Education, 69* (2), 126-141.

Jeynes, W. H. (2003). A meta-analysis – The effects of parental involvement on minority children's academic achievement. *Education and urban society, 35*(2), 202-218.

Kracke, Jr, E. A. (1947). Family vs. Merit in Chinese Civil Service Examinations Under the Empire. *Harvard Journal of Asiatic studies*, 10 (2), 103-123.

Lee, V. E. & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

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.

Li, J. (2009). *Woguo Gaokao zhengce bianqian yanjiu: jiyu liyi xiangguanzhe lilun de fenxi [The evaluation of China's National Marticulation Tests Policies: based on the theory of stakeholders*. (Unpublished Doctor Dissertation). Huazhong University of Science & Technology, Wuhan, P.R. China.

Lin, N., & Xie, W. (1989). Occupational Prestige in Urban China. *American Journal of Sociology, 93*(4), 793-832.

Lin , N.. & Bian, Y. J. (1991). Getting Ahead in Urban China. *American Journal of Sociology,* 97, 657-688.

Lin, N. & Ye, X. L. (1997). *Occupational Prestige of Rural China*. Paper presented at Conference on the Chinese Triangle of Mainland-Taiwan-Hong Kong: Comparative Institutional Analysis, Toronto: North America Chinese Sociologists Association, 1997-08.

Liu, C. F., Zhang, L.X., Luo, R.F., Rozelle, S. & Loyalka, P. (2010). The effect of primary school mergers on academic performance of students in rural China. *International Journal of Educational Development, 30* (6), 570-585.

Liu, X. F. & Lu, K. (2008). Student performance and family socioeconomic status: Results from a survey of compulsory education in Western China. *Chinese Education and society, 41*(5), 70-83.

Lockheed, M. & Zhao, Q. H. (1993). The empty opportunity: local-control and secondary-school achievement in the Philippines. *International Journal of Educational Development, 13* (1), 45-62.

Louis, V. V., & Zhao, S. Y. (2002). Effects of family structure, family SES, and adult experiences on life satisfaction. *Journal of Family Issues*, 23, 986–1005.

Lu, X. Y. (2001). *Dangdai zhongguo shehui jieceng yanjiu baogao [The report on social class in contemporary China]*. Beijing: Social Sciences Academic Press.

Marks, G. N. (2006) Are between- and within-school differences in student performance largely due to socio-economic background? Evidence from 30 countries. *Educational Research*, 48(1), 21-40.

McEwan, P. J. & Trowbridge, M. (2007). The achievement of indigenous students in Guatemalan primary schools. *International Journal of Educational Development, 27*(1), 61-76.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53,* 185–204.

Ministry of Education of the P.R. China (MOE). (2001). *Decision on Reforming and Developing Basic Education by State Council*. Retrieved Auguest, 15, 2009, from http://www.moe.edu.cn/edoas/website18/26/info126.htm.

Ministry of Education of the P.R. China (MOE). (2008). *National Education Development Report to the 48th session of the international conference on education*. Retrieved Auguest, 15, 2009, from http://www.ibe.unesco.org/National_Reports/ICE_2008/china_NR08_ec.pdf.

Ministry of Education of the P.R. China (MOE). (2010). *Opinions on further promote the balanced development of compulsory education based on the scientific development idea. (Guanyu Guanche Luoshi Kexue Fazhanguan Jingyibu Tuijin Yiwujiaoyu Junheng Fazhan de Yijian).*Retrieved Auguest, 15, 2009, from http://www.moe.edu.cn/edoas/website18/32/info1263091462131332.htm.

Maslak, M. A., Kim, J. & McLoughlin, A. S. (2010). Educational engagement in China: A case from the Northwest. *International Journal of Educational Development, 30* (3), 254-262.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2007) *TIMSS 2007 Technical Report.* Retrieved from http://timss.bc.edu/isc/publications.html.

Organization for Economic Cooperation and development (OECD). (2004). *PISA 2003: Learning for tomorrow's world – First results from PISA 2003*. Paris.

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

Parish, W. L. (1981). Egalitarianism in Chinese Society. *Problems of Communism, 29*, 37-53.

Park, A., & Hannum, E. (2001). Do Teachers Affect Learning in Developing Countries? Evidence from Matched Student-Teacher Data from China." Paper prepared for the conference *Rethinking Social Science Research on the Developing World in the 21st Century*. Social Science Research Council, Park City Utah, June 7-11, 2001.

Peng, S. S. & Hall, S. T. (1995). *Understanding racial-ethnic differences in secondary school science and mathematics achievement* (NCES No. 95710). Washington D.C.: U.S. Department of Education.

Perry, Laura B. and Andrew McConney 2010 "Does the SES of the School Matter? An Examination of Socioeconomic Status and Student Achievement Using PISA 2003." *Teachers College Record, 112(4)*: 1137-1162. Retrieved Auguest, 15, 2009, from http://www.tcrecord.org ID Number: 15662.

Rowe, K. J. (1992). *Identifying Type I errors in educational and social research: Comparisons of results from fitting OLS and multilevel regression models to hierarchically structured data*. Paper presented at the Third National Social Research Conference, The University of Western Sydney, Hawkesbury, June 29 to July 2, 1992.

Sanbonmatsu, L., Kling, J. R., Duncan, G. J., & Brooks-Gunn, J. (2006). Neighborhoods and academic achievement: results from the moving to opportunity experiment. *The Journal of Human Resources. 41(*4), 649-691.

Schiller, K. S., Khmelkov, V. & Wang, X. Q. (2002). Economic development and the effects of family characteristics on mathematics achievement. *Journal of marriage and the family*. *64*(3), 730-742.

Silinskas, G., Leppanen, U., Aunola, K., Parrila, R. & Numi, J-R. (2010). Predictors of mothers' and fathers' teaching of reading and mathematics during kindergarten and Grade 1. *Learning and Instruction, 20*(1), 61-71.

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

Webster, B. J. & Fisher, D. L. (2000). Accounting for variation in science and mathematics achievement: A multilevel analysis of Australian data - Third International Mathematics and Science Study (TIMSS). *School effectiveness and school improvement, 11*(3), 339-360.

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

Wößmann, L. (2003). *Educational Production in East Asia: The Impact of Family Background and Schooling Policies on Student Performance.* Kitakyushu: Working Paper Series for The International Centre for the Study of East Asian Development, Vol. 2003-17.

Wu, X. G. (2007). Zhongguo de huji zhidu yu daiji zhiye liudong [The Chinese household registration and intergenerational social mobility]. *Shehui xue yanjiu[Sociological Studies], 6*, 38-65.

Xu, X. X. (2000). Changes in the Chinese Social Structure as Seen from Occupational Prestige Ratings and Job Preferences. *Shehuixue Yanjiu [Sociology Research]. 2000* (3), 67-85.

Xue, H. P. (2007). *Zhongguo xibu jiaoyu shengchan hanshu yanjiu – guansu nongcun chuzhong xuexiao chengji yingxiang yinsu fenxi [A study on educational production function in Western regions of China-The analysis of determinant factors of student achievement for rural secondary schools in the Gansu Province]*. (Unpublished Doctor Dissertation). Peking University, Beijing, P.R. China.

Yi, L. (In press). Auditing Chinese higher education? The perspectives of returnee scholars in an elite university. *International Journal of Educational Development*, Retrieved Auguest, 15, 2009, from: http://dx.doi.org/10.1016/j.ijedudev.2011.03.003.

Zeng, T. S.., Deng, Y. C., Yang, R. Y., Zuo, X. M., Chu, Z. H. & Li, X. J. (2007). Balanced development of compulsory education: Cornerstone of education equity. *Frontiers of Education in China, 2*(4), 469-493.

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

Zuo, F. A. (1994). Jiating shejing diwei yu xueye chengjiu zhi guanxi [The relationship between family social status and academic performance: meta-analysis]. *Jiaoyu yanjiu zixun [Education research consultation], 2*(3), 38-47.

Chapter 6

Effect of teacher's classroom teaching on mathematics performance: video analysis*

Abstract

The present study explores the nature and quality of the social interaction and the questioning approach during mathematics lessons in Chinese primary schools. A sample of 601 students and 9 teachers participated in the study. A multiple regression and multilevel analysis of mathematics performance reveal that: (1) the interaction raised by teacher contribute to the mathematics performance, and the interaction between student-student in public and teacher-one student have significantly positive impact on students' mathematics performance; (2) for the teacher questioning, the evaluation and problem-solving play important roles on mathematics performance. The results suggest that the teacher-centered interaction or student-centered interaction have their own conditions under which they can might have a positive impact. Questioning of the teachers should be focused on the cognitive thinking on the mathematics problems.

1. Introduction

In current schooling approaches, classroom teaching and learning activities take up most of the time in schools. This observation is true for education in general and mathematics education in particular. Effective teaching and learning are expected to foster the construction of mathematical knowledge and skills and foster mathematical literacy (Campell, Kyriakides, Muijs, & Robinson, 2003; Wayne & Youngs, 2003). The crucial importance of the particular teaching and learning approaches that determine related performance have invoked a growing interest in evidence-based research focusing on the detection of key instructional approaches (see e.g., De Corte, 2004; Seidel & Shavelson, 2007). For instance, the 1995 and 1999 Third International Mathematics and Science Study (TIMSS) explored teaching patterns and effective teaching approaches in mathematics instruction on the base of video studies (Hiebert

---

* This chapter is based on the submitted paper of Zhao, N.N., Valcke, M., & Desoete, A., Sang, G.Y. & Zhu, C. (submitted). Teacher's classroom teaching on students' performance in Chinese context: A video analysis. *Asia-Pacific Journal of Teacher Education*

et al., 2003; Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). This type of studies have helped to develop a basic understanding of particular instructional processes. But they have also helped to develop a knowledge base about shared pedagogical cultures in a particular country of setting. Researchers figure out that within the same country, the teachers sometimes adopted the same pedagogical culture that are linked to similar teaching patterns (Hugener, Pauli, Reusser, Lipowsky, Rakoczy, & Klieme, 2009).

In developing countries, limited studies are available about how teachers actually teach in the classroom and how this is related to teaching quality and learning outcomes (Glewwe & Kremer, 2005; Heneveld & Craig, 1996; Stephens, 1997). Within these limited studies, a "paradox" is mentioned in relation to mathematics classroom instruction in the Chinese context. This paradox stresses that a pedagogical culture that is based on "passive transmission" and "rote drilling" approaches nevertheless produces students who outperform their counterparts in the Western (Biggs, 1991; Biggs & Watkins, 1996; Marton, Dall'Alba, & Lai, 1993; Huang & Leung, 2005; Watkins & Biggs, 2001). This questions that status of the "not effective" teaching approaches and also questions what actually contributes to the high learning performance? Clearly, more empirical research is needed. The latter is also important because also within the Chinese educational culture, we observe students with very different attainment levels. It arises questions that while "rote drilling approaches" or "passive transmission approaches" seem to work in the Chinese context, there might be additional qualities linked to these instructional approaches.

The particular observations about the Chinese instructional approaches have to be reconsidered in the context of the new curriculum reform that was launched in the year 2001 (MOE, 2001). According to the new curriculum, the Ministry of Education (MOE) encourages teachers to change their classroom teaching approaches from a "teacher-centered" to a "student-centered" approach. The Chinese educational authorities want to align instruction with trends in other countries (Cuban, 1983; Schuh, 2004). This call for a redirection has raised a heavy debate in the Chinese context about the role of teachers and students in the instructional setting (Wang, 2004; Zhong, 2006). The debate is particularly heavy since Asian classes have found to be typically teacher-centered (see related research in the 1990s by the University of Michigan, Stevenson, Chen, & Lee, 1993; Stevenson & Stigler, 1992). It has also proven to be difficult to change teaching approaches teachers have been familiar with. The discussion is also fuelled due to the lack of empirical evidence that demonstrates – within the Chinese setting which approach has a differential positive and significant impact on performance.

On the one hand, Western researchers find it difficult to grasp what in the Chinese traditional teaching approaches nevertheless leads to high performance. On the other hand, the MOE in China wants to change traditional teaching approaches in favor of "good" teaching approaches adopted from the West. The exploration on the important point which is behind the teaching approaches is raised to the time schedule. It is necessary to explore the classroom interactions and the discourse between teachers and students which activate the students' cognition.

This brings us to the general aims of the present paper that present a qualitative/quantitative study, focusing on two key questions. Firstly, Is it possible to identify particular instructional interaction patterns between teachers and students that contribute to students' mathematics performance? Secondly, what patterns in instructional questioning can be identified and how are these related to mathematics performance?

2. Theoretical Background

Numerous studies have been conducted to explore operational variables that are related to classroom interaction and consistently are associated with student performance (Marzano, Pickering, & Pollock, 2001). But research remains inconclusive about satisfactory solutions to grasp the nature of classroom interaction and to find (Reusser, 2001).

2.1 Classroom interaction and mathematics performance

Analysis attempts of classroom interaction aim at exploring the observable "sight structures" (Seidel & Prenzel, 2006). A key assumption is that quality of classroom teaching and learning depends on the quality of the instructional interaction (Wang, Haertel, & Walberg, 1993). Studies have shown how different interaction patterns have a clear effect on the packaging, presentation and (re)formulation of curriculum knowledge and related learning performance (Barnes & Todd, 1995; Cazden, 2001; Inagaki, Hatano, & Morita, 1998).

Different frameworks exist to study classroom interaction. Some studies focus on the "function" of the classroom interaction: social, procedural, expository, explanatory and cognitive (Offir & Lev, 2000; Oliver & McLoughlin, 1996). Other studies focus on the role of the actors in the interaction: emitter, target, audience and residue members (Adams & Biddle, 1970, p.20). The latter approach was already adopted by Chinese researchers in the middle of the 1990s. Chinese sociologists studied the interactions in elementary schools and identified a variety of prototypical interaction models: (1) the transmitter-listener model, which refers to one teacher addressing a whole class, (2) the target-target model, which refers to the questioning between a teacher and one student or between the teacher and a group of students, (3) the audience-player model, referring to students presenting their discussion results in public and students discussing with each other in private while the teacher is part of the audience and (4) the guide-learner model, referring to students' individual work guided by a teacher (Wu, Wu, Cheng & Liu, 1995; Gao, Zhao, & Liang, 2003).

The Transmitter-Listener model refers to interaction patterns in which one teacher is the information transmitter and the whole class of students acts as listeners. Researchers state that this type of whole-class interaction does improve the effectiveness of the mathematics teaching process (Graham, Rowlands, Jennings, & English, 1999; Reynolds & Muijs, 1999). But other researchers call for caution because the relationship between this model and student performance is unclear (Alexander, 1996; Keys, 1997).

The ***Target-target*** model refers to the public interaction between teacher and one a particular student or between the teacher and a specific group of students. This interaction is also labeled as dialogic inquiry (Wells, 1999) or discussion based teaching (Boaler, 2000). It is hypothesized to enable students to acquire ways of solving the questions. In this interaction model, the teacher plays an important role since he/she facilitates a high-quality student processing level or "math-talk" (Cobb & Bauersfield, 1995). Teachers should – according to this model – especially intervene because they encourage students to verbalize their math thinking, to reconstruct solution processes and to resolve conflicting math ideas.

The ***Audience-player*** Model refers to the teachers being an audience for the work presented by single students or group of students. Since this model explicitly builds on student-student interactions, previous studies stresses that this small group work clearly gives rise to learning opportunities and is superior as compared to traditional classroom teaching (Yackel, Cobb & Wood, 1991). Students sharing their reasoning with one another is considered to be lead to cognitive activation and results in higher student performance (Redfield & Rousseau, 1981). Sharing aloud their analytical ideas about a problem, the problem context, their reasoning, and their problem solving approaches is said to activate cognitive development. During the small group interaction, teachers are expected to scaffold the the interaction (Webb, Nemer & Ing, 2006).This guarantees evaluative input that helps the groups and individual learners to improve and develop.

The ***Guide-learner*** model refers to teachers acting as a guide for their students who carry out individual work. Individual work is expected to be responsible for high levels of active thinking. Naturally, the effect will depend on the particular level of a students' engagement in the individual activity (Rimm-Kaufman, La Paro, Downer, & Pianta, 2005; Stevenson & Lee, 1995). This type of individual thinking "in private" is a very typical type of interaction in the Chinese educational context (Gao, Zhao & Liang, 2003).

A clear underlying dimension in the presentation of these interaction models is the extent to which they are "teacher-centered" or "student-centered". In the "teacher-centered" models, the teacher controls the whole class directs the learning processes to obtain a desired level of involvement and type of engagement (Wagner & McCombs, 1995). In "student-centered" model, the activities are raised by the students. For example, students or groups of students determine the nature and of and organization of the content and classroom rules (McCombs & Whisler, 1997). In general, the transmitter-listener and target-target model are in line with a teacher-centered perspective, while the audience-player and guide-learner model rather reflect a student-centered approach.

2.2 Teacher questioning during the interaction and mathematics performance

Moving to the content level of instructional classroom interactions, the discourse between teacher and students is mainly shaped by the nature and quality of teacher questioning (Green & Dixon, 1993; Nathan & Knuth, 2003). The analysis questioning behavior has already proven to be a promising direction. In the TIMSS 1995 video analysis study, it was found that compared to German lessons, students in Japanese classes reach a higher level cognitive activation due to the different nature of teacher questioning (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). In many ways, the quality of instruction seems to depend on the quality of teacher questions and especially how the latter activate mathematical cognitive processes (Black, 2007; De Corte, 2004; Seidel & Shavelson, 2007). Through questioning and answering, students have to develop and apply language structures. Language help to makes meaning explicit when students make attempts to be precise, brief and logically coherent in their mathematical language when they think, explain, argue, defend and reason (Marton & Tsui, 2004; Walshaw & Anthony, 2008). Effective classroom discourse depends as such a lot on the level of the mathematical language (Cobb & Yackel, 1996; Sherin, 2002).

Several typologies of classroom questions are available in the literature. Some are especially geared to the domains of mathematics instruction (Chapman, Rosenshine, & Meister, 1996; King, 1991). For instance, Perry, Van der Stoep and Yu (1993) constructed a questioning framework for grade one students in mathematics lessons. They distinguish between computation questions, rule recall questions, computing in context questions, make up a problem questions, problem-solving questions, identification questions, and conceptual knowledge questions. We build on the latter typology in the present study.

The question typology can also be approached from another perspective by structuring them into three clusters in the present study.

A first cluster centers on questions about ***specific mathematical knowledge and skills***, such as computations, rule recall, conceptual knowledge, computing in context and making up a problem. These questions can vary depending the particular mathematics sub-domain of the lessons.

A second cluster of questions centers on ***mathematics learning processes***, such as problem-solving strategies, and evaluation. Reviewing is used by the Chinese teachers and we will add this questioning type to the present study.

Recent research illuminates how these two clusters of questions impact cognitive learning processes due to an explicit activation of particular declarative or procedural knowledge (Daniels, 2001; Vygotsky,1978). First, due to these teacher questions students progress from concrete to more abstract ideas, they rethink their assumptions and develop clear hypotheses, they make efforts to adapt and as such (re)construct new mathematical knowledge (Hinrichsen & Jarrett, 1999). Secondly, classroom questioning dialogue also focuses on the logics of mathematical thinking when that is needed to go beyond what is currently known. This is, e.g., mirrored in students being asked to reflect on differing solution strategies (Barnes, 1976; Edwards & Mercer, 1987). Thus, on the one hand students construct new declarative and procedural knowledge, and on the other hand students develop the methods for learning mathematics (De Corte, 2003; Kramarski, Mevarech & Arami, 2002).

A third cluster of questions is related to ***classroom management***, such as identification. These questions help to direct student attention. Though this type of questions can very often been found during instructional activities, research stressed that questions that rather forstered active academic involvement, are more beneficial for classroom performance as compared to questions that emphasize classroom management (Brophy & Good, 1986; Doyle, 1977).

On the base of the theoretical and empirical findings about interaction models and the question typology, a study has been set up that links classroom interaction to mathematics learning performance. Next to data about mathematics performance, the study builds on the coding of video recordings of actual mathematics classroom sessions.

3. Research design

3.1 Participants

In the present study, nine teachers and their classes (in total 601 students) from eight schools in either a rural or an urban region, participated in the present study. During school year 2008-2009, mathematic lessons of these nine mathematics teachers were videotaped. These videotaped lessons represent the data, analyzed in view of the present study. Each mathematics lesson lasted between 35-45 minutes. Forty minutes is the average duration of a lesson in Chinese primary schools. But teachers can extend a lesson with a few more minutes during the lesson pause. In each of the participating class, one teacher did work with between 43 to 98 students (*M*≈66, *SD*≈19). Table 1 gives a summary of the sample characteristics.

Table 1
Description of the cases

Lessons	Duration of lessons (Seconds)	Number of students	Urban/rural	Grades	Knowledge domain
611200	2560	75	Urban	2	Number: multiplication
721100	2500	79	Rural	1	Number: calculation
611600	2370	75	Urban	6	Number: function
821300	2740	52	Rural	3	Number: fractions
624500	2230	47	Rural	5	Geometry: parallelogram
512200	2350	49	Urban	2	Statistics: frequency
711300	2100	98	Urban	3	Statistics: probability
822500	2350	78	Rural	5	Problem-solving
811100	2720	43	Urban	1	Number: Time and clock
Mean	2435.56	66		-	-

3.2 Methods

A mixed-method approach was adopted, combining both qualitative and quantitative methods to study the impact of the particular teaching approaches and developing a picture that was respectful for the particular culture context (Morse, 2003). Our analysis framework consisted of two approaches towards the video analysis to interpret classroom instructional processes. The first approach helped to identify the particular nature of the interaction between teachers and students in the mathematics lessons. The second approach centered on the nature and typology of the questioning behavior of the teachers. In the context of both approaches, the same dependent variable was considered: mathematics performance.

3.2.1 Dependent Variable: Mathematics performance

A newly developed, standardized instrument was used to measure the students' mathematics performance (Zhao, Valcke, Deosete, Verhaeghe & Xu, in press). This test covers the three mathematics domains of the 2001 curriculum syllabus (MOE, 2001) and focuses on twelve sub-domains: 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). In the design of the research instrument, 10,959 primary school students were involved from 20 schools in five Chinese provinces with different economic levels. The different primary school grade tests are calibrated along one continuous scale with the Bilog-MG3 program. Internal consistency (Cronbach's α) values of the grade level tests range from 0.93 to 0.96. The test helps to map student abilities ranging between -5.30 to 3.30 (M = .57, SE = .26).

From the dataset, nine teachers and their 601 students were selected in view of the present study. Considering the fact the classes were from different school grades, the student scores were standardized in order to determine student's relative position within their grade.

3.2.2 Independent variable: controlling for the school variance

The aggregated variable of the school means for the mathematics performance is used as the controlling variable to decrease the variance of the school resources and situation for the effect of the classroom teaching (Greenwald, Hedges & Laine, 1996).

3.2.3 Independent variable 1: Interaction patterns

As a first independent variable, expected to influence math performance, the study focused on the nature of the interaction patterns between teachers and students. The interaction typology was based – as discussed earlier - on the approaches adopted in previous studies (Adams & Biddle, 1970; Belleck, 1967; Offir & Lev, 2000; Gao, Zhao, & Liang, 2003). In view of the coding, the video recordings were split up into units of 10 seconds per unit and coded according to a procedure described below. Table 2 gives an overview of the coding categories.

Table 2
Interaction patterns in the mathematics classroom

Interaction	Interaction	Examples
Transmitter - Listener	One Teacher - whole class	Direct instruction
Target - Target	One Teacher - one student	- Teacher asks questions or reacts to student questions - Teacher evaluates students or manages classroom
	One Teacher - group of students	- Teacher asks questions or reacts to student questions - Teacher evaluates students or manages the classroom
Audience - Player	Students - students in public	- Students present, explain reasons
	Students - students in private	- Students discuss with partner
Guide - Learner	Individual thinking in private	- Students exercise with the help of the teacher

3.2.3 Independent variable 2: Teacher question types

The second part of the video analysis focused on teacher question types. Based on the previous studies (Perry, VanderStoep, & Yu, 1993), we identified eight question categories. The typology and related examples are presented in Table 3. Though we can expect that§ question types might vary according to the mathematics content dealt with during the lesson. Bringing the coding together of the different lessons, focusing on different math contents, we expect to deal with this problem. In addition, to be able to balance question types in relation to math content and the specific lesson setting and the total number of questions, the percentage of each question type was used as a measure in the study. The following equation was used to calculate the relative proportions:

$$\text{Percentage of each categories} = \frac{\text{Times of this categories in each class}}{\text{The sum of questioning in this class}}.$$

Table 3
Interaction patterns of the mathematics classroom

	Question type	Example
Cluster 1	1. computation/rote recall	7+3 equals?
	2. rule recall	What is the rule for two-digit addition?
	3. conceptual knowledge	Why do you use subtraction for this problem?
	4. computing in context	There are 30 pieces of red folding paper, 20 green. How many are there altogether?
	5. make up a problem	Create a problem that leads to the equation 8-3-2=3
Cluster 2	6. review	What did you learn in today's lesson?
	7. problem-solving strategies	How did you arrive at this answer?
	8. evaluation	Do others agree with the answer? Why?
Cluster 3	9. identification	Identify what is known when the teacher is directing a students' attention to important properties of a problem

3.2.4 Coding procedure

The coding was based on the following procedure: Firstly, all video sequences were transcribed by the author, taking care to note both verbal and non-verbal behaviours. Transcript conventions were adapted from approaches devised by Jefferson (2004), Goodwin and Goodwin (1986), and Oelschlaeger and Damico (2000). Secondly, two observers were trained and independently coded the lessons. The transcripts of every 10 second interval was approached as the unit for coding. The software program Nvivo 8.0 was used to carry out the coding of both the classroom interaction and the question types. An inter-observer reliability of >.80 was established (=Total agreements/Total observations).

Table 4
Coding booklet on the classroom teaching

Areas of the study	Coding strategy	Sampling scheme
A. Classroom interaction	Category system	Time units: 10 second intervals
B. Teacher questioning	Combination of category system and rating scale	Event: lessons Time units: 10 second intervals

3.3 Statistical analysis

Since the coded observation data are nested at two levels - class level and student level - a regression and multilevel regression analysis was adopted. Calculations were carried out with the software program MPlus 5.0. In addition, the analysis was also carried out on the base of a multi-level analysis considering the nested nature of the data (students, teachers, classes).

4. Result

4.1 Classroom interaction and mathematics performance

The time spent in relation to a typical classroom interaction category is depicted in Figure 1. Results in the figure 1 reveal that the interaction of one teacher-whole class (*M*=15.33, *SD*=3.76) and one teacher and individual student (*M*=12.91, *SD*=6.85) are the most popular in mathematics classes in primary schools. The results also show that interaction of student and student in public (*M*=2.09, *SD*=4.62) appears to be the third strategy used in mathematics classes. Teachers interacting with the whole class occupies about 38.66% of the time, while the teacher interacting with individual students or with a group of students, occupies 32.54% and .93% of the time. The Transmitter-Listener and Target-Target interaction type that both reflect a teacher-centered approach, determine over 70% of the lesson time.

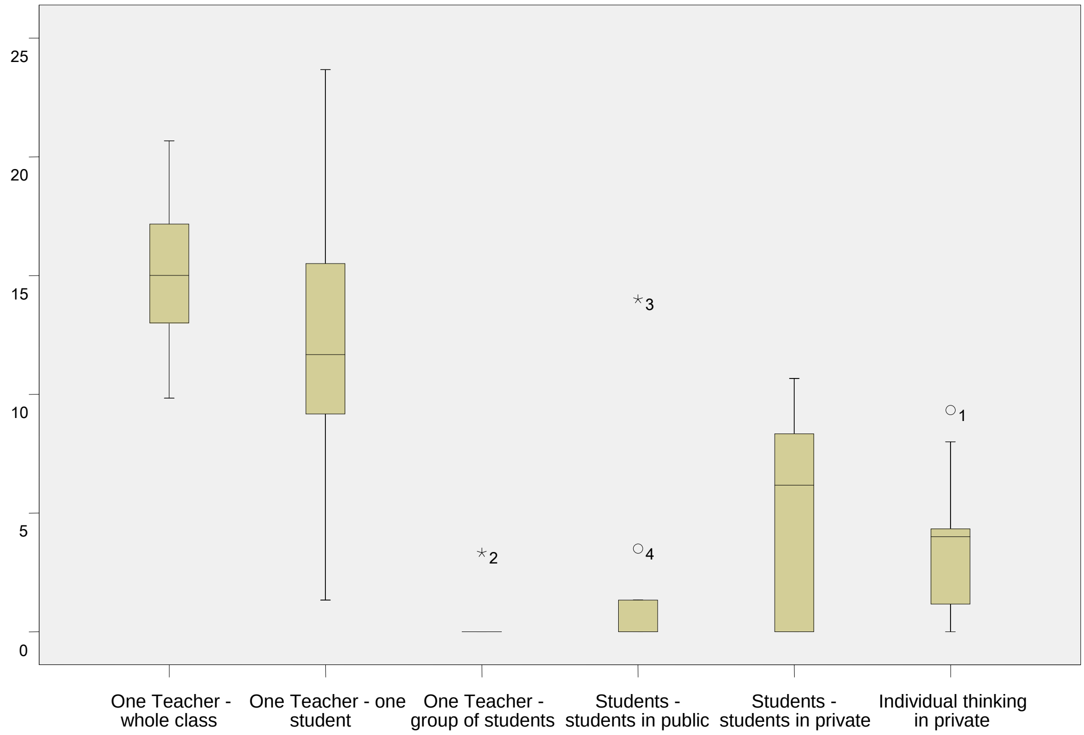


Figure 1 Boxplot of means of time on interaction between teachers and students.

Prior to the calculation of the stepwise regression, correlations between independent and dependent variables were checked (Bonferroni correction was applied). The analysis results are summarized in Table 5. Mathematics performance is significantly correlated with the different independent variables (Table 5).

Table 5
Correlations between classroom interaction types and students mathematics performance

	2	3	4	5	6	7	8
1. Mathematics performance	.72*	-.16*	.25*	.46*	.25*	-.35*	-.20*
2. School mean math performance		-.37*	.15*	.59*	.46*	-.16*	-.37*
3. One Teacher - whole class			-.12*	.44*	-.25*	.31*	.13*
4. One Teacher - one student				.60*	-.72*	-.75*	.03
5. One Teacher - group of students					-.18*	-.50*	.13*
6. Students - students in public						.58*	-.35*
7. Students - students in private							-.24*
8. Individual thinking in private							1
Mean		15.33	12.91	.37	2.09	5.30	3.67
SD		3.76	6.84	1.11	4.62	4.33	3.28

Note: * $p<(.01/9=.001)$
Bonferroni correction was done for the correlations.

The stepwise regression analysis on the individual student mathematics performance reveals that the 77% of variance can be explained by the interaction types between teachers and students. Model 1 - considering the school mean of mathematics performance - accounts for 52% of the variance (Adjusted R^2 = .52). In the consecutive models, inclusion of interactions between student-student in private, student-student in public, teacher-one student results explaining additional variance of .52, .06, .02 and .01. In the final model, 77% of the variance is due to the four independent variables. The overall statistics test of the regression indicates a significant joint contribution effect ($F(4,596)=222.15$, $p<.01$) which means that the coefficients of each independent variables are not equal to zero.

Table 6
Interaction types and the impact on mathematics performance

Step	Variables	R	R^2	ΔR^2	B	B
1	Mean school math performance	.72	.52	.52	.60	.45
2	Student-student in private	.76	.57	.06	-.08	-.35
3	Student-student in public	.77	.59	.02	.08	.39
4	Teacher-one student	.77	.60	.01	.03	.20
	Constants					

If we carry out a multilevel analysis on the same data, an analogue picture is found. In a first step, we add the school mean in mathematics performance is added to the null model. Compared with the null model, the variance at classroom level reduces from .501 to .093. In a second step, we add – step-by-step - all independent variables to the model: transmission-listener, target-target, audience-player, guide-learner. The results show that the interaction type "one teacher-one student" is an important predictor for students' mathematics performance. The variance at the class level drops from .093 to .073. Next, the interaction types "student-student in public" and "in private" are added to the model. The AIC and BIC (adjusted to the sample size) drop to the lowest coefficients of 1086.545 and 1095.112 respectively, implying this final model is most optimal. The interaction types "student-student in public" and "in private" are as such the best predictors for mathematics performance with an effect size of .356 and -.256. Except for the school means of math performance, up to 76.34% (=(.093-.022)/.093) of the variance at class level is explained by these three types of instructional interaction. The results repeatedly show teacher-centered approaches improve student performance while student-centered approaches (e.g., student-student in private) decrease student performance.

4.2 Teacher's questions and mathematics performance

As could be derived from the former section, classroom interaction types that build on teacher input, occupy most of the lesson time. In this section, we focus on teacher's questioning behavior and its impact on mathematics performance in the classroom. Figure 2 shows the percentage each types of questioning was observed during particular lesson. Questioning of problem solving strategies occupied 30.16% of the total frequencies. The following is identification and evaluation which occupied 26.61% and 11.53% of the whole frequencies.

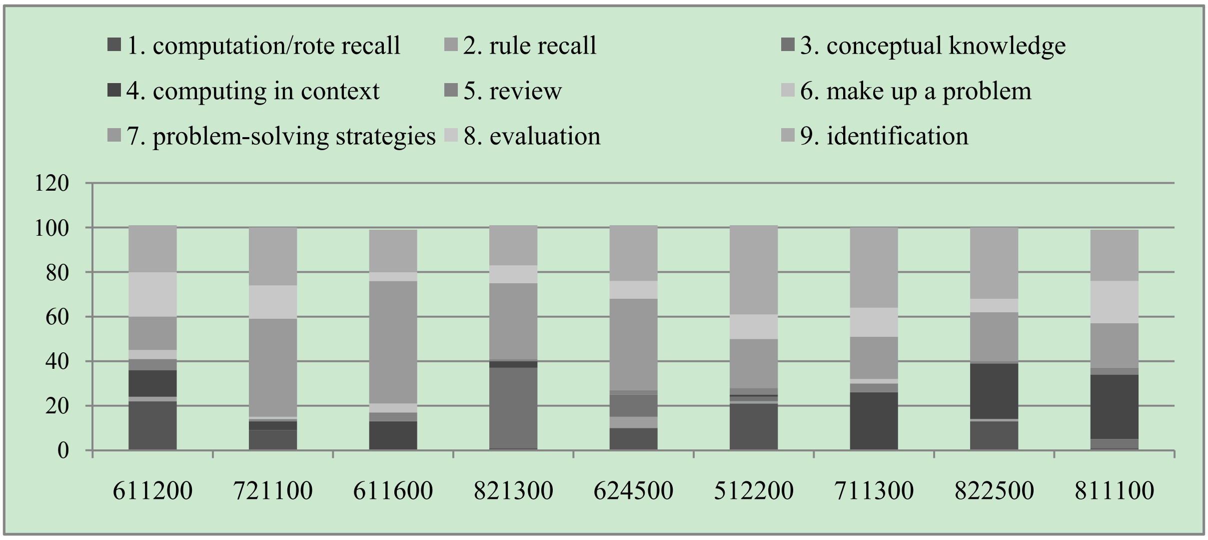


Figure 2
Lesson pattern codes on the classroom interaction.

Prior to the calculation of the stepwise regression analysis, correlation analysis indicates that independent and dependent variables are significantly correlated with one another.

Table 7
Correlations between types of teacher questions students' mathematics performance

Questioning	2	3	4	5	6	7	8	9	10	11
1.Mathematics Performance	.72*	.23*	.06	-.04	-.46*	.23*	.52*	.32*	.23*	-.49*
2. School mean math performance		.30**	.28*	-.06	-.52*	.28*	.69*	.44*	.01	-.63*
3. Compute			.50*	-.26*	-.34*	.08*	.02	-.41*	.35*	.15*
4. Rule recall				.02	-.37*	.03	-.10*	-.07	.01	-.07
5. Concept learning					-.39*	-.40*	-.39*	.12*	-.22*	-.40*
6. Compute in context						.26*	.09*	-.48*	.09*	.29*
7. Make up a problem							.80*	.13*	.18*	-.38*
8. Review								-.29*	.37*	-.04
9. Problem-solving strategies									-.55*	-.48*
10. Evaluation										-.01
11. Identification										1
Means		.03	.09	.01	.13	.01	.30	.06	.27	.12
SD		.02	.09	.02	.12	.02	.14	.12	.08	.06

Note: * $p<(.01/9=.001)$
Bonferroni correction was done for the correlations.

The results of the stepwise regression analysis are presented in the Table 8. In a first step, the inclusion of the school means of the mathematics performance accounted for 52% of the variance (Adjusted $R^2 = .52$). The inclusion of question type “rule recall” in a second step, results in an additional 5% of variance being explained (R^2 change = .05). Additng other question types “evaluation”, “computing in context”, “problem solving strategies” account for additional proportions of explained variance of 4%, 1% and 1%. The analysis results show that the sample B coefficients for the teacher questions reflect a significant joint regression coefficient effect.

In the final model, 79% of the variance in mathematics performance can be explained by the independent variables: mean school math performance, and the question types “rule recall”, “evaluation”, “computing in context” and “problem-solving strategies”. The residual R^2 =.38 can be attributed to the effects of other variables in the nine classes being observed. The overall *F*-test reveals that the sum of the square for the regression model is 338.13 while the sum of the square for the residual is 202.42. The result of the overall statistical test of the regression of the dependent variable on the independent variables shows a significant joint contribution effect($F(5,595)=198.78$, $p<.01$). There is a strong evidence that the coefficients are not equal to zero.

Table 8
Dependent variable: Students mathematics performance

Step	**Variables**	**R**	$\mathbf{R^2}$	$\mathbf{\Delta R^2}$	**B**	**β**
1	Mean school math performance	.72	.52	.52	.81	.62
2	Evaluation	.75	.57	.05	.06	.33
3	Problem solving strategies	.78	.61	.04	.01	.15
4	Rule	.78	.61	.01	-.10	-.15
5	Computing in Context	.79	.62	.01	-.01	-.15
	Constant				-.62	

Building on a multilevel analsyis approach, comparable results are found. After adding the school mean mathematics performance to the null model, all teacher question types are included step by step to the model. Rule recall questions have a negative effect on mathematics performance. However, problem solving and the evaluation questions have a positive effect on mathematics performance. The variance at class level drops from .093 to .005 in the final model (AIC and BIC - adjusted by the sample size - are 1080.479 and 1090.270 respectively). Compared to Model 2, except for the school mean of math performance, 94.6%(=(.093-.005)/.093) of the variance at class level is explained by the question types “rule recall”,” problem solving” and “evaluation”. When controlling for school diversity on the base of the school mean performance, it is cledar that the more teacher ask questions building on problem solving and evaluation, the higher the mathematics performance of the students.

5. Discussion

5.1 Teacher or students? Who directs the instructional processes?

From 2001 on, the Ministry of Education (MOE) in China tried to change the teaching and learning approaches and encourage the teachers to adopt innovative instruction all strategies (such as small group collaborative learning) instead of the dominant traditional approach that could be labeled ad a direct instruction approach. This explicit call from the educational authorities caused a strong debate and doubt as to the relevance of innovative instructional models. Although classroom interaction which is dominated by the teacher and a whole-class instructional approach is adopted as a mainstream modus in Chinese mathematics lessons (Wang, 2011), the debate indicates that this approach is not in itself a negative way of setting up instruction. Though the latter approach is often labeled as “passive transmission”, actual teacher behavior is open for student input. Both the teaching approaches propagated by the MOE and the traditional teaching approaches seem to be adopted by the teachers and mixed into the mathematics teaching and learning context.

In the present study, it nevertheless revealed that teachers dominate the Chinese classroom interaction. This is in line with the findings of previous studies (Stevenson & Lee, 1995; Lim, 2007). The results show that teacher-one student interaction in public and student-student interaction in public play a positive role in attaining mathematics performance. In contrast, the student-student interaction in private reveals a negative impact on mathematics performance. Interactions between small groups of students in private also provide learning opportunities for mathematics (Yackel, Cobb, & Wood, 1991). The question now arises why these particular interaction pattern have this particular impact in the Chinese context?

First, interactions types have specific advantages or disadvantages under different conditions (Tan, Sharan & Lee, 2007). On the one hand, interactions roused by the teacher seem to have a positive influence in classes with a larger class size. Large classes are a common feature in the Chinese context (see Table 1 where classes up to 79 pupils are listed). Teacher-centered interaction seems to support student interaction from a socio-culturual theory perspective (John-Steiner & Mahn, 1996). As beginning mathematics learners, primary school students can hardly depend on each other’s experience and need guided participation (Rogoff,1990). Teachers function as such in the zone of proximate development. They demonstrate problem solving strategies and apply questioning, supported with visual representations and verbalizations (Gersten, Chard, Jayanthi, Baker, Morphy, Flojo, 2009). After fostering student-student interaction in public, teachers direct students' individual reflection towards their problem-solving strategies. For example,

> Teacher: Do you understand the student 1 and student 2's presentation?
> Students chorus: Yes.
> Teacher: Do you have any questions about this presentation?
> Students chorus: No.
> Teacher: No? Then, I have a question for student 1 and student 2. I want to ask you. How do you divide the parallelogram into two triangles? (Teacher hold up the paper parallelogram. Student 1 and student 2 point at the diagonal and say nothing.)
> Teacher: OK. Look at this? The paper parallelogram? (Teacher points at the diagonal in the parallelogram.) I cut the parallelogram into two triangles along the diagonal of the parallelogram.
> Teacher: Come here and stick your triangles on the blackboard. Come.
> (Retrieved from transcripts 624500-Time zone 111-113)

In this case, after the student-student interaction in public, the teacher checks whether the students understand the students' presentation. Though students say "in chorus" they understand, teachers explicitly check whether individual students understand this particular insight. Often, after the interaction in public, teachers also give individual remarks and invoke student reflection on the discourse. In the latter case, since student 1 and student 2 couldn't explain the approach by using a mathematics language, the teacher supports them by presenting to them the concept of "diagonal". Our findings indicate that this type of teacher-one student and student-student in public interaction have a positive impact on performance; even though they are rather teachers-centered.

Student-student in private interaction is out of control of the teacher and much depends on the quality of the student involvement during the interaction (Christle & Schuster, 2003). Mere interaction is not valuable in itself.

Secondly, the present results raise the interesting question how to deal with the relationship between traditional teaching approaches and new teaching approaches. In the Chinese context, student-centered interaction is a new teaching strategy. Since 1980s, studies about student-centered interaction involving small groups of students, indicate this improves higher order thinking and can lead to higher achievement (Sharan & Ackerman, 1980). However, there is no clear relationship between the duration of the student-centered interaction and the quantity and quality of students' engagement. Especially in cases where teachers don't know how to make good use of the student-centered interaction approach in the class, research points out that the latter approach will require additional instructional time (see e.g., Tan, Sharan & Lee, 2007). The question is critical when we consider Chinese classroom sizes. Though – as reflected in the results – teachers try to partly adopt new interaction approaches, it is yet unclear how this can be achieved in large classes while still attaining satisfactory levels of higher-order thinking and performance (Zhong, 2005). And to adopt or not for the new approaches in the curriculum reform is not the only problem in Chinese reform but it occurs in all over the country following the curriculum reform (Lloyd, 1999). The variance in the time allowed for student-student interaction in private that is larger than variance of the teacher-whole class interaction (See figure 1), partly explains how teachers differ in their willingness to adopt new interaction patterns. The results clearly indicate that attention should be paid to teaching training in the context of the curriculum reform and to set up additional research to monitor the educational impact of the adoption of innovative instructional strategies. The lack of an evidence-base makes teachers reluctant to embrace the new pedagogies in the Chinese context.

5.2 Teacher questions and the scaffolding of students' mathematical thinking

Since teacher-centered approaches seem to be related to higher mathematics performance in the present study, the next research aim focused on how teacher questions scaffold students' mathematical thinking. The impact of questioning on performance has been a recurrent research theme during the past decades (Gall, 1970). Evaluation and problem-solving questions seem to have of positive impact on performance, while others question types have no impact or even a negative influence.

There are several ways to explain the particular impact of these question types. First, teachers adopting problem solving related questions ask their students to make explicit their procedural thinking and how they reach the results. At this stage, experienced teachers can emphasize important points of the mathematics content. The following is the example about the hierarchy in the calculation of arithmetic operations:

> Teacher: For this operation, who knows which part we should calculate first? The addition or the subtraction? Student 58.
> Student 58: I think we should first calculate the subtraction and then addition.
> Teacher: Ok, take your seat please. See if we put addition at the front, we should first calculate the addition. If we have subtraction t the front, then …
> Student chorus: We should calculate the subtraction first.

Teacher: Ok, just like we stand in a line: the one who is at the front, then ….
Student chorus: We should calculate it first.

(Retrieved from the transcripts 721100-Time zone 165-167)

In the above example, the teacher catches the important points of the mathematics calculation process and uses a metaphor to construct the mental image of "order in calculations" for the students. The present research results are in line with previous studies. The teacher make use of the questions that monitor and check students' understanding (Garza, 2009).

Second, teachers' evaluation questions also have a clear impact on performance. The teachers asks in this way that students evaluate their own behavior. This seems to be one of the best strategies to extract from the students information about their own understanding of the problems or tasks. The questions also give students an opportunity to reflect on their own and their peers' cognitive processes; example:

Teacher: He thinks that the area occupied 1/2 of the space. Any other idea?
Student 17: I think that the area of the circle is 2/3.
Teacher: Then, do all of you agree with student 17?
Student chorus: Yes.
Teacher: Ok, you can continue to argument, why is the area of the circle 2/3 of the entire circle? How do you think about that?
Student 17:….

(Retrieved from the transcripts 821300 -Time zone 182-184)

In this example, the teacher asks the students to reflect on the answer of a pervious student. This is followed by another question about the problem-solving process in which students have to verbalize their own cognitive processes. Evaluation questions and problem-solving questions are combined during the lesson. Teachers who adopt questions that invoke dynamic assessment (Swanson & Lussier, 2001) and influence performance in a positive way.

6. Conclusions, limitations and directions for future research

The present study reports about the result of a video analysis of classroom interaction and teacher question types and their impact on mathematics performance in Chinese primary schools. Especially the lack of empirical evidence about the relationship between interaction patterns and performance – gathered within the Chinese context – and the recent call of the Chinese educational authorities to adopt a more student-centered approach underline the relevance of the present study. The findings of the present study show how the current instructional pattern is still predominantly teacher-centered, but nevertheless already incorporates elements of student-centered approaches. These interaction patterns and the teacher centered instructional questioning approach is nevertheless related with higher performance. As a key factor, the class size of Chinese classroom settings was put forward as a contextual variable that explains the need for a precarious balance between student- and teacher-centered approaches. In addition, the results also invoke a discussion about the need for additional teacher training at pre-service and in-service level, to make teacher better acquainted with alternative instructional interaction approaches. In the present study, qualitative and qualitative approaches were combined to develop a richer picture of the relationship between instructional processes and mathematics performance.

The present study reflects some limitations. First, we have to note that our study of instructional effectiveness only centered on mathematics performance, and did not include other performance domains, such as language, sciences. In addition, no attention was paid to critical mediating variables, such as beliefs or student's attitudes. Secondly, the results are gathered from 9 teachers and their classes and a limited number of lesson periods. The results could reflect the particulars of the present sample and be less applicable to the population. Additional studies that involve a larger number of teachers and classes during a longer period of time, in view of being able to generalize the results. Third, in our analysis approach we only focused on two levels in the interaction. In large-scale studies students should be studied in their classroom setting and within their schools to be able to take into account processes and variables related to school policies or school culture. Future studies also need to pay attention to the social attitudes and skills of students. Collaborative learning research has underpinned the need to develop social interaction skills in learners in view of effective learning (Johnson & Johnson, 1989, 1969; Slavin, 1996). The adoption of innovative instructional approach cannot neglect – next to teacher development – the additional development of generic competences in learners, such as learning to interact, learning to report, setting personal objectives, etc.

Despite these limitations, the present articles repeats the paradox that is consistently observed in East-Asian mathematics classroom settings. Teacher-centered instructional approaches seem to be an effective approach for mathematics learning in Chinese classroom settings. Further research is needed to study the Chinese instructional classroom setting that helps to understand this paradox condition.

References

Adams, R.S., & Biddle B.J. (1970). *Realities of teaching: Explorations with video tape*. New York: Holt, Rinehart, and Winston.

Alexander, R. J. (1996). *Other primary schools and ours: Hazards of International Comparison*. CREPE Occasional Paper 1 (Warwick, Centre for research in Elementary and Primary Education).

Barnes, D. (1976). *From communication to curriculum*. London: Penguin Books Ltd.

Barnes, D., & Todd, F. (1995). *Communication and learning revisited: Making meaning through talk*. Portsmouth, NH: Boynton/Cook Publisher Inc.

Belleck, A.A. (1967). *The Language of the Classroom*. New York: Teachers College Press.

Biggs, J. B. (1991). Approaches to Learning in Secondary and Tertiary Students in Hong Kong: Some Comparative Studies. *Educational Research Journal, 6*, 27-39.

Biggs, J. B., & Watkins, D. A. (2001). The Chinese learner in retrospect. In D. A. Watkins & J. B. Biggs (Eds). The Chinese Learner: Cultural, psychological, and contextual influences (pp.269-285). Hongkong: Comparative Education Research Center, The University of Hong Kong; Melbourne, Australia: Australian Council for Education Research.

Black, L. (2007). Analysing cultural models in socio-cultural discourse analysis. *International Journal of Educational Research, 46*. 20-30.

Boaler, J. (2000). Mathematics from another world: Traditional communities and the alienation of learners. *Journal of Mathematical Behavior, 18*, 1-19.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.). Handbook of research on teaching (3rd ed.,pp.328-375). New York: Simon & Schuster.

Campbell, R. J., Kyriakides, L., Muijs, R. D., & Robinson, W. (2003). Differential teacher effectiveness: Towards a model for research and teacher appraisal. *Oxford Review of Education, 29*, 347-362.

Cazden, C. (2001). Classroom discourse: The language of teaching and learning. Portsmouth, NH: Greenwood Press.

Chapman, S., Rosenshine, B. & Meister, C. (1996). Teaching Students to Generate Questions: A Review of the Intervention Studies. *Review of Educational Research,* 66 (2), 181-221.

Christle, C. A., & Schuster, J. W. (2003). The effects of using response cards on student participation, academic achievement, and on-task behavior during whole-class, math instruction. *Journal of Behavioral Education, 12*, 147-165.

Cobb, P. & Bauersfield, (1995). Introduction: The coordination of psychological and sociological perspectives in mathematics education. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 1–16). Hillside, NJ: Lawrence Erlbaum Associates.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*, 175-190.

Cuban, L. (1983). How did teachers teach, 1890-1980. Theory into Practice, 22(3), 160-165.

Daniels, H. (2001) *Vygotsky and Pedagogy*. London: Routledge.

De Corte, E. (2003). Designing learning environments that foster the productive use of acquired knowledge and skills. In E. De Corte, L. Verschaffel, N. Entwistle, & J. J. G. van Merrienboer (Eds). Powerful learning enviornments: unravelling basic components and dimensions (pp. 21-33). Amsterdam: Pergamon.

De Corte, E. (2004). Mainstreams and perspectives in research on learning (mathematics) from instruction. *Applied psychology – an international review. 53*, 279-310.

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

Doyle, W. (1977). Paradigms for research on teacher effectiveness. In L. S. Shulman (Ed.). Review of research in education (Vol. 5, pp. 163-198). Washington, DC: American Educational Research Association.

Edwards, D., & Mercer, N. (1987). *Common knowledge: The development of understanding in the classroom*. London: Routledge.

Fried, M., & Amit, M. (2005). The unfolding of lessons and their pedagogical rationale as a criterion for international comparisons: the case of teaching systems of linear equations. Paper presented at the *11th conference of the European Association for Research on Learning and Instruction*, Nicosia, Cyprus.

Gall, M. D. (1970). The use of the questions in teaching. *Review of educational research, 40*, 707-721.

Gao, L. B., Zhao, N. N., & Liang, C. F. (2003). The studies on the observation on classroom interaction in primary school [Ketang guancha jiaowang yanjiu]. In *Educational Science Research[*jiaoyu kexue yanjiu*]. 6*, 25-28.

Garza, R. (2009). Latino and white high school students' perceptions of caring behaviors: Are we culturally responsive to our students? *Urban Education, 44*, 297-321.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: a meta-analysis of instructional components. Review of Educational Research, 79, 1202-1242.

Gersten, R., Chard, D.J., Jayanthi, M., Baker, S.K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*, 1202-1242.

Graham, T., Rowlands, S., Jennings, S., & English, J. (1999). Towards whole-class interactive teaching. *Teaching Mathematics and Its Application, 18*, 50-60.

Green, J., & Dixon, C. (1993). Introduction to special issue. Talking knowledge into being: Discursive and social practices in classroom. *Linguistics and Education. 5*: 231-239.

Greenwald, R., Hedges, L. V., Laine, R. D. (1996). The effect of school resources on student achievement. *Review of educational research, 66* (3), 361-396.

Heneveld, W., & Craig, H. (1996). *School count: world bank project designs and the quality of primary education in sub-Ssharan Africa*. Washington DC, The World Bank.

Hiebert, J., Gallimore, R., Garnier, K., Bogard Givvin, K. Hollingsworth, J., & Jacob, J., et al. (2003). *Teaching mathematics in seven countries. Results from the TIMSS 1999 video study*. Washington, DC: National Center for Education Statistics, US Department of Education.

Hinrichsen, J.,& Jarrett, D. (1999) *Science inquiry for the classroom: A literature review*. Oregon: Northwest Regional Educational Laboratory.

Huang, R. J. & Leung, K. S. F. (2005). Cracking the paradox of Chinese learners: looking into the mathematics classrooms in Hong Kong and Shanghai. In Fan, L.H. (Eds). How Chinese learn mathematics: perspectives from insiders.(pp. 348-381) World Scientific Publishing Company.

Hugener, I., Pauli,C., Reusser, K., Lipowsky, F., Rakoczy K., & Klieme, E..(2009). Teaching patterns and learning quality in Swiss and German mathematics lessons. *Learning and Instruction, 19*, 66-78.

Inagaki, K., Hatano, G..,& Moritas, E.: 1998, Construction of mathematical knowledge through whole class discussion, *Learning and Instruction, 8*, 503–526.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. Lerner (Ed.), Conversation analysis: Studies from the first generation. Amsterdam/Philadephia: John Benjamins.

Johnson, D. & Johnson, R.T. (1996). Cooperation and the use of technology. In: D. Jonassen (Ed.). *Handbook of Research for Educational Communications and Technology*, 1017-1044. London: MacMillan.

Johnson, D.W. & Johnson, R.T. (1989). *Cooperation and Competition: Theory and Research*. Edina, MN: Interaction.

John-Steiner, V., & Mahn, H. (1996). Sociocultural approaches to learning and development: A Vygotskian Framework. *Educational Psychologist, 31*, 191-206.

Keys, W. (1997). Behind the headlines: what do the results of TIMSS really tell us? *Education Journal, September,* 20-22.

King, A. (1991). Effects of training in strategic questioning on Children's problem-solving performance. *Journal of Educational Psychology, 83*(3), 307-317.

Kramarski, B., Mevarech, Z. R., & Arami, M. (2002). The effects of metacognitive instruction on solving mathematical authentic tasks. *Educational Studies in Mathematics, 49*, 225-250.

Kunter, M., & Baumert, J., (2006). Linking the TIMSS to research on learning and instruction: a reanalysis of the German TIMSS and TIMSS video data. In S. J. Howie, & T. Plomp (Eds.), *Learning mathematics and science: lessons learned from TIMSS* (pp.335-351). London: Routledge Falmer.

Leung, F. K. S. (1995). The mathematics classroom in Beijing, Hong Kong and London. *Educational Studies in Mathematics, 29,* 197-325.

Leung, F. K. S. (2001). In search of an East Asian identify in mathematics education. *Educational Studies in Mathematics, 47*, 35-51.

Lim, C. S. (2007). Characteristics of Mathematics Teaching in Shanghai, China: Through the lens of a Malaysian. *Mathematics Education Research Journal, 19*(1), 77-89.

Lloyd, G. M. (1999). Two teachers' conceptions of a reform-oriented curriculum: implications for mathematics teacher development. *Journal of Mathematics Teacher Education, 2,* 227-252.Marton, F., & Tsui, A. (2004). *Classroom discourse and the space of learning*. Mahwah, NJ: Lawrence Erlbaum.

Marton, F., Dall'Alba, G., & Lai, K. T. (1993). The paradox of the Chinese learner (Occasional Paper 93.1). Melbourne, Australia: Educational Research and Development Unit, Royal Melbourne Institute of Technology.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

McCombs, B.L., & Whisler, J.S. (1997). The learner-centered classroom and school: strategies for increasing student motivation and achievement. San Francisco: Jossey-Bass.

Morse, J. M. (2003). *Principles of mixed methods and multimethod research design*. A. Taskhkkori, & C. Teddlie. (Eds). Handbook of mixed methods in social and behavioral research (pp. 189-208). Thousand Oaks: Sage Publications.

Nathan, M. J., & Knuth, E. J. (2003). A study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction, 21*, 175-207.

Nelson, K. (1995) From spontaneous to scientific concepts: continuities and discontinuities from childhood to adulthood. In: Martin, L., Nelson, K. and Tobach, E., Editors, (1995). *Sociocultural psychology: theory and practice of doing and knowing*. (pp. 229–249). New York: Cambridge Univ. Press.

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*, 147-166.

Perry, M., VanderStoep, W., & Yu, S. L. (1993). Asking questions in First-Grade Mathematics Classes: Potential influences on mathematical thought. *Journal of Educational Psychology*, *85,* 31-40.

Redfield, D. L. & Rousseau, E. W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research*, 51, 237–245.

Reusser, K. (2001). Bridging instruction to learning – where we come from and where we need to go: A research strategy and its implementation in a cross-cultural video survey in Switzerland. 9th EARLI conference Fribourg/CH, September.

Reynolds, D., & Muijs, D. (1999). *Numeracy matters: contemporary policy issues in the teaching of mathematics*. In I. Thompson (Ed.). Issues in teaching numeracy in primary schools. Buckingham: Open University Press.

Rimm-Kaufman, S. E., La Paro, K.M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in the kindergarten classroom. *Elementary School Journal*, *105*, 377–394.

Rogoff, B. (1990). Apprenticeship in thinking. New York: Oxford University Press.

Schuh, K. L. (2004). Learner-centered principles in teacher-centered practices? *Teaching and Teacher Education, 20,* 833-846.

Seidel, T., & Prenzel, M. (2006). Stability of teaching patterns in physics instruction: findings from a video study. *Learning and Instruction*, *16*, 228-240.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis research. *Review of Educational Research*, 77, 454-499.

Sharan, S. ,& Achkerman, Z. (1980). Academic achievement of elementary school children in small-group versus whole-class instruction. *Journal of experimental education*, 48, 125-129.

Sherin, M. G. (2002). When teaching becomes learning. *Cognition and Instruction, 20*, 119-150.

Shimizu, Y. (2002). Discrepancies in perceptions of lesson structure between the teacher and the students in the mathematics classroom. Paper presented at the interactive symposium, "International Perspectives on Mathematics Classroom", at the Annual meeting of the American Educational Research Association. New Orleans, April 1-5.

Slavin, R.E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. Contemporary Educational Psychology, 21 (1), 43-69.

Stephens, D. (1997). Quality of primary education, in: K. Watson, C. Modgil & S. Modgil (Eds). Educational dilemmas: debate and diversity, quality in education. (Vol. 4) London, Cassell, 51-64.

Stevenson, H. W., & Lee, S. Y. (1995). The East Asian version of whole-class teaching. *Education Policy*, *9*, 152-168.

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: The best ideas from world's teachers for improving education in classroom*. New York: Free Press.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: NY, Free Press.

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.

Swanson, H. L., & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research, 71*, 321–363.

Tan, I. G. C., Sharan, S., & Lee, C. K. E. (2007). Group Investigation Effects on Achievement, Motivation, and Perceptions of Students in Singapore. *The journal of Educational Research, 100* , 142-154.

Vosniadou, S. (2007). The cognitive-situative divide and the problem of conceptual change. *Educational Psychologist, 42*, 55-66.

Vygotsky, L.S. (1978). *Mind in society*, Cambridge: Harvard Univ. Press, MA.

Wagner, E.D. , & McCombs, B.L.(1995). Learner centered psychological principles in practice: designs for distance education. Educational Technology, 35(2), 32–35.

Walshaw, M. ,& Anthony, G. (2008). The teacher's role in classroom discourse: a review of recent research into mathematics classrooms. *Review of Educational Research, 78* , 516-551.

Wang, C. (2004). Renzhen duidai qingshi zhishi de jiaoyu sichao: zaiping you yingshi jiaoyu xiang suzhi jiaoyu zhuangui tifa de taolun [A Critical Reflection on the Thought of "Despising Knowledge" in Chinese Basic Education]. Peking University Education Review, 2(3), 5-23.

Wang, D. (2011). The dilemma of time: student-centered teaching in the rural classroom in China. *Teaching and Teacher Education, 27*, 157-164.

Wang, M.C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 6*, 249-294.

Watkins, D. A., & Biggs, J. B. (2001). The paradox of the Chinese learner and beyond. In D. A. Watkins & J. B. Biggs (Eds). The Chinese Learner: Cultural, psychological, and contextual influences (pp.3-26). Hongkong: Comparative Education Research Center, The University of Hong Kong; Melbourne, Australia: Australian Council for Education Research.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: a review. *Review of Educational Research, 71*, 89-122.

Webb, N. M., Nemer, K. M. & Ing, M. (2006). Small-Group Reflections: Parallels Between Teacher Discourse and Student Behavior in Peer-Directed Groups. *The Journal of the Learning Sciences, 15*(1), 63–119.

Wells, G. (1999). D*ialogic inquiry: Towards a sociocultrual practice and theory of education*. Cambridge University Press.

Wu, Y.J., Wu, K. N., Cheng, X.Q. & Liu, Y. S. (1995). Woguo Xiaoxue Jiaowang Shijian Goucheng de Shehuixue Fenxi. Shanghai Jiaoyu Keyan [Structure of the time on the interaction in elementary school: from the sociological perspective]. *Shanghai Research on Education, 5,* 25-29.

Yackel, E., Cobb, P., & Wood, T. (1991) Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22* (5), 390-408.

Zhao, N., Valcke, M., Desoete, A., Verhaeghe, J.P. & Xu, K. (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*.

Zhong, Q. (2006). Curriculum reform in China: Challenges and reflections. *Frontiers of Education in China. 3* (1), 370-382.

Zhong, Q.Q. (2005). Curriculum reform in China: Challenges and reflections. *Frontiers of Education in China,* 1, 370-382.

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

Appendix 1.

Transcripts of the video recordings in the classes.

Time		**Interaction Units**	**questioning Units**
00:00:00	Tape begins. Teacher entered the classroom and said “Good morning, classes.” Students stood up and replied:”Good morning, Sir”.		
00:10:00	Teacher bids for student attention		

Chapter 7

Can homework compensate for disadvantaged environments?*

Abstract

Homework aims to improve variables and processes related to academic achievement. Homework assignments are not only given by teachers. Also parents and learners themselves develop homework tasks. In the present study we focus on the nature and impact of additional homework developed by parents and learners. The analysis of data from 10,959 students enrolled in Chinese elementary schools, reveals that parents and students from disadvantaged families approach homework as a way to compensate for an unprivileged background. Students develop homework assignments depending on the extent to which their parents develop homework. The results show that the learning performance of students of parents with low level jobs and that assign a moderate level of homework, improves significantly. In contrast, student achievement of learners is significantly lower when their parents with high level jobs don't assign any homework. Students from disadvantaged families benefit largely from homework involvement. Although there is a certain compensatory impact of homework, educational authorities should nevertheless provide additional support to learners from disadvantaged families.

1. Introduction

In many countries, homework is a daily task for most students and it often requires a large amount of their daily study time. In general, it is believed that homework accelerates knowledge acquisition and related academic achievement. Previous studies mainly focused on the quality of homework and its effect on academic achievement, depending on age and academic level (Cooper, 2007). Recent studies additionally stress the impact on mediating variables, such as motivation, expectancies, mental efforts and taking responsibility for one's own learning (Eccles, 1983; Hong & Milgram, 2000; Zimmerman & Kitsantas, 2005; Warton, 1997). The latter processes and variables have shown to be positively related with academic achievement (Hong, Peng, & Rowell, 2009; Zimmerman & Kitsantas, 2005). From a

---

* This chapter is based on the submitted paper of Zhao, N.N., Valcke, M., Desoete, A. & Verhaeghe, J. (submitted) Can homework compensate for a disadvantaged socioeconomic background? *Educational Research*.

self-regulation perspective, homework can be approached as booster for handing over responsibilities to learners(Hong & Milgram, 2000; Warton, 1997; Zimmerman & Kitsantas, 2005). In Asian students, homework is considered as investing additional "effort" as is as such a key to academic success (Chen & Stevenson, 2008; Sue & Okazaki, 1990). In China, it is believed that effort can compensate for a lack in ability (Hau & Salili, 1996) and that effort could break the socio-economic barriers affecting disadvantaged families and might be helpful to promote generation and class mobility (Blau & Ruan, 1990). Thus, engaging in homework is expected to provide students - with a disadvantaged family background - with the means to attain a higher occupational level.

Despite the available research evidence about the potential impact of homework, less research is available focusing on the Chinese context. A study that considers particularities of the Chinese cultural context can incorporate the nature of the socio-economic differences (Authors, submitted) and also consider the different way homework is implemented in the Chinese setting where next to homework assigned by the teacher, also the parents and the children themselves assign homework. This results in a more complex picture of the relationship between school and family related variables and processes that affect the impact of types of homework on achievement.

1.1 Homework revisited

Homework can be defined as the tasks assigned to students by school teachers that are meant to be performed during non-school hours (Cooper, 1989). Findings about the relationship between homework and academic achievement vary on the basis of how homework is measured. A meta-analysis of previous research, set up in the period 1987-2003 yield inconsistent results (Cooper, Robinson, & Patall, 2006): some studies observed a positive effect of homework on academic achievement, others claim that there are no differences in student achievement as a function of time spent on homework and still others found a negative relationship between the workload of homework and attitude on achievement of pupils.

But homework can be approached differently in different cultures. In most western countries, studies mainly focus on homework assigned by school teachers. In Asian countries, homework is not only assigned by the teacher, but often also by the parents and the students themselves (Guo, Liu & Zhao, 2007). Parental homework involvement as well as the students' engagement in homework can be expected to be an important moderator in the relationship between socio-economic status and academic achievement. In the present paper, we focus in particular on the alternative approaches towards homework that is elaborated by parents and/or students.

Homework has been measured in various ways (Cooper, Robinson, & Patall, 2006). Mainly, three measures are adopted to determine whether and how much time students spend on their homework. First, some researchers ask the person assigning homework how long the homework should take. Second, one asks these persons how much time students actually spend on their homework. Thirdly, students are asked themselves to report about the time spent in working on homework.. The latter self-reported measures appear to be less reliable (Hallam, 2004) because the learners' perception of time might be less accurate as compared to adults. In addition, the experienced workload can bias these self-reported measures.

In China, homework is seen as an important indicator of engagement in learning; even in primary school. Current educational policies tend to emphasize homework to a lesser extent at primary school level. The Ministry Of Education (MOE, 1994) puts forward educational principles to: "reduce too heavy homework assignments for primary and secondary school students". This policy requires that teachers should not assign any homework for grade 1 students, homework time should not exceed 30 minutes for learners in grade 2 and 3, the time should not exceed 45 minutes for grade 4, the time should not exceed 60 minutes for grade 5-6. However, actual homework time requirement are still very high in practice. Guo, Liu and Zhao (2006) present evidence that homework takes 80 to 90 minutes a day for learners in grade 3 to grade 5. A self-report survey involving 6th graders of primary school reveals that 47.8% of learners spend 2 to 3 hours and 34.8% of 6th graders spend 3 to 4 hours on homework. In addition, 72.2% of the 6th graders feel that these homework requirements are acceptable (Li, 2009). For fifth graders, available data show that the 38.27% and 46.91% students in grade 5 spend either more than 2 hours or between 1.5 and 2 hours on homework. For 4th graders, the percentage are 30.56% and 50% respectively (Luan, 2007).

1.2 Variables in parents, schools and learners affecting the involvement in and impact of homework

Characteristics of parents have been known to be related to academic achievement (Sirin, 2005; White, 1982). Previous studies reveal that parents in poorer families give less support to the development of their children's autonomy (Cooper, Lindsay, & Nye, 2000). Some studies demonstrate these parents to have more difficulties getting involved in students' learning (Scott-Jones, 1984). Parental involvement seems to depends on the expectancies and values of the family (Eccles & Harold, 1996). Studies reveal a positive correlation between parental and student attitudes towards learning (Cooper, Lindsay, & Greathouse, 1998). Parental beliefs and behavior linked to academic achievement seem to be learned by children (Eccles, 1983). It is therefore not surprising that parents' socioeconomic status influences students' engagement in homework (Chiu & Zeng, 2008). The students from families with a lower SES have access to fewer educational resources and reflect a lower academic motivation (Hampden-Thompson & Johnston, 2006). The relationship between SES and students' effort and perseverance seems to be stronger in richer countries as compared to poor countries. This reinforces the access to resources effect mentioned above (Chiu & Zeng, 2008).

Though it is a consistent finding that learners in a disadvantaged situation are at risk of attaining lower achievement levels, there is also the positive finding that parental involvement in children's homework will affect children' s academic success (Fan & Chen, 2001; Hoover-Dempsey, Battiato, Walker, Reed, DeJong, & Jones, 2001). Parental involvement seems to interact with the relationship between SES and student achievement.

It is known that parents' involvement in homework does not only vary with parents' SES, but also depends on the school context (Cooper, Lindsay, & Nye, 2000; Grolnick, Benjet, Kurowski, & Apostoleris, 1997; Kerbow & Bernhardt, 1993). Parents of learners enrolled in schools in rural areas, report less learning involvement as compared to parents of children in urban or sub-urban schools (Prater, Bermudez, & Owens, 1997; Mulntire, Marion, & Quaglia, 1990). This can be linked to the extent to which parents and teachers share the same cultural beliefs (Bourdieu, 1977; Bowles & Gintis, 1976). Parents with similar beliefs were found to be more highly involved in homework as compared to other parents. Lastly, Hoover-Dempsey and Sandler (1997) found that the homework assigned by the teacher to children also reinforces the involvement of parents in homework. Special characteristics of the schools seem to affect the way students are involved in homework (Lee & Bryk, 1989). For example, a Catholic school with a strong emphasis on academic achievement, reinforces students' active participation in homework (Bryk, Holland, Lee & Carriedo, 1984). However, when school achievement is not of central student value, involvement in homework is less probable (Maehr & Nicholls, 1980).

The latter links homework involvement to student characteristics. Student attitudes towards homework seem to depend on school grade level (Cooper, Lindsay, Nye, & Greathouse, 1998). In addition, Trautwein and Lüdtke (2009) claim that students' homework motivation and efforts are determined by the value they attach to the school tasks and their expectations as to the abilities to achieve well. In this way, homework becomes a component in a complex interplay of variables that affect the engagement in self-regulated learning.

The available research is clear about direct relationships between homework and achievement and the potential impact of parent and school characteristics. But there is hardly evidence about the complex interplay between all the variables involved. In addition no research is available set up in the Chinese setting and focusing on the nature and impact of teacher and student assigned homework. This brings us to the focus of the present study.

1.3 Research questions

Building on the introductory part of this study, the present study focuses on the following key research questions. Firstly, to what extent do parents assign homework? And how is this affected by parent and school characteristics? Secondly, to what extent do students assign homework? And how is this affected by parents' assignment of homework, and parent and school characteristics? Thirdly, we question the impact of parent and student engagement in homework on achievement and how this relationship is affected by school an parent related characteristics.

2. Research design

2.1 Sample

The data used in the present study are derived from a large scale study focusing on mathematics achievement in Chinese primary schools (Zhao, Valcke, Desoete, Verhaeghe & Xu, in press). The data were obtained from pupils enrolled in one 20 schools, geographically positioned in five provinces or cities with a varying level of Gross Domestic Product (GDP). The sample was representative with respect to developmental level and level of urbanization. In total, 10,959 students from Grade 1 to Grade 6 participated in the study, of which 51.88% were male. Table 1 represents basic sample characteristics.

Table 1
Description of Sampling

	Low level job	Blue collar	White collar	Total
Rural school	852	999	176	2027
Rural-Urban school	1180	2070	743	3993
Urban school	702	2446	1791	4939
Total	2734	5515	2710	10959

2.2 Research measures

Student mathematics achievement level. All students were tested as to the mastery of three mathematics domains, reflected in the most recent Chinese curriculum standards in China (MOE, 2001): number and algebra, shape and space, statistics and probability. The test covered a range 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). The internal consistency (Cronbach's α) of the tests, at each grade, ranged from 0.93 to 0.96 (Authors, submitted). In view of the present study, on the base of their mathematics achievement score, students were categorized in one of three achievement levels: low achiever (lower 25%), average achiever (middle 50%) and high achievers (higher 25%).

School characteristic: level of urbanization. Schools are categorized based on the level of GDP and urbanization level. Three categories were distinguished: rural schools (schools in a region with a low urbanization level and low GDP), rural-urban schools (schools in a region with an average urbanisation and GDP-level), urban schools (schools in a region with a high urbanization and GDP-level)). The potential impact of the urbanization level of a school's region cannot be underestimated. Urban schools are expected to be have available more resources, better qualified teachers, a stronger impact of school policies and school leadership.

Level of parents' job. Father and mother's' jobs were initially coded according to one of the 27 hierarchical levels of Li (2005a; 2005b). In view of the present study, the father and mother's job with 27 levels each were carried out by applying the WLSM method for the factor analysis (Mplus 5.1). And the final factor scores are record by 25%, 25-75% and 75-100% categories : low job level (unschooled manual labor jobs), blue collar level (schooled manual labor jobs) and white collar job level.

Homework assignment. In the original study, questions focused on homework assignment by the teacher, the parents or the student themselves. In addition, in relation to each type of homework, students indicated whether homework lasted less than or equal or lasted more than 30 minutes to carry out.

Up to 95% of the students replied that they received homework from their teachers that required between 0 to 30 minutes to carry out. Homework, assigned by the teacher, was not incorporated in the conceptual and research models. But a larger variation was observed in the proportion of students being assigned none, short homework (= or < than 30 minutes) or longer homework (> 30 minutes) by their parents (P) or by themselves (S). Given the focus of the present study, these two variables and values were used to carry out the analyses.

2.3 The theoretical model

Five variables were linked together to develop a conceptual model. Students' mathematics achievement level (A) is adopted as the single dependent variable. In total four variables are expected to influence mathematics achievement: urbanization level of the region of the school (U), parents' job level (J), the extent to which parents assign homework (P) and the extent to which the students assign homework to themselves (S).

2.4 Data analysis

Considering the complexity of the conceptual model in which - next to direct relationships between variables - also a large number of interaction effects can be defined, a specific analysis approach was adopted. Models were designed and evaluated in view of looking for the most parsimonious model. To evaluate the goodness of the fit of the consecutive models, likelihood ratio chi-square (L^2) and Akaike's information criteria (AIC) and (BIC) were calculated with the LEMWIN package (Vermunt, 1997), a prerelease version of Latent Gold 5 (Vermunt & Magidson, 2007), and Mplus Version 5 (Muthe´n & Muthe´n, 1998–2007). The model selection processes started from the most complex full model incorporating 5-way interaction effects of mathematics achievement lavel (A), urbanisation level of the region of the school (U), parents' job level (J), the extent to which parents present homework (P) and the extent to which the students assign homework to themselves (S).

Generally, the M_0 and M_1 are said to be the nested model if all of the effect in M_1 are subset of the value contained in M_0. The difference between the L^2 between M_0 and M_1 is the test whether deleting some interactions causes the model become significantly worse as compared to the previous model. All of the information about the model selection can be asked for the authors if needed.

3. Results

3.1 Model selection

As stated above, the model selection procedure started with the most complex model. Tables, documenting the consecutive steps in the model testing procedure, can be obtained from the corresponding author. The procedure aimed at looking for the most parsimonious model , but still achieving a good fit between the theoretical model and the model reflected in the data. The final model (model 13) reflects the best fit, leaving out the impact of particular variables without losing significant information. The model represents an AIC value of -89.63 and BIC value of -819.83, being the lowest values tested. Building on this model (See Figure 1), three different logit models will be tested in the following sections of this article: (1) P | UJ {UP}; (2) S |

UJP {UPS,UJS}; and (3) A | UPSJ {UJSA, UPJA}. The following equations were applied to test the relationships in the model 13:

$$\text{Log } P_{ji} = \mu + \lambda_i^{U} + \lambda_j^{J} + \lambda_{ij}^{UJ} \qquad \text{[Model P|UJ]}$$

$$\text{Log } S_{jik} = \mu + \lambda_i^{U} + \lambda_j^{J} + \lambda_k^{P} + \lambda_{ij}^{UJ} + \lambda_{ik}^{UP} \qquad \text{[Model S|UJP]}$$

$$\text{Log } A_{ijklm} = \mu + \lambda_i^{U} + \lambda_j^{J} + \lambda_k^{P} + \lambda_l^{S} + \lambda_{ij}^{UJ} + \lambda_{ik}^{UP} + \lambda_{il}^{US} + +\lambda_{kl}^{PS} + \lambda_{ijk}^{UJP} + \lambda_{ijl}^{UJS} \qquad \text{[Model A|UJPS]}$$

To test the three logit models, Kaufman and Schervish method was further used to analyze the findings. For each categorical variables, the first category was used as the reference category to compare the logit equations (e.g., for variable J, the lowest job level was used as a reference category). Then, the odd ratios of the logit equations was determined by calculating the natural log. For instance, given a Lambda value of .07 (expected value), the natural log is calculated (=1.6032), implying that the odds of observing a particular value is 1.62 times higher than the value in the reference category. Detailed tables, reporting all the logit equation results and odds ratios, can be obtained from the author. In the next sections only summary information will be reported and some exemplary tables.

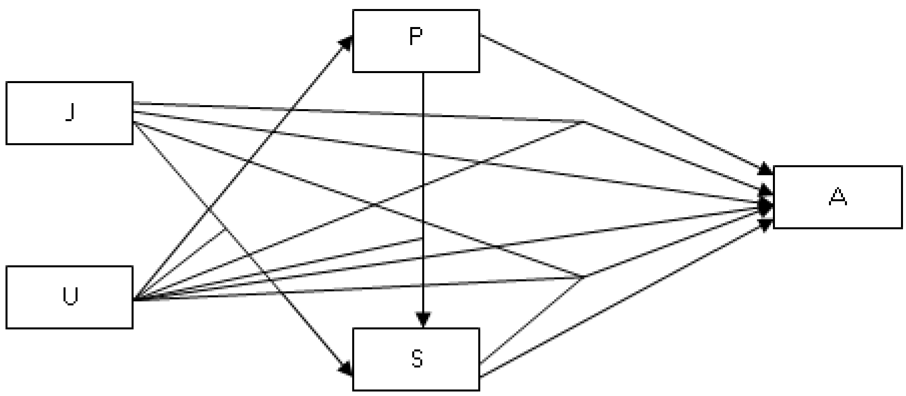


Figure 1 Final path model, resulting from the loglinear analysis

Note: Dependent variable: academic achievement (A);

Independent variables: Urbanisation level of school region (U), level of homework assignment by parents (P), level of student homework assignment (S), and parents' job level (J).

3.2 Parents' and student's homework

3.2.1 To what extent do parents assign homework to students? What other variables in the model affect the level of homework assignment by parents?

In general, parents seldom assign homework to their students. The odds ratio of parents who don't assign homework to their children is 1.6 times larger than expected on the basis of the overall effect (exp(.07)=1.6032). Parents also seldom assign homework that lasts more than 30 minutes homework. The odds ratio of parents assigning longer homework is 1.03 less than expected on the basis of the overall effect (exp(-.03)=.97=1/1.03).

Does the urbanization level of the school (U) affect the extent to which parents assign homework? First, compared to rural schools (reference category), the odds that parents in rural-urban schools assign long homework versus parents who give no homework is.02 more (=1.02-1). In urban schools, the odds that parents give long homework increases up to .38 (=1.38-1) times more as compared to parents in rural schools. The odds that parents give short homework (between 0 to 30 minutes homework), is .14 (=1.14-1) times more higher in rural-urban schools as compared to rural schools. This also increases in urban schools where the odds is .21 (=1.21-1) higher as compared to rural schools.

3.2.2 To what extent do students assign homework to themselves? How is this affected by school and parent characteristics?

An example of the impact of urbanization levels (U) and parents' job level (J) on the students' level of homework assignment (S) is shown in Table 2. Student homework assignment varies by the urbanization level of the school and the job level of the parents.

Table 2 Logit equation for the odds of students assigning "long" homework to themselves, considering the urbanization level of the school's region and parents' job level.

School /Parents	Logit Low level job	Blue collar	Equation White collar	Logit Low level job	Blue collar	Difference White collar	Odds Low level job	Blue collar	ratio White collar
Rural	.23	-.23	0						
school	-.02	-.02	-.02	.52	-.43	-.12			
	.31	-.18	-.1	0	-.95	-.64	1[a]	.39	.53
rural-urban	.23	-.23	0						
school	.2	.2	.2	.14	.02	.44			
	-.29	.05	.24	-.38	-.5	-.08	.68	.61	.92
Urban	.23	-.23	0						
school	-.18	-.18	-.18	.05	-.26	-.33			
	0	.15	-.15	-.47	-.78	-.85	.63	.46	.43

[a] Students enrolled in rural schools with low job parents represent the reference category.

The odds that students assign "long" homework to themselves (taking more than 30 minutes to tackle), is larger in students of parents with low level jobs and enrolled in rural schools. In table 2, the odds ratio is consistently lower than the reference category for students enrolled in rural-urban or urban schools and for students whose parents have a blue collar or white collar job. For instance the odds ratio that students assign long homework to themselves when their parents have a blue collar job and when they are enrolled in an urban school, is 2.17 times smaller (2.17=1/.46).

The odds that students assign "short" homework to themselves (requiring up to 30 minutes time to carry out) is higher in rural and urban schools when the parents job level is low or a blue collar job. In rural-urban schools, especially students with white collar parents assign short homework to themselves. The odds ratio is .39 times higher as compared to students with low job parents in rural schools (.39=1.39-1).

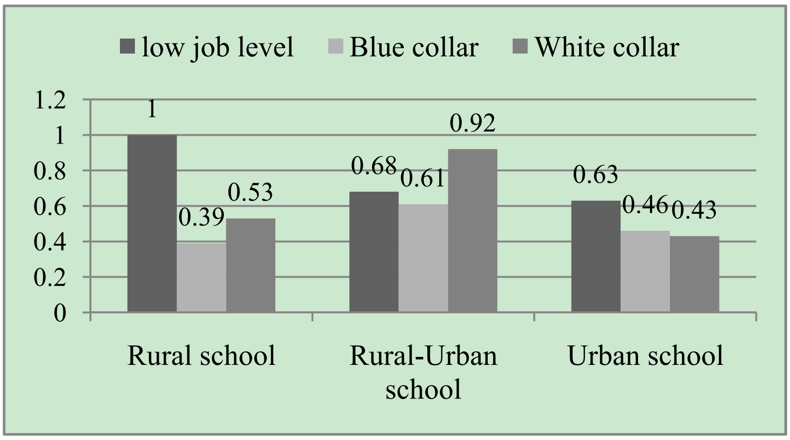


Figure 2 Odds that students assign long homework to themselves, depending on school characteristics and parent's job level.

3.2.3 Does assignment of homework by parents affect student's assignment of homework? And is this affected by the urbanization level of the school?

There is a clear relationship between parents' homework assignment on students' level of homework assignment. This is exemplified by the analysis results reported in Table 3 that focuses on the odds ratio that students assign "short" homework to themselves, depending the level of homework assigned by their parents. Compared to the reference category of parents assigning no homework to students enrolled in rural schools, the odds that students assign short homework to themselves is clearly larger when also parents assign short homework to them. For instance, these odds are .43 times higher when parents assign short homework to learners enrolled in rural schools, (.43=1.43-1); .20 times higher when they are enrolled in rural-urban schools (.20=1.39-1.19), but only .02 times higher when students are enrolled in urban schools (.02=.96-.94).

Table 3
Odds of students assigning homework to themselves, depending on parents' homework assignment and urbanization level of the school's region

	Logit		Equation	Logit		Difference	Odds		ratio
School /Parents	no	0-30	>30	no	0-30	>30	no	0-30	>30
Rural school	.05	.23	-.28						
	-.07	-.07	-.07	-.05	.31	-.46			
	-.03	.15	-.11	0	.36	-.41	1	1.43	.66
Rural-Urban school	.05	.23	-.28						
	.06	.06	.06	.12	.28	-.22			
	.01	-.01	0	.17	.33	-.17	1.19	1.39	.84
Urban school	.05	.23	-.28						
	-.18	-.18	-.18	-.11	-.09	-.34			
	.02	-.14	.12	-.06	-.04	-.29	.94	.96	.75

The last column in the table also shows how the odds of assigning short homework to themselves become smaller when parents assign long homework (taking > 30 minutes to carry out). This could be interpreted as parents "overloading" their children with homework. For instance, the odds of students assigning short homework to themselves is 1.52 times lower as compared to parents who assign no homework (1.52 (=1/.66).

If we consider the impact on students' "long" homework assignment, and we consider parents who give no homework to their students enrolled in rural schools as a reference category, the following picture emerges. The odds become smaller when parents give "short" or "long" homework. For example, these odds are 1.44 smaller (=1/(0.91/1.31)) as compared to parents who assign no homework.

The former results suggest that student homework assignment compensates parents' behavior. If the parents already give long homework, the students are less likely to assign homework to themselves. But, when parents give short homework to their children, this increases the likelihood that these students also assign short homework to themselves. The analysis results in this section seem not to be affected by the parents' job level.

3.3 Homework and mathematics performance

3.3.1 Does parent assignment of homework affect student mathematics performance level?

First, the general picture indicates that (1) if parents assign at least short homework to their children, (2) when the job level of the parents increases and (3) the urbanization level of the school increases, also the odds increase of being a high achiever. There are two exceptions: when students are from low job families and are enrolled in rural-urban schools, the odds of being a high achiever becomes very high; and, when the students are from white collar families and are enrolled in rural-urban schools, the odds of being a high achiever decrease.

When it comes to the particular impact of parents assigning homework to the students, a particular picture emerges (see Figure 3). Parent homework assignment does not seem to have an impact, especially not for the students in the rural schools. The odds of being a high achiever do not change, whatever the amount of the homework given by the parents.

When the parents' job level is aligned with the urbanization level of the region where the school is located (e.g., parents with a low level job and students are enrolled in a rural school), a U-shaped relationship can be observed in the odds of being a high achiever. In this context, assignment of short homework by the parents will contribute to students' achievement, while no homework or long homework have a small effect.

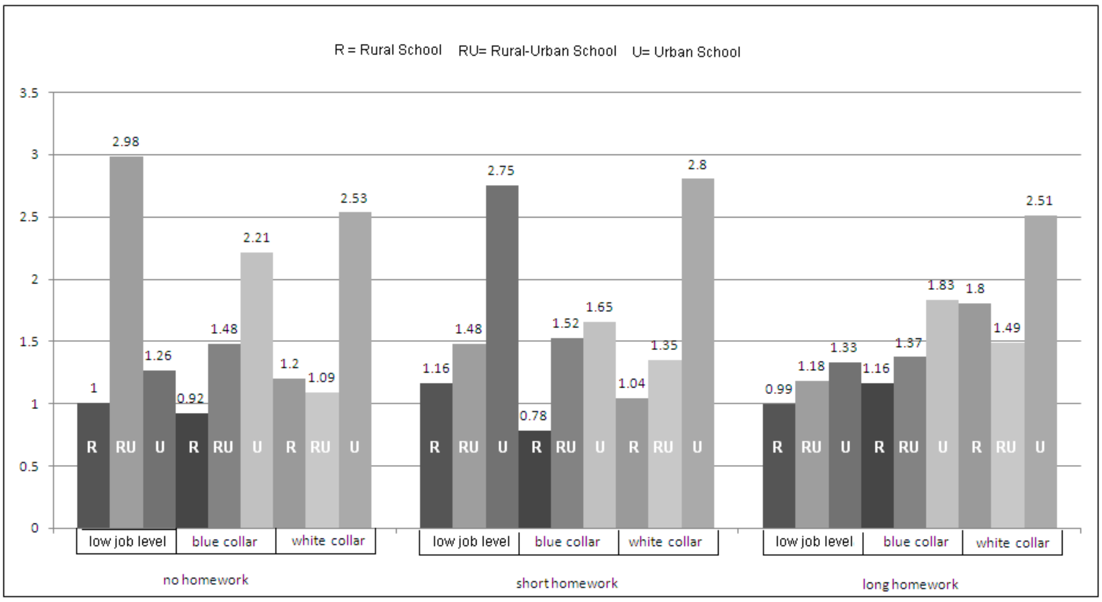


Figure 3
Odds ratios of being a high achiever depending on the level of homework assignment by parents (P) and considering urbanization level of the school's region and parents' job level.
Note: R refers to rural school; RU refers to rural-urban school; U refers to Urban school.

3.3.2 Does student assignment of homework affect their mathematics performance level?

Figure 4 shows in a transparent way the impact of student homework assignment on achievement levels. When we first focus on students with blue collar parents, the odds of being a high achiever mainly depends on the urbanization level of the school. In the figure, we see how the slope in the odds increases depending on the urbanization level of the school and hardly differs depending on the level of student homework assignment. But, when students have parents with a low level job or a white collar job, the picture becomes different.

In general, the odds of being a high achiever are larger in urban schools as compared to schools in rural or urban-rural regions. The odds of being a high achiever in rural schools, are larger when the parents have either a low level job or a white collar job level. The impact of the urbanization level of the school's region is obvious when we focus in Figure 4 on the changes in odds of students who don't assign homework to themselves. The odds of being a high achiever is larger in rural-urban and urban schools as compared to the odds in students enrolled in rural schools. In rural schools, the odds for being a high achiever for students with low job level parents, are 1.11 (=1/.9) times higher as compared to students with white collar parents. In rural-urban schools, the odds of being a high achiever for students with low level job parents is 1.42 (=2.06/1.45) times higher compared to these odds for students with white collar parents.

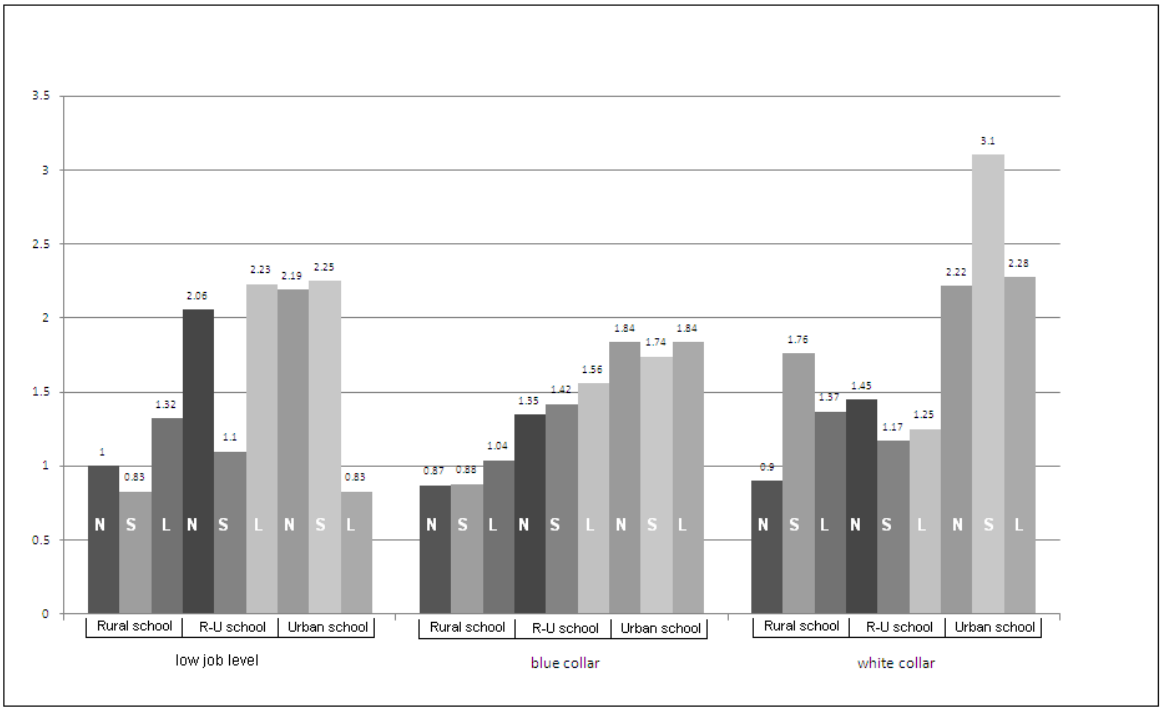


Figure 4
Odds ratios for being a high achiever, considering the level students assign themselves homework, parents' job level and urbanization level of the school.
Note: N refers to no homework; S refers to short homework; L refers to long homework.

A key question is whether the level of homework assignment by students compensates for a weaker context (e.g. rural schools and/or parents with low level jobs)? The results as reflected in Figure 4 show that also students in these less advantageous situations can change the odds of being a high achiever if they assign more homework. But, compared to the students in an advantageous situation (white collar parents and enrolled in urban schools), these extra efforts only become apparent when the latter students don't make efforts and don't assign homework to themselves. In addition, we also observe that students of parent with low level jobs even suffer when they assign long homework to themselves when they are enrolled in schools in rural-urban and in urban regions. For instance, when assigning long homework, the odds of being a high achiever is - for a student from a disadvantaged family and enrolled in an urban school - 1.2 (=1/.83) times smaller as compared to the same type of students in a rural school. The results suggest that disadvantaged students in urban schools might suffer from a dual pressure that comes from both their family background and the school context, resulting in lower achievement.

Figure 4 also show that also students from white collar families are affected by contextual variables. These odds of being a high achiever for these student are clearly higher when they are enrolled in urban schools. When they are enrolled in urban schools, they can compensate for this contextual disadvantage by assigning e.g. a sufficient level of homework. As an example, when they assign short homework, the odds that these students are high achievers become 1.96 (=1.76/.9) times higher as compared to the reference category.

4. Discussion, Limitations and Conclusions

Previous studies consistently underpinned the linear relationship between socioeconomic status and achievement (Sirin, 2005). This relationship is also confirmed in relation to the impact of homework on achievement (Cooper, Robinson, & Patall, 2006). The present study differs from earlier research in a number of ways. First, it focused on particular types of homework especially found in Asian context: homework assigned by parents and/or by students. Second, the study differs to the extent that next to family background variables (job level of the parents), we also focused on contextual school variables, such as the urbanization level of the school's region. In general, the findings suggest that students can build on homework to compensate for a disadvantaged situation. Students with parents with lower job levels assign to a higher extent short or long homework to themselves. Secondly, homework assignment seems to mediate in the relationship between parents' job level and academic achievement.

4.1 Parents' and student's homework assignment a way to compensate for a disadvantageous situation

As we can see from the results, parents' job level hardly affects the assignment of homework by the parents. This can be explained by referring to the Chinese context. Research confirms that parents in the Chinese context believe that "effort" can help to improve one's achievement level and situation (Chen & Stevenson, 2008). This belief is present at every socio-economic level, implying an openness to class mobility in the Chinese context and a favorable personal attitude towards class mobility (Blau & Ruan, 1990). But, as our results point out, variables at the school level influence parents' behavior. Parents of children enrolled in urban schools seem to be more focused on giving additional homework to their children. This is in line with findings of previous studies that urban school settings foster parents' involvement, more rather than in rural schools (Prater, Bermudez & Owens, 1997).

The assignment of homework by students to themselves reflects a more complex pattern and seems to depend on an interaction between the urbanization level of the school's region and parents' job level, next to the and interaction between school urbanization level and the extent to which parents assign homework.

The interaction effect of school characteristics level and parents' job level reflects a curvilinear relationship that was also found in earlier studies in the Chinese context (see Cooper et al, 2006; Zhao, Valcke, Deosete, Verhaeghe & Xu, in press). This curvilinear relationship is clearly exemplified in Figure 2 that represents the odd ratio of students assigning long homework to themselves. The odds that students of parents with a low job level – in rural and rural urban schools – assign long homework to themselves is clearly larger.

The latter findings are also relevant when discussing the interaction effect of school characteristics and the extent to which parents assign homework. Students that are being assigned short homework, also seem to assign short homework to themselves. This reflects the results of previous studies that parents' beliefs about the homework influence students' beliefs about homework (Cooper, Lindsay & Greathouse, 1998). But when parents too much homework, students are less likely to assign homework to themselves. The behavior of parents and students seems to be balanced.

4.2 The impact of homework on academic achievement

Many variables contribute to academic achievement. The present paper only focused on homework assignment as a catalyst for achievement. Our research results reiterate the recurrent finding in the literature that school level variables and parents' job level play an important role in mathematics performance in Chinese elementary schools. Nevertheless, our results also indicate how parents and students try to compensate for a disadvantageous situation.

The compensatory impact of homework has a small effect. This is in line with earlier studies concluding parents' and students' homework assignment have a limited impact on mathematics achievement (Cooper, Lindsay, Nye & Greathouse, 1998). The impact of homework assignment on achievement seems to fluctuate with the interaction patterns in urbanization level of the school and the parents' job level. Only when the homework assigned by the student is long, in families with low job level parents, the odds of being a high achiever become higher. But the head start of students from white collar families is still present. The moment also these students assign homework to themselves, the odds that they are a high achiever clearly become larger. These results re-confirm the results from previous study on the relationship between socioeconomic status and mathematics achievement (Zhao, Valcke, Deosete, Verhaeghe & Xu, in press).

In the Chinese context, families and society stress the importance of education (Chao, 1996). Homework has been regarded as a valued strategy to improve the achievement. Although giving too much homework can put too much pressure on students and restricts the improvement of their achievement, homework affects achievement (Lin & Chen, 1995). That clarifies why in previous studies it is found that even elementary school students spend a lot time on homework (Guo, Liu & Zhao, 2006). Although Chinese parents and students seem to believe that hard work – e.g., by assigning homework - will compensate for a disadvantaged situation, our results suggest that this is only true to a limited extent. This implies that educational authorities should be aware of this situation and that additional efforts should be made to compensate for the unfavorable background of groups of students in the school system. The results suggest that the extra efforts of the parents and/or the students themselves might not be sufficient to improve their opportunities for a better future.

4.3 Limitations and directions for future research

The present study provided specific cultural explanations about the impact of homework assignment and academic achievement. Though interesting results could be presented, also some limitations of the study should be considered. Firstly, although the study covered both students' and parents' homework assignment, it ignored the actual homework completion and its quality. With respect to quality, previous studies reveal that completion and a high quality improve academic achievement to a larger extent as compared to simply giving more homework (Cooper, Lindsay, Nye & Greathouse, 1998). A second limitation is that we studied homework assignment behavior, but ignored parents' and students' attitudes towards homework (Trautwein & Lüdtke, 2009). Our model implies that homework reflects a students' motivation. But this assumption has not been empirically tested. Our results should therefore be approached with cautiousness. Attitudes towards homework can have a different influence on children's motivation and as such affect achievement. Thirdly, we should address the issue of homework and its relationship with achievement in a more comprehensive way. Next to parents' and students' homework, other variables and processes at the student, teacher, and school level should be considered. Especially when we want to focus on compensatory mechanisms to counter the effects of a disadvantaged situation, more comprehensive and culture related issues should be focused upon in the Chinese context.

References

Blau, P. M. and Ruan, D. C. (1990). In-equality of Opportunity in Urban China and America. In Kalleberg, A. L. (ed). *Research in Stratification and Mobility*, vol 9, CT: JAI Press.

Bryk, A. S., Holland, P. B., Lee, V. E. & Carriedo, R. A. (1984). Effective Catholic schools: an exploration. Washington, DC: National Catholic Education Association.

Chao, R. K. (1996). Chinese and European American mothers' beliefs about the role of parenting in children's school success. *Journal of Cross Cultural Psychology, 27*, 403–423.

Chen, C. & Stevenson, H. W. (2008). Motivation and mathematics achievement: a comparative study of Asian-American, Caucasian-American, and East Asian high school students. *Child Development, 66* , 1215-1234.

Chiu. M. M. & Zeng, X. H. (2008). Family and motivation effects on mathematics achievement: Analyses of students in 41 countries. *Learning and Instruction, 18*, 321-336.

Cooper, H., Lindsay, J. J. & Nye, B. (2000). Homework in the home: how student, family, and parenting-style differences relate to the homework process. *Contemporary Educational Psychology, 25*, 464-487.

Cooper, H., Lindsay, J. J., Nye, B., & Greathouse, S. (1998). Relationships among attitudes about homework assigned and completed, and student achievement. *Journal of Educational Psychology, 90*, 70–83.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research, 76*, 1–62.

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

Eccles, J. S. & Harold, R. D. (1996). Family involvement in children’s and adolescents’ schooling. In A. Booth & J. F. Dunn (Eds)., Family-school links: how do they effect educational outcomes? (pp. 3-14). Mahwah, NJ: Erlbaum.

Eccles, J. S. (1983). Expectancies, values, and academic choice: Origins and changes. In J. Spence (Eds), Achievement and achievement motivation (pp. 87-134). San Francisco: Freeman.

Fan, X. & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*, 1-22.

Grolnick, W. S. & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology, 81*, 143-154.

Grolnick, W. S., Benjet, C., Kurowski, C. O., & Apostoleris, N. H. (1997). Predictors of parent involvement in children's schooling. *Journal of Educational Psychology,* 89, 538–548.

Guo, H., Liu, X. M. & Zhao, J. P. (2006). Guonei si shengqu putong zhongxiaoxue zuoye qingkuang diaocha yanjiu [A survey on students' homework in comment elementary and secondary schools of four areas of China]. Jiaoyuxuebao[Journal of Educational Studies], 6 (2), 29-39.

Hallam, S. (2004) *Homework: the evidence*, London: Institute of Education.

Hampden-Thompson, G., & Johnston, J. S. (2006). *Variation in the relationship between nonschool factors and student achievement*. Washington, DC: NCES.

Hau, K-T, Salili, F. (1996). Prediction of academic performance among Chinese students: effort can compensate for lack of ability. Organizational behavior and human decision process. 65 (2), 83-94.

Hong, E., & Milgram, R. M. (2000). *Homework: Motivation and Learning Preference*. Westport, CT: Bergin & Garvey.

Hong, E., Peng, Y. & Rowell, L. L. (2009). Homework self-regulation: Grade, gender, and achievement-level difference. *Learning and Individual Difference, 19*, 269-276.

Hoover-Dempsey, K. V., Battiato, A. C., Walker, J. M. T., Reed, R. P., DeJong, J. M. & Jones, K. P. (2001). Parental involvement in homework, *Educational Psychologist, 36*(3), 195-209.

Hoover-Dempsey, K.V., & Sandler, H.M. (1997). Why do parents become involved in their children's education? *Review of Educational. Research, 67*, 3-42.

Kerbow, D., & Bernhardt, A. (1993). Parent intervention in the school: The context of minority involvement. In J. Coleman & B. Schneider (Eds.), *Parents, their children, and schools* (pp. 115–145). Boulder, CO: Westview.

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

Li, S. Y. (2009). Shanghai diqu xiaoxue gaonianji jiating zuoye wenti de diaocha yu yanjiu [Survey on the students' homework in higher grade in shanghai]. Master thesis in Shanghai normal university.

Lin, J., & Chen, Q. (1995). Academic pressure and impact on students' development in China. *McGill Journal of Education, 30*, 149–168.

Luan, X. F. (2007). Xiaoxuesheng jiating zuoye wenti yanjiu [The study of the primary school pupils' homework]. Master Thesis in Northwest Normal University.

Maehr, M. & Nicholls, J.G. (1980). Culture and achievement motivation: a second look. In N. Warren (Ed.), *Studies in cross-cultural psychology* (vol 2). London: Academic Press.

Mclntire, W., Marion, S., & Quaglia, R. (1990). Rural school counselors: Their communities and schools. *The School Counselor,* 37(3), 166-172.

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

National educational commission. (1994). Quanmian guanche jiaoyu fangzhen, jianqing zhongxiaoxue guozhong keye fudan de yijian [Carry out the educational principle, effectively reduce too heavy homework assignments for primary and secondary school students]. Retrieved from http://www.educationalreform.net.cn/article/5.htm.

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.

Patall, E. A., Cooper, H. & Robinson, J. C. (2008). Parent involvement in homework: A research synthesis. *Review of Educational Research, 78,* 1039-1101.

Prater, D. L., Bermudez, A. B. & Owens, E. (1997). Examining parental involvement in rural, urban and suburban schools. *Journal of research in rural education, 13*(1), 72-75.

Scott-Jones, D. (1984). *Family influences on cognitive development and school achievement*. In E. Gordon (Ed.), Review of research in education (vol. 11). Washington DC: American Educational Research Association.

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

Sue, S. & Okazaki, S. (1990). Asian-American educational achievements: a phenomenon in search of an explanation. *American Psychologist, 43*, 913-920.

Trautwein, U., & Lüdtke, O. (2007). Students' self-reported effort and time on homework in six school subjects: Between-student differences and within student variation. *Journal of Educational Psychology, 99,* 432–444.

Trautwein, U., & Lüdtke, O. (2009). Predicting homework motivation and homework effort in six school subjects: the role of person and family characteristics, classroom factors, and school track. *Learning and Instruction, 19*, 243-258.

Vermunt, J.K. (1997). LEM: A General Program for the Analysis of Categorical Data. Department of Methodology and Statistics, Tilburg University.

Warton, P. M. (1997). Learning about responsibility: Lessons from homework. *British Journal of Educational Psychology, 67,* 213–221.

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

Zimmerman, B.J., & Kitsantas, A**.** (2005). Students' perceived responsibility and completion of homework: The role of self-regulatory beliefs and processes. *Contemporary Educational Psychology, 397-417.*

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

Chapter 8

Determining the variables for the children at risk of being learning difficulties*

Abstract

The purpose of the present study is to explore the effect of individual (gender, metacognition and self-efficacy) and environmental variables (family SES, teacher certification level, GDP of school region) on being at risk for experiencing mathematics learning difficulties in Chinese primary schools. A hierarchical logistic regression was applied on the data from 10959 students, of which 2738 were labeled as students experiencing learning difficulties (lower 25 percentile). A low self-efficacy level and low metacognitive experiences increase the likelihood of experiencing learning difficulties. Secondly, a lower socio-economic status of the family social status, a restricted teacher certification level and lower GDP level of the school region also increase the likelihood of being a student at risk. A school region's GDP causes the school-level unexplained variance of the random intercept to fall to .582. Adding a school's region GDP and the teacher certificate level to the model decreases the unexplained random effect variance from 20.65% and 34.7% to 9.20% and 35.3%, respectively. This underscores the additional importance of environmental variables to explain being at risk for mathematics learning difficulties.

1. Introduction

Mathematics literacy is a vital component of children's literacy in the primary school. Mathematics literacy is also viewed as a key factor in students' educational success and in acquiring future educational opportunities (Rivera-Batiz, 1992; Trust, 2009). During the past two decades, more and more studies focused on mathematics learning difficulties (Desoete, Roeyers, & De Clercq, 2004). Here, we adopt Elkins's definition for "learning difficulties". For Elkins, the term of learning difficulties is different from learning disabilities. Learning difficulties focus more on the

---

* This chapter is based on the submitted paper of Zhao, N.N., Valcke, M., & Desoete, A. (submitted). Detecting the children at risks on academic achievement in Chinese primary school by hierarchical logistic regression model. *Educational Studies*.

experiences of students. Thus, the analysis of learning difficulties focus on the interactive model of school achievement.

Part of mathematics related research focused on identifying early indicators to identify potential students at risk for learning difficulties in primary school (Geary, 2007). The latter author applies an evolutionary educational psychology perspective to introduce two types of interrelated variables explaining cognitive development: biological-primary and biological-secondary variables (Geary, 1995). The biological-primary variables comprise “individual” learner characteristics, such as gender, ability, … (Butterworth, 2003; Geary, 2003), while the biological-secondary variables link the psychological variables to the “environmental” variables (Mazzocco & Myers, 2003). The latter relationship is of educational importance, given the possibility to intervene and influence future development of learners on the base of educational interventions.

Given the state-of-the-art in the literature, there is clearly a need for paying attention to learning difficulties in educational research set up in developing countries. Most available studies about special education set up in developing countries, focus on dealing with physical disabilities (visual, speech, hearing, orthopedic), intellectual retardation, psychiatric disturbances and/or gifted children (Desai & Kothare, 2009; Qian, 2005). The same observation is true for research set up in China. In the Chinese context, the focus is hardly on students with learning problems (reading, writing, mathematics etc.), emotional problems, ADHD, Autism and so on (Qian, 2005).

The former is in sharp contrast with the observation that, in developing countries, about 200 million primary school children fail to accomplish expected attainment levels considering their developmental and cognitive potential, and this partly due to a disadvantageous setting environment (Grantham-McGregor, Cheung, Cueto, Glewwe, Richter,& Strupp, 2007). These children often mirror learning difficulties at elementary school level and are consequently less prepared for and less able to profit from further educational opportunities (Engle, Grantham-McGregor, Black, Walker, & Wachs, 2007). Early intervention programs are expected to compensate for risk that are run by these students (Walker et al.,2007; Wachs, et al., 2007). But the latter requires early screening and a thorough understanding of individual and environmental variables playing a role in the development of learning difficulties. Interventions starting too late are less likely to counter the detrimental effect of learning difficulties (Najman, Aird, Bor, Williams, & Shuttlewood, 2004). The present study extends the existing research by exploring the full complexity of individual and environmental variables in the Chinese primary school context that can be associated with the observation of learning difficulties.

1.1 Individual variables

Learning is a multifaceted process resulting from the interaction of a large set of individual variables, such as biological, cognitive and affective variables (Geary, 1995; Shuell, 1990). In the next paragraphs, we review available evidence that has related a series of individual variables to learning difficulties.

1.1.1 Demographic characteristics: Gender

A meta-analysis about the impact of gender on the acquisition of mathematics consistently points out that males outperform girls (Else-Quest, Hyde & Linn, 2010). Nevertheless, recent research repeatedly demonstrates that the impact of gender tends to decrease (Hyde & Mertz, 2009).

Limited research is available about the relationship between gender and the occurrence of learning difficulties. A particular study about dyslexia, revealed that more males than females experience this type of learning difficulty (Finucci & Childs, 1981). Some authors critiqued the related studies considering the representative nature of the subsample of learners with learning difficulties were defined (Sowder, 1984; Vogel, 2001).

1.1.2 Metacognition

Metacognition has repeatedly been found to be an adequate predictor of successful learning (Desoete & Roeyers, 2001; Trainin & Swanson, 2005). The concept "metacognition" can be described as a three-dimensional construct. The following interrelated components can be dsitinguished: metacognitive knowledge, metacognitive experiences and metacognitive skills (Efklides, 2008). Metacognitive "knowledge" has been described as the knowledge and deeper understanding of cognitive processes and products (Flavell, 1976). Brown (1987) further specifies the following metacognitive knowledge components: knowledge about oneself and others, knowledge of the task and knowledge about metacognitive strategies.Metacognitive "experiences" are what the person is aware of and what she or he feels when coming across a task and when processing the information related to it (Efklides, 2008). Metacognitive "skills" refer to the deliberate use of strategies (metacognitive regulation strategies) in order to monitor cognition (Chi, 1987; Efklides, 2008). Students reporting above average levels of metacognitive expericences and metacognitive skills, were found to achieve higher as compared to peers with below average metacognitive skills and experiences (Butler, 1998a; Cardelle-Elawar, 1992; Desoete & Roeyers, 2001; Maqsud, 1998; Schoenfeld, 1992). As a result, students with learning difficulties easily experience problems in working memory (Kruger, 2002). In past several years, some intervention on the metacognitive knowledge for the lower mathematics achiever have a good result that the intervention can improve their performance in the primary school (Cardelle-Elawar, 1995; Teong, 2003).

1.1.3 Mathematics Self-efficacy

Motivational variables comprise – among others - beliefs, attitudes and emotions that reflect the affective reactions (McLeod, 1990).

In average performing students, high self-efficacy has proven to be related to better academic achievement (Pajares & Graham, 1999). Studies involving students with learning difficulties reveal that they report a rather low self-efficacy level (Kurtz & Hicks-Coolick, 1997; Hampton & Mason, 2003). In addition, interventions focusing on students with learning difficulties showed that students' task-specific self-efficacy increases with their increase in academic skills, although their global self-efficacy did not improve (Bulter, 1998b). The latter calls for a careful and planned approach of self-efficacy in view of attaining the intervention targets (Klassen, 2002a; 2002b).

1.2 Environmental variables

1.2.1 Social economical status of the family

In the literature, learning achievement is reported to be strongly related to the socio-economic status (SES) of the learner's family (Fan, 2001; Reyes, & Stanic, 1988). Recent research questions this linear relationship. According to this research, family SES is a predictor with a rather small effect on the learning performance of average performing students (Sirin, 2005). Neverthless, previous studies also revealed that students with learning difficulties - in the mainstream classrooms – are from families with a lower social status as compared to peers without learning difficulties (Stone & Greca, 1990). Children from poor families were found to be 1.5 times more likely to reflect a learning disability as compared to peers from richer families (Duncan & Brooks-Gunn, 2001).

In addition, SES is found to be associated with a wide array of individual variables (biological, psychological and affective variables) that might lead to the occurrence of learning difficulties (Bradley & Corwyn, 2002). Multiple and complex explanations have been put forward to explain the SES-performance relationship. For instance, when learners with a disadvantageous family background are stimulated to a lesser extent. It is hypothesized that this might influence the hippocampus development in the brain (Hanson, Wolfe, Pollak, 2011); health problems (Hadley &

Patil, 2008); higher levels of stress (Conger, Conger, Elder, Lorenz, & Simons, 1994; Feldman & Walton-Allen, 1997); inadequate parenting (McClellan, 2000) etc.. Children from low income families are four times more likely than children from middle-income families to start in schools at a lower level (Jordan & Levine, 2009).

1.2.2 School related variables: teacher certification level

The quality of the "school" has been found to be a good predictor for student performance. As a consequence, in developing countries, the results of remedial interventions are less conclusive (Hanushek, 1995). The first problemis the opportunity gap between students of high and low socioeconomic status (Akiba, LeTendre & Scribner, 2007). Also, schools in poor regions were found to have more uncertified or unqualified teachers, reflecting poor academic achievement in learners (Darling-Hammond, 2004). Unqualified teachers increase the probability of student drop-out in developing countries (Bergmann, 1996). Underqualified teachers in developing countries seem not to be able to unlock children's abilities and foster academic achievement in a variety of knowledge domains (O'Sullivan, 2001). Especially teachers' content knowledge, pedagogical content knowledge and instructional experiences seem to matter in this context (Glewwe & Kremer, 2006).

1.2.3 Economic development level of the school's region

In addition to family and teacher related variables, many studies center on economics related contextual variables. For instance, the economic development level of the region where a school is located, has been found important to be a relevant predictor of academic performance (Glewwe & Kremer, 2006). Research points at the decisive influence of school policies, the social organization and the size of schools on academlic performance of learners (Bosker, Kremer, & Lugthart, 1990; Opdenakker & Van Damme, 2001; Sammons, Hillamn, & Moretimore, 1995). The latter characteristics are influenced by the economic development level of the region of a school. This is apparent in developing countries such as China. Schools in "richer" regions can develop a stronger policies, are better equipped, attract (and pay) better teachers, … as compared to schools in the "poorer" regions.

Students living in a poor region have a higher risk to be or become an underachiever (Najman, Aird, Bor, Williams, & Shuttlewood, 2004). Explanations fcor this impact refer to families in developing regions who face extra pressure due to limited community resources to cater for school demands (Miles, 2011). In addition, the educational system as a whole (such as the financial school system, the school evaluation system, …) is rather weak in developing regions with a lower GDP

(Glewwe, 2004). Research clearly shows how this affects the family and students' motivations for school (Glewwe & Kremer, 2006).

1.3 Research question

Most of the empirical literature points at the impact of particular individual and/or environmental variables that increase the risk of students to develop learning difficulties. As stated above, research is scarce that focuses on this problem in developing countries. The same applies to research set up in the Chinese context. This brings us to the central research question of this study. To what extent can inidivdual and/or environmental variables be considered as significant predictors of mathematics learning difficulties?

2. Methodology

2.1 Samples and sampling procedure

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. Of the 10959 students, 5227 are female. The age of students ranges from 5 to 14 years old.

Table 1
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%

2.2 Procedure

The research procedure was based on a protocol to standardize test administration. A time frame of 120 minutes was negotiated with school principals. Students worked individually when filling out questionnaires and test booklets. The session started with a background questionnaire, gathering student background information (age, gender) and indicators of family social status. Teachers added information about their personal background (e.g., years of teaching experience, teacher certification level).

Next, the mathematics test was presented. This test was developed on the base of the new Chinese Mathematics curriculum in the context of a large scale study, reported elsewhere (Zhao, Valcke, Desoete, Verhage, & Xu, in press). More details about the mathematics test are presented below. During a separate session, students were presented with the self efficacy test and questions related to metacognition.

2.3 Measures

2.3.1 Dependent variable: Academic achievement and mathematics learning difficulties

As explained above, all participants were presented with a newly designed mathematics test that resulted from a large scale performance indicator study in the Chinese context. The latter study resulted in six calibrated grade level tests (80 test items), covering the curriculum domains number and algebra, space and geometry, and statistics.

On the base of the test score, students were classified into one of two groups: a Normal Achievement group (NA) or a Mathematics learning Difficulties group (MD). The grouping criterion builds on the total mathematics test score. MD group allocation depended on test scores below the 25th percentile. This approach is in line with criteria presented by other authors (Geary, 2004; Mazzocco, 2007). Application of this criterion resulted in 2738 of the 10959 students being positioned in the MD category.

2.3.2 Independent variables

To develop a better understanding of mathematics performance and being at risk as a student in the MD group, individual and environmental variables were operationalized to be incorporated in a test model predicting mathematics learning achievement.

2.3.2.1 Individual variables

Individual-level variables incorporated students' demographic characteristics, students' self-efficacy, and students' metacognition.

Gender was coded as a dichotomous variables.

Mathematics Self-efficacy: The Mathematics self-efficacy SCale (MSC), designed by Marat (2005) was used to determine students' self-efficacy. A typical item is: "How well do you believe you can calculate accurately numerical problem mentally?" Students respond on the base of a 5-point Likert scale from *Not Well at all* (coded 1) to *very well* (coded 5). Cronbach's alpha reliability of the MSC is .97 and the confirmatory factor analysis reflected high goodness-of-fit indices (x^2= 41654.53, *df*=2818, *p* <.00, *GFI*=.91, *CFI*=.99, *RMSEA*=.04).

Metacognition: Metacognition measurement was based on the post diction paradigm, typically used for metacognition measurement in primary school settings (Desoete & Roeyers, 2006). Students were asked to predict the level of their test performance (e.g., 'I think I will obtain 70/100 on the test'). The Metacognition score is equal to the square of the difference between predicted scores minus the true score.

2.3.2.2 Environmental variables

Environmental variables include the family's socioeconomic status, the teachers' certification level, and the economic development level of the region where a school is located.

Family socioeconomic status: In the literature, a variety of indexes have been presented to develop an SES indicator (OECD, 2009; Olson, Martin, & Mullis, 2008).

On the present study, we build on the most dominant ingredients to develop an SES index; mother's/father's job level (ranging from 1=highest to 27 lowest), and family wealth indicators (indicated by TV, washing machine, computer and refrigerate). On the base of an Exploratory factor analysis (EFA), two factors are extracted: family wealth (FSES_W) and family job level (FSES_J). Satisfactory goodness-of-fit indexes are observed (x^2=265.80, *df*=8, *p* <.00, *CFI*=.98, *TLI*=.96, *RMSEA*=.048) on the base of a confirmatory factory analysis (CFA).

Teachers' certification level: Teachers' certification level was coded along a categorical scale expressing teacher training diploma levels: 1 (=secondary normal school), 2 (=pre-Bachelor) and 3 (= Bachelor).

Regional economic development level: The Gross Domestic Product (GDP) of the region in which the school was located, was assessed with an ordinal scale ranging from 1 (average income per person 40000+ Yuan) to 5 (average income per person 10000-15000 Yuan), building on the data from the 2005 Annual Report of the Chinese Economic Bureau.

2.4 Statistical model

The study aimed to investigate to what extent individual and environmental variables are associated with the likelihood of being positioned in the MD group with a low mathematics score. A binomial logistic regression analysis, adopting hierarchical generalized linear modeling, was carried out with the by MLWin 2.15 package (Raudenbush & Bryk, 2002).

The dependent variable in the model is a dichotomous variable distinguishing between being positioned in the NA group (0) or in the MD, at risk group (1) that mirrors mathematics learning difficulties. The independent variables are represented by the individual and environmental variables, discussed above. The model specification is as follows:

$$\text{Logit}\ \left(\frac{\text{At risk}_{ijk}}{\text{Normal}_{ijk}}\right) = \text{Logit}\ \left(\frac{\text{At risk}_{ijk}}{1-\text{At risk}_{ijk}}\right) = \beta_{0jk} + \beta_{1ijk} X_1 + \ldots \beta_{2ij} X_2 + \ldots \beta_{3i} X_{3+\ldots}$$

$$\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$$

where,

At risk_{ijk} which equal 1 if student i is at risk, and 0 if he/she is not.

In the hierarchical logistic regression, the predictive quasi-likelihood (PQL), based on 2nd order terms of the Taylor series expansion, was used for a three level logit calculation in order to avoid downward bias caused by marginal quasi-likelihood

(MQL) (Goldstein, 2003). Although the parameter of -2loglikelihood statistics for the generalized linear models with discrete responses is sometimes considered as an approximation, we will also use it as a indicator for model comparison (Goldstein, 2003). For the logistic model, Intraclass Correlation Coefficient (ICC) for each model is computed in the following way (Goldstein, 2003):

$$\mathrm{ICC}^{(s)} = \frac{\mathrm{Var}\left(v_{\mathrm{k}}^{(s)}\right)}{\mathrm{Var}\left(v_{\mathrm{k}}^{(s)}\right) + \mathrm{Var}\left(u_{\mathrm{k}}^{(s)}\right) + \frac{\pi^2}{3}}$$

3. Results

3.1 Descriptive statistics

Descriptives in relation to all variables of both students in the Normal Achievement group (NA) and the Mathematics learning Difficulties group (MD) are summarized in Table 2. According to the data in this tabel, there is a higher proportion of male students with learning difficulties, and a higher proprotion of teachers with a pre-Bacholar degree and living in a school region with a GDP between 10000 and 15000 Yuan.

Table 2 Descriptive statistics of predictors in overall, NA and MD samples (N=10959)

Variables	All M or %	SD	NA M or %	SD	MD M or %	SD
Gender						
Male	52.3%		51.9%		53.6%	
Female	47.7%		48.1%		46.4%	
Metacognitive experience	.086	.107	.057	.001	.172	.002
Self-efficacy	296.89	44.02	300.02	.482	287.48	.835
FSES_Job	.069	2.828	-.079	.031	.517	.054
FSES_W	-.014	.319	.001	.004	-.062	.006
Teacher certification level						
Secondary Normal school	2.7%		2.7%		2.8%	
Pre-Bacholar	33.0%		28.5%		46.5%	
Bacholar	64.3%		68.8%		50.7%	
GDP (in Yuan)						
Aver. per person 40000+	28.1%		28.4%		27.3%	
Aver. per person 30000-40000	15.0%		15.8%		12.8%	
Aver. per person 20000-30000	20.2%		21.4%		16.7%	
Aver. per person 15000-20000	15.3%		17.8%		7.8%	
Aver. per person 10000-15000	21.3%		16.6%		35.5%	
Total	10959		8221		2738	

Note: M=means; *SD*=standard deviation; MD= student at risk of mathematics learning difficulties; NA= normal achieving student.

3.2 Hierarchical logistic regression model

The individual-level predictors and environment-level predictors were entered step-by-step into the hierarchical logistic regression model. All continuous variables were centered around the means and the last category of the ordinal variables was used as a reference category. The logit coefficients and converted odds ratios from the three-level random-intercept binomial logistic models are listed in Table 3. Environmental variables were entered in two sub-steps: first teacher certificate level was entered, next the GDP level of the region where the school is located. The three levels are therefore labeled as individual-level, class level and school level variables.

Table 2 The logit coefficients and converted odds ratios from the three-level random-intercept binomial logistic models.

		Model 1	**Individual**	**Model 2**	**Class**	**Model 3**	**School**	**Model 4**	**Interaction**
		Logit	OR	Logit	OR	Logit	OR	Logit	OR
Level 1 Individual level									
Male		.037 (.063)	1.038	.037 (.063)	1.040	.037 (.063)	1.038	.040 (.063)	1.041
Self-efficacy		-.010 (.001)***	.990	-.010 (.001)***	.990	-.010 (.001)***	.990	-.006 (.002)**	.994
Metacognition		13.736 (.462)***	923568.300	13.737(.463)***	924492.300	13.649(.478)***	843234.200	13.641 (.474)***	839868
Family Job		.038 (.016) **		.038 (.016)**		.037 (.016)**		.037 (.016)**	
Family Job^2		-.006 (.003)*		-.006 (.003)*		-.006 (.003)*		-.006 (.003)*	
Family wealth		.450 (.135)***	1.568	.452 (.135)***	1.571	.459(.135)***	1.582	.461 (.135)***	1.586
Level 2 Class level									
(Environmental level 1)									
Teacher certificate level[a]	Normal school			-.519 (.665)	.596	-.521 (.654)	.594	-.504(.654)	.604
	Bachelor			-.617 (.261)***	.536	-.525 (.258)**	.592	-.547(.256)**	.579
Level 3 School level									
(Environmental level 2)									
Regional GDP level[b]	40000+					-1.165(.580)*	.312	-1.207(.575)**	.299
	30000-40000					-1.432(.648)**	.239	-1.477(.642)**	.228
	20000-30000					-1.384(.648)**	.251	-1.433(.642)**	.239
	15000-20000					-2.106(.812)***	.122	-2.154(.805)***	.116
Cross-level interactions									
Teacher _Normal school X	Self-efficacy							-.001(.006)	.999
Teacher _Bachelor X	Self-efficacy							-.006(.002)**	.994
Intercept		-1.994(.256)**	.136	-1.589 (.303)***	.204	-.508 (.469)	.602	-.461(.465)	.631
Intercept variance	School level	.998 (.401)**		.959 (.389)**		.582 (.267)**		.572(.262)**	
Intercept variance	Class level	2.527 (.286)**		2.471(.282)***		2.405 (.274)***		2.343(.268)***	
-2*Loglikelihood		-190.557		-293.468		-372.563		-640.355	
ICC	School level	.146		.143		.093		.092	
ICC	Class level	.370		.365		.359		.353	

Note: a: The pre-Bachelor degree is considered as the reference group; b: The region with an average GDP of 10000-15000 Yuan per person is considered as the reference group. Numbers between brackets are standard errors.

When entering the consecutive variables, the models change (see model 1 to model 4 in the table). After controlling for individual level variables, estimates of the between-school variance (regional GDP) and between-class variance (teacher certificate level) are estimated at 1.455 and 2.287, respectively. The variability at the level of the teacher certificate is higher than the variability at the regional GDP level.

Individual-level variables account for a large proportion in the overall class and school level variability in Model 1. When gender is the single variable in the model, analysis result already indicate that boys are 16.5% more likely to fall into the MD group as compared to girls. However, this effect disappears when the variable metacognition is entered into the model. While students with a lower self-efficacy level are more likely to fall into the MD category, also increasing inaccuracy in metacognition (a larger distance between predicted score and real score) highly increases the likelihood of falling into the mathematics learning difficulties group. As to the impact of SES variables, the results present a strange picture. The probability of being classified as an MD student increases with the increase in family wealth. When controlling for family wealth with the mean of the whole sample, the probability to be classified as an MD student mirrors a quadratic relationship with the parents' job level. The impact of individual-level variables are rather constant when class-level (teacher certificate) and school-level (region GDP) are included in the model, suggesting good stability of the parameter estimates.

Class level and school level variables clearly affect the likelihood of being classified as a student with mathematics difficulties. Model 2 and model 3 incorporate the environmental class-level and school-level variables. In model 2, teacher's certification level is included. Teacher's bachelor degree is significantly and negatively associated with students being at risk to fall into the MD group. Students of teachers with a pre-bachelor degree are 1.87(=1/.536) times more chances to be in the mathematics learning difficulties group as compared to students with a Bachelor degree teacher. The latter suggests that teacher's certification level contributes to the quality of the mathematics learning process. In model 3, the economic development level of the region also influences the likelihood of being at risk. Adding this variable causes the school-level variance of the random intercept to fall from .959 to .582. Students enrolled in schools located in regions with a income of 15000-20000, 20000-30000, 30000-40000 and 40000+ per person are 8.20 (=1/.122), 3.98 (=1/.251), 4.42 (=1/.239), and 3.21 (=1/.312) times less likely to be situated in the MD group as compared to students enrolled in a school that is located in a region with an average income of 10000-15000 Yuan per person.

In Table 3, also the interactions between the variables have been analysed. The results reveal that only the interaction between teacher's certification level and students' self-efficacy level significantly contributes to the model. Being taught by a teacher with a Bachelor Degree decrease the likelihood of falling into the MD group for students with lower self-efficacy levels; as compared to the impact compared to students being taught by teachers with a pre-Bachelor Degree. The predicted probabilities - based on the model 4 estimates - have been graphically represented in Figure 1.

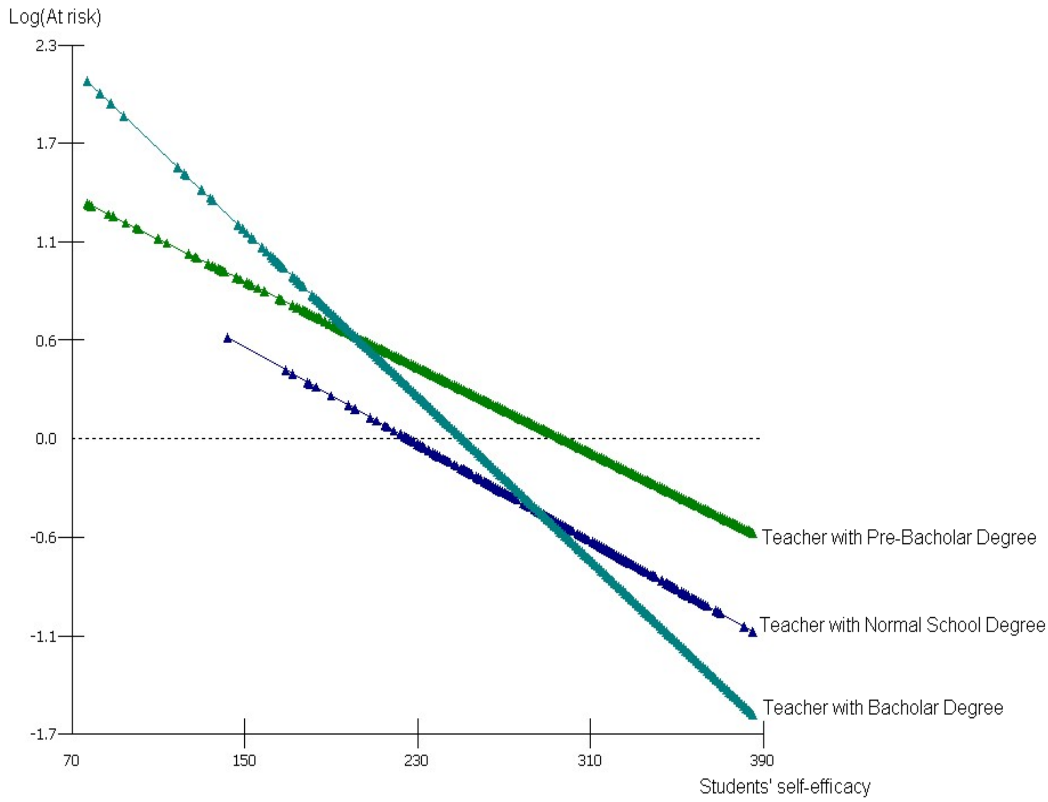


Figure 1.
Prediction of the probability of falling into the learning difficulties group, considering teacher certification level and student's level of self-efficacy.

The random-effect variances at the school and class level among students who are at the risk of falling into the MD group are .998 and 2.527, in the first model, and .572 and 2.343 in the final model. In the null model, according to the value for the intraclass correlation coefficient (ICC; another name for the variance partition coefficient in multilevel logistic regression modeling), the school and class random effect variance explain respectively, 20.65% and 34.7% of the total variance in the equation contrasting MD students with NA students. In the final model, according to ICC, the school and class random effect variances decreased into 9.20% and 35.3% of the total variance in the equation contrasting MD students with normal students.

4. Discussion, implication and directions for future research

4.1 Individual variables and learning difficulties

As explained earlier, individual variables where added to our model ,in a step-by-step way. The analysis results show that the effect of the gender disappeared when the variable metacognitive experiences is added to the model. This result implies that especially boys are assigned to the MD group. This finding is in line with the results of previous research (Else-Quest, Hyde & Linn, 2010). However, when we control for metacognitive experiences, the risk that more boys are located in the MD group disappears. This implies that there is a stronger variation in the metacognitive experiences of girls and that metacognition is important in girls with mathematics learning difficulties (Vogel, 2001; Vogel & Walsh, 1987). This finding is of importance in the broader context of diagnosing learning difficulties. The results suggest that next to performance variables, also other individual variables should be taken into account and assessed in view of remediation learning difficulties.

Students reporting a higher level in metacognitve experiences and higher self-efficacy levels are less likely to fall into the group with learning difficulties. To explain this particular effect, is less easy since self-efficacy can be influenced by earlier lower achievement (Schunk, 1991). The impact of metacognitive experiences is consistent with research already discussed above. Both individual variables, self-efficacy and metacognitive experiences have also proven to be relevant components of intervention programs as reported in the in the previous studies might also provide some good suggestions for the those two variables (Klassen, 2002a; Maqsud, 1998).

Our results pointed at the weaker impact of social economic status variables on the likelihood of being assigned to either the NA or MD group. This looks surprising, but can be explained when considering the particular Chinese context. In line with the findings of previous research set up in Chine (Authors, in press), the relationship between SES and academic achievement is not linear, but rather a curvilinear relationship is observed. Students from families with a lower SES level seem to outperform students from families with an average SES level, but still underperform as compared to students from high SES families.

Of course, in the present study, only specific SES related family variables have been considered (job level, wealth indicators, …). Future could center in addition on socio-emotional variables and motivational variables that play a role in the family climate in relation to schooling and academic performance.

4.2 Environmental variables and learning difficulties

After controlling for the individual variables, environmental variables have been entered in two steps. As expected, teacher's certificate level plays a role. This is a result that has clear policy related implications (MOE, 2011). In the Chinese context, measures have been taken to upgrade initial teacher training qualifications. In addition, clear investments can be observed to upgrade in-service teachers in general and in less developed regions in particular (MOE, 2011).

Differences in region's economic development level have been found to be associated with the likelihood of being assigned in the NA or MD group. Students enrolled in scholos that are located in poor regions have a higher risk to fall into the learning difficulties group. The latter is important when the focus is on students with learning difficulties, since earlier research could not present evidence as to the linkage between -GDP and learning performance (Authors, in press). The former results should be revisited and future research could check whether differences between the economic development level of a region is masked by the high performing students in the school within a low GDP region. In fact, in developing regions, already average performing students seem to have a risk to underachieve as was shown in previous studies (Najman, Aird, Bor, Williams, & Shuttlewood, 2004). This stresses again the need for policy makers to consider additional investments in compensation programs to b set up in particular region in China.

5. Conclusions and Limitations

The present study aimed at identifying critical variables that can be related to the risk of developing mathematics learning difficulties in the Chinese primary school context. To address this issue, logistic regression models were tested to estimate the particular impact of individual and/or environmental variables on the likelihood of being positioned in the group of normal achieving students or students with mathematics learning difficulties. Three conclusions can be presented: (a) individual variables no doubt to play most important role in being at risk for learning difficulties; (b) family SES and teacher certification level should be considered as causes that might influence future learning opportunities; (c) individual variables and environmental variables interact and explain together a large proportion of the unobserved class variability of being positioned in the group of students with learning difficulties.

However, we have to stress some limitations of the present study. Firstly, in our final model, the random-effect variances remain significantly different from zero. This implies that other - yet undefined - variables at the class and school level should be included in future models that are expected to influences the occurrence of learning difficulties. Secondly, although the interaction between teacher's certification level and students self-efficacy was discussed, the path between other environmental variables and other individual variables was not explored in this study.

Despite the former limitations, the results presented in the present paper provide a new more comprehensive perspective about students at risk of having/developing learning difficulties. from a theoretical perspective, the present study contributes to the literature that stresses the complex interaction between individual and environmental variables. From a practical perspective, the present study provides an insight into unequal opportunities between schools in different Chinese regions. GDP is a critical variable at school level that should be considered when developing future educational policies. But, as was stressed before, the Chinese context especially calls for contxt specific measures that take into account non-linear relationships between environmental and individual variables.

Reference

Akiba, M., LeTendre, G. K. & Scribner, J. P. (2007). Teacher Quality, Opportunity Gap, and National Achievement in 46 Countries. *Educational Researcher, 36* (7), 369-387.

Bergmann, H. (1996). Quality of education and the demand for education - Evidence from developing countries. *International Review of Education, 42* (6), 581-604.

Boekaert, M. (1993). Being concerned with well-being and with learning. *Educational Psychologist, 28,* 149-167.

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.Bradley, R. H. & Corwyn, R. F. (2002). Socioeconomic status and child development. Annu. Rev. Psychol. 53, 371-99.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In: F.E. Weinert & R.H. Kluwe (Eds.). *Metacognition, Motivation, and Understanding*, 65-116. Hillsdale, NJ: Lawrence Erlbaum.

Bulter, D. L. (1998). Promoting strategic learning by postsecondary students with learning disabilities. *Journal of Learning Disabilities, 28,* 170-190.

Bulter, D. L. (1998a). Metacognition and learning disabilities. In Wong, B.Y. L. (ed.) Learning about learning disabilities, Second edition (pp. 277-307). New York: Academic Press.

Butterworth, B. (2003). *Dyscalculia screener*. London: NFER Nelson Publishing Company Ltd.

Cardelle-Elawar, M. (1992). Effect of teaching metacognitive skills to students with low mathematics ability. *Teaching and Teacher Education*, 8 (2), 109-121.

Chi, M. (1987). Representing knowledge and metaknowledge: implications for interpreting metamemory research. In: F.E. Weinert & R. Kluwe (Eds.), *Metacognition, Motivation and Understanding*, 239-266. Hillsdale: Lawrence Erlbaum.

Conger, R.D., K.J. Conger, G.H. Elder Jr., F.O. Lorenz & R.L. Simons (1994) Economic stress, coercive family process and developmental problems of adolescents, *Child Development, 65,* 541-561.

Darling-Hammond, L. (2004). Inequality and the right to learn: Access to qualified teachers in California's public schools. *Teachers College Record, 106*(10), 1936–1966.

Desai, S. A. & Kothare, S. V. (2009). Neuropsychological and behavioral evaluations. *Journal of the Indian Academy of Applied Psychology, 35*, 163-172.

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

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. and De Clercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities, 37*(1), 50-61.

Duncan G. J. & Brooks-Gunn, J. (2001). Poverty, welfare reform, and children's achievement. In Biddle, B. J. (eds). Social class, poverty, and education: Policy and Practice. (pp. 49-76). New York: Routledge Almer.

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13,* 277-287.

Elkins, J. (1983) The concept of learning difficulties: An Australian perspective. In J.D. McKinney & L. Feagans (Eds.), Current topics in learning disabilities. Norwood, N.J.: Ablex Publishing Corporation.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*, 103-127.

Engle, P., Grantham-McGregor, S., Black, M., Walker, S., & Wachs, T. (2007). How to avoid the loss of potential in over 200 million young children in the developing world. *Child health and education, 1* (2), 67-87.

Every Child a Chance Trust. (2009). The long-term costs of numeracy difficulties. Available from : http://www.everychildachancetrust.org/downloads/ecc/Long%20term%20costs%20of%20numeracy%20difficulties.pdf.

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

M. A., & Walton-Allen, N. (1997). Effects of maternal mental retardation and poverty on intellectual, academic, and behavioral status of school-age children. *American Journal on Mental Retardation*, *101,* 352–264.

Finucci, J. M. & Childs, B. (1981). *Are there really more dyslexic boys than girls?* In A. Ansara, N. Geschwind, A., Calaburada, M. Albert, & N. Gartrell (Eds.), Sex differences in dyslexia (pp. 1-10). Baltimore: Orton Dyslexia Society.

Geary, D. (2003). Learning disabilities in arithmetic: problem-solving differences and cognitive deficits. In L. Swanson, K.R. Harris & S. Graham (Eds.) *Handbook of learning disabilities* (pp. 199-212). NY: Guilford Press.

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist, 50*, 24-37.

Geary, D. C.(2007). An evolutionary perspective on learning disability in mathematics. Developmental Neuropsychology, 32 (1), 471-519.

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

Glewwe, P., & Kremer, M. (2006). *Schools, teachers, and education outcomes in developing countries*. In Hanushek, E. A. & Welch, F. (Eds). Handbook of the economics of education (2rd). Amsterdam and Oxford: Elsevier.

Glewwe, Paul (2004) *An Investigation of the Determinants of School Progress and Academic Achievement in Vietnam*, in P. Glewwe, D. Dollar and N Agrawal, eds., Economic Growth, Poverty, and Household Welfare in Vietnam. The World Bank. Washington, DC.

Goldstein, H. (2003). *Multilevel Statistical Models (3rd Edition).* London: Edward Arnold: New York, Halstead Press.

Grantham-McGregor S, Cheung Y, Cueto S, Glewwe P, Richter L, Strupp B. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet,* 369, 60–70.

Hadley, C. & Patil, C. L. (2008). Seasonal changes in household food insecurity and symptoms of anxiety and depression. American Journal of Physical Anthropology, 135: 225-232.

Hamptona, N. Z. & Mason, E. (2003). Learning Disabilities, Gender, Sources of Efficacy, Self-Efficacy Beliefs, and Academic Achievement in High School Students. *Journal of School Psychology, 41,* 101– 112.

Hanson, J. L., Chandra, A., Wolfe, B. L., Pollak, S. D. (2011). Association between Income and the Hippocampus. *PLoS One, 6* (5), 1-8.

Hanushek, E. (1995) Interpreting Recent Research on Schooling in Developing Countries, *World Bank Research Observer, 10* (2), 227-246.

Hyde, J. S. & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *PNAS, 22*, 8801-8807.

Jordan, N. C. & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental disabilities research review, 15:* 60-68.

Klassen, R. (2002a). A question of calibration: A review of the self-efficacy beliefs of students with learning disabilities. *Learning Disability Quarterly, 25*, 88-102.

Klassen, R. (2002b). Writing in early adolescence: A review of the role of self-efficacy beliefs. *Educational Psychology Review, 14* (2), 173-203.

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

Kurtz, P., & Hicks-Coolick, A. (1997). Preparing students with learning disabilities for success in postsecondary education: needs and services. *Social Work in Education, 19*, 31-42.

Lillard, A. & Else-Quest, N. (2006). The early years: evaluating Montessori Education. *Science, 313*, 1893-1894.

Maqsud, M. (1998). Effects of metacognitive instruction on mathematics achievement and attitude towards mathematics of low mathematics achievers. *Educational Research, 2*, 237- 243.

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.

Mazzocco, M. M. M. (2007). Defining and differentiating mathematical learning disabilities and difficulties. In D. B. Berch & M. M. M. Mazzocco (Eds). *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 29-47). Baltimore: Brookes.

Mazzocco, M.M. and Myers, G.F.(2003). Complexities in identifying and defining Mathematics learning disability in the primary school-age years. *Annals of Dyslexia, 53,* 218-253.

McClelland, A. (2000) *Impacts of poverty on children,* Brotherhood Comment, Brotherhood of St. Laurence, Australia.

McLeod, D. B. (1990). Information-processing theories and mathematics learning: The role of affect. *International Journal of Educational Research, 14*, 13-29.

Miles, M. (2011). Children with hydrocephalus and Spina Bifida in East Africa: Can family and community resources improve the odds? *Disability & Society. 17* (6), 643-658.

Ministry of Education (MOE) in P. R. China. (2011). Jiaoyubu guanyu dali jiaqiang zhongxiaoxue jiaoshi peixun gongzuo de yijian [The suggestions on the improve the teacher training for the middle and elementary school from the ministry of education]. Retrieved from http://www.moe.edu.cn/publicfiles/business/htmlfiles/moe/s4559/201101/114220.html.

Najman, J.K., Aird, R., Bor, W., O'Callaghan, M., Williams, G.M., & Shuttlewood, G.J. (2004). The generational transmission of socio-economic inequalities in child cognitive development and emotional health. *Social Science and Medicine, 58,* 1147–1158.

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

O'Sullivan, M. C. (2001). The inset strategies model: an effective inset model for unqualified and underqualified primary teachers in Namibia. *International Journal of Educational Development, 21,* 93-117.

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

Qian, Z. L. (2005). Zhongguo tesu ertong jiaoyu de xianzhuang baogao [Report on the Chinese studies on special education for children]. Retrieved 20 April, 2011, from http://www.edu.cn/20011112/3009417.shtml.

Raudenbush, S. W. & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods. (2nd ed.). Thousand Oaks, CA: Sage.

Reyes, L.H., & Stanic, G.M.A.(1988). Race, sex, socioeconomic status, and mathematics. *Journal for Research in Mathematics Education*, 19(1), 26-43.

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].

Schoenfeld, A. H. (1992). *Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics*. In D. Grouws (Ed.), Handbook for Research on Mathematics Teaching and Learning (pp. 334-370). New York: MacMillan.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*, 207-231.

Shuell, (1990). Phases of meaningful learning. *Review of Educational Research, 60*, 531-548.

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

Sowder, L. (1984). Format variables and learner characteristics in mathematical problem solving. Final technical report. (ERIC document reproduction service, No. ED 238 735).

Teong, S.K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning* 19, 46-55.Trainin, G. and Swanson, H. L.(2005). Cognition, metacognition, and achievement of college students with learning disabilities. *Learning Disabiiity Qiiarteriy, 28*, 261-272.

Vogel, S. A., & Walsh, P. (1987). Gender differences in cognitive abilities in learning disabled females and males. *Annals of Dyslexia, 37*: 142-165.

Voyel, S. A. (2001). Gender differences in intelligence, language, visual-motor abilities, and academic achievement in students with learning disabilities: A review of the literature. *Journal of Learning Disabilities, 23* (1), 44-52.

Walker, S.P, Wachs, T.D, Gardner, J.M, Lozoff, B., Wasserman, GA, Pollitt, E., Carter, JA, and the International Child Development Steering Group. (2007) Child development: risk factors for adverse outcomes in developing countries. *Lancet, 369,* 145-157.

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.*

Chapter 9

Influence of numerical facility ability on the mathematics performance[*]

Abstract

The aim of the present study is to investigate the impact of cultural variables and the mathematics ability levels on the mastery of numerical facility skills. In this light, two questions were asked: (1) is there a difference in numerical facility skills between Flemish and Chinese children and (2) how do children in both countries differ in the mastery of specific numerical facility skills? With regard to the first question, the research results reveal that Chinese students outperform Flemish students. Although the performance difference between Chinese and Flemish students decreases with increasing age, the difference is still significant at the end of elementary school. A MANOVA reveals that low achievers experience comparable learning difficulties in solving multiplication tasks in both countries. With regard to the second question, the results suggest that Flemish and Chinese learners differ in the way they master particular numerical facility skills. In general, students perform better on addition than on subtraction or multiplication tasks, the latter being easier than division or mixed exercises. The test results of Flemish learners reflect consistently smaller differences in the mastery of types of numerical facility skills (e.g., addition vs. subtraction, subtraction vs. mixed operation) as compared to Chinese learners. Finally, in both countries high achievers performed good and stable on all numerical facility tasks, whereas low achievers do not attain a stable achievement level.

1. Introduction

For over more than 100 years, individual differences in numerical, arithmetical and mathematical performance have been included in educational and psychological studies (Brownell, 1928; Geary, 2006; Thorndike & Woodworth, 1901; Thorndike, 1922). About fifty years ago, the first systematic cross-national study on difference in mathematics performance was conducted to explore the role of cultural and social

---

[*] This chapter is based on the submitted paper of Zhao, N.N., Burny, E., Valcke, M., Desoete, A. & Imbo, I. (submitted). The mastery of numerical facility skills in Chinese and Flemish primary school learners. *European Journal of Psychology of Education*

differences (Husen, 1967). Since then, differences in mathematics performance due to different educational systems have been confirmed in ongoing international assessments (TIMSS, PISA). Moreover, a number of worldwide comparative studies provides important information on the role of the educational system that is adopted in a specific culture (Dowker, Bala, & Liloyd, 2008). Recently, an increasing number of studies focuses on both psychological and sociological explanations for differences in mathematical skills across countries (Imbo & LeFevre, 2009, in press; Zhou et al., 2009). Since both Flemish and Chinese students reflect high performance levels in the international mathematics performance indicator studies such as PISA (Prais, 2003), and this irrespective of differences in their curriculum and/or instructional approach, it is interesting to analyse differences in mathematics skills between these two countries. But, as will be explained below, the present study goes beyond a general comparison of mathematical performance.

Previous studies mainly focused on two mathematical domains: the mastery of numerical facility skills and mathematical reasoning (Chein, 1939; Dowker, 2005; Thurstone & Thurstone, 1941). The focus of the present study is on the mastery of numerical facility skills, as these skills are a prerequisite to solve everyday problems. Moreover, numerical facility skills form the basis for dealing with a variety of arithmetical problem-solving tasks. Thus, the present study aims to explore whether the development of numerical facility skills differs in Chinese and Flemish primary school children with different levels of ability. Presenting the development of numerical facility skills as the basics of mathematics, the current study addresses two research questions:

(a) Is there an impact of cultural background (e.g., language, curriculum, teaching practices, etc.) on learners' development of numerical facility skills? And if so, are these differences also apparent in the mastery level of particular grades, or particular sub-groups of learners (i.e., low, average or high achieving)?

(b) Are the differences in the mastery of the five numerical facility skills the same in both countries, in all school grades and for learners with different mathematical abilities? For example, is addition easier than division in both cultural settings? Are the differences in mastery between particular numerical facility skills the same between grades and in learners with different mathematical abilities?

1.1 Differences in the development of numerical facility skills

Numerical facilities are defined as the combination of arithmetical computation and a conceptual understanding of number relationships and arithmetical concepts (Geary, 2006; Thurstone & Thurstone, 1941). In the next section, we focus on the development and mastery of numerical facility skills in Flemish and Chinese primary school children.

1.1.1 The impact of culture in Flanders and China

Cross-national comparative studies of mathematics performance consistently show how East-Asian children outperform Western learners in numerical and arithmetical skills; more specifically in addition and subtraction tasks (De Corte, Greer, & Verschaffel, 1996; Robitaille & Travers, 1992; Geary, Bow-Thomas, Fan, & Siegler, 1993). These differences in mathematics performance have been attributed to a number of cultural variables and processes, such as language (Colome, Laka, & Sebastian-Galles, 2010; Whorf, 1956) and the educational system (Campbell & Xue, 2001; Greay, Bow-Thomas, Fan, & Siegler, 1996). Since previous studies show that cultural differences are not only observed in school-age children but also in adults and preschoolers, this suggests that schooling is only one of the factors explaining differences (Imbo & LeFevre, 2009, in press; Siegler & Mu, 2008), and additional explanations could be related to e.g., linguistic differences (Colome, Laka, & Sebastian-Galles, 2010).

Linguistic differences can play a role in a number of ways. First, it is argued that the transparency of the counting system in a particular language, influences working memory span (Baddeley, 2000; Raghubar, Barnes, & Hecht, 2010), and as such influences performance in mathematics. In this light, several studies show an advantage of about two items in the digit span for Chinese, which helps to explain their higher performance in basic arithmetic (Geary, Bow-Thomas, Liu, & Siegler, 1996; Stigler, Lee, & Stevenson, 1986). Second, previous studies explored the role of specific mathematical language in mathematics performance. According to Seron and Fayol (1994), the way numbers are represented in a language influences the processing of numbers and as such affects students' mathematical performance. The role of this specific mathematical language is reflected in the triple code model of Dehaene and colleagues (Dehaene, 1992; Dehaene & Cohen, 1995). According to this model, there are three internal representations of numbers: an analogue magnitude system, a visual Arabic sketchpad, and a verbal system. Especially this verbal code would be affected by the language used in a cultural setting.

Next to the exploration of language as a determinant of differences in mathematical performance, also school related variables could play a role. Differences in the curriculum structure (sequencing of curriculum topics), textbook design and didactical strategies adopted by teachers are expected to affect the development and mastery of numerical facility skills (Geary, Bow-Thomas, Fan, & Siegler, 1996; Xin, 2007). If there are clear differences in the exposure to basic numerical facility problems during primary school, it can be expected that learners will evolve in different ways (speed, timing) from an explicit procedural strategies usage during the early years to memory retrieval strategies in later years (Koshmider & Ashcraft, 1991; Siegler, 1988). Since school related variables interact with language factors, a preliminary analysis was carried out to analyze the curriculum in Flanders and China in view of the curriculum content, the didactical approaches being adopted and the time weekly spent for mathematics. The results of this preliminary analysis are summarized in appendix 1a, 1b and 1c. The school related differences seem to be limited between Flanders and China. Only one key difference is apparent. The curriculum in China is clearly more demanding during the initial school grades. It is therefore argued that differences in numerical facility skills are – next to expected differences during early years - mainly to be explained through differences in e.g., the language domain.

1.1.2 Different mathematics ability levels

Additionally, it is interesting to explore the differences in mastery and development of numerical facility skills between different ability groups. While the majority of comparative studies solely focused on the development of numerical facility skills in "normal" achieving students (Wang & Lin, 2009), little comparative research has been set up that centres on learners with varying mathematics skills (Desoete, Stock, Schepens, Baeyens, & Roeyers, 2009; Geary & Hoard, 2005). Therefore, the present study will additionally – next to the role of culture - focus on the performance of low, average and high performing children in mathematics. Since previous studies indicate that students with mathematical learning difficulties show deficits in particular numerical skills (i.e., arithmetic procedural skills, number fact retrieval, place value concept, and number sense) and two domain-general processing skills (i.e., working memory and processing speed) (Chan & Ho, 2010), it is expected that students with poor mathematics skills will experience more difficulties in executing basic number tasks that require automation of number retrieval from long-term memory than students with profound mathematics skills (Desoete et al., 2009; Geary & Hoard, 2005). It will be interesting how this interacts with cultural variables.

1.2 Patterns in the mastery and development of the five numerical facility skills

Numerical facility skills comprise four operations: addition, subtraction, multiplication, and division. Several studies have indicated that these tasks differ in terms of their difficulty level (Campbell, 1999; Siegler, 1996). Differences between these tasks have been examined from an educational and psychological perspective. Educational studies stress the distinction between skills acquired through informal learning and skills obtained through formal schooling. For example, Fayol (1990) and Siegler (1996) stated that addition skills are mostly the result of informal learning and develop along with children's acquisition of counting skills at the early beginning of their school career, whereas multiplication skills are the results of formal learning. Another example is provided by Campbell (1997; 1999), building on a mediation effect between two pairs of numerical facility skills: addition and subtraction on the one hand and multiplication and division on the other hand. Campbell (1997; 1999) argued that knowledge of one operation mediates children's performance in the related operation. These research findings can be linked to the sequencing in the attention paid to particular operations in the curriculum. Addition is treated prior to multiplication, divisions, etc.

Psychological studies on the other hand focus on differences between the four operations in terms of retrieval strategies and indicate that retrieval processes are more likely to appear in multiplication and addition tasks as compared to subtraction and division tasks, which might be caused by structural task characteristics and prior experience (Campbell & Xue, 2001). A study of Thevenot, Castel, Fanget and Fayol (2010), for example, showed that only high achieving students use retrieval strategies in solving mental subtraction skills. Studies from a neuropsychological perspective, show that, while addition tasks rely on visuo-spatial processing, multiplication tasks mainly build on verbal processing (Imbo & LeFevre, 2010; Zhou, 2007). From varying perspectives, these studies provide evidence for important differences between the development and execution of the five distinct numerical facility skills.

In the context of the present comparative study, a more fine-grained analysis of the mastery and development of the four operations can be interesting. Next to general differences in mastery of the four operations, a more fine-grained comparison – in relation to country, grade and mathematical ability level – is a step forward as compared to earlier research. When differences are found, they can influence future decisions as to the design of the curriculum and the teaching and learning environment.

2. Methodology

2.1 Sample

A total of 7247 Chinese students and 913 Flemish students, enrolled in grade 3 to grade 6 of primary education, were involved in this study. Flemish students were enrolled in twenty one schools. The 7247 Chinese students were selected from twenty schools in five different provinces and cities. After test administration (see below), all students were asked to provide information on gender and grade. Descriptive statistics for grade and gender are presented in Table 1.

Table 1
Sample characteristics

Country	Gender	Grade 3	Grade 4	Grade 5	Grade 6	Total
Flanders	Female	119	146	100	67	432
	Male	150	155	73	103	481
China	Female	872	852	847	917	3488
	Male	933	979	904	943	3759
Total		2074	2132	1924	2030	8160

2.2 Procedure

All ability measures were obtained in the setting of the classroom of the pupils and following a standard protocol. All participants were first administered the numerical facility tests, followed by the mathematical ability test. The test administration lasted about 5 minutes for the numerical facility test and 40-50 minutes for the mathematics abilities test.

2.3 Numerical facility skills

The standardized tests of numerical facility tasks (Tempo Test Rekenen, TTR; De Vos, 1992) was administered to both Flemish and Chinese primary school children. The TTR is a mental calculation test including items in relation to all four operations (i.e., addition, subtraction, multiplication, and division). There are 40 additions (e.g., 2 + 3 = ?), 40 subtractions (e.g., 5 – 2 = ?), 40 multiplications (e.g., 2 x 9 = ?), 40 divisions (e.g., 12 : 2 = ?), followed by 40 problems where all operations are intermixed. Each series of 40 problems is vertically presented in a column. For each series of problems, the students are given one minute to solve as many problems as possible. In summary, the TTR consists of 200 arithmetic problems, of which a maximal number has to be solved within five minutes. The test administrator controls the timing for each type of operation. The TTR test is in line with the Flemish and Chinese curriculum that stresses knowledge of simple arithmetic, including knowledge of addition, subtraction, multiplication and division within the multiplication tables (See Appendix 1 a).

2.4 Determining mathematical ability levels

In view of differentiating between different mathematical ability levels, learners in both countries were screened on the basis of a general mathematical ability test that is aligned with the local curriculum. In Flanders, the students were assessed with the Kortrijk Arithmetic Test–Revised (Kortrijkse Rekentest Revision, KRT-R; Baudonck et al., 2006). The KRT-R (Baudonck et al., 2006) is a standardized test to measure arithmetical achievement that has an established position in Flemish education (e.g., Desoete & Grégoire, 2007; Desoete, Roeyers, & De Clercq, 2004). In China, a new standardized test, covering the new curriculum syllabus from 2001 was administered to all students. This test was calibrated for all grade levels and considering the full range of math abilities in the Chinese context. The test reflects a high empirical reliability; for alfa values range between.95 and .93 .

Both tests were used to distinguish between three mathematics performance levels: a low performing group (25%), an average performing group (50%), and a high performing group of learners (25%).

2.5 Data Analysis

In a first step, univariate analyses of variance (ANOVA) were applied to assess differences in numerical facility skills. A second step involved the application of a 4(Grade) x 2 (country) x 3 (skills) MANOVA on five numerical facility skills to evaluate differences across the four operations (addition, subtraction, multiplication, division) and mixed-operation skills between Flemish and Chinese students, considering the different levels of mathematical ability. In a third step, for each set of ability tests, countries and grade differences were assessed by means of a 4(Grade) x 2 (country) x 3 (skills) MANOVA on ten difference between each two numerical facilities. Next, if the MANOVAs performed in step 2 or 3 were significant, 2 (country) X 4 (grade) ANOVAs were carried out, in order to study the nature of the differences in the mastery of the five numerical facility skills. Simple mean values were contrasted by means of the HSD procedure. A significance value of $p < .001$ level was stated in view of all analyses.

3. Results

3.1 Descriptive results

Descriptive results about the numerical facility scores across countries and grades are represented in Figure 1. There are no significant differences between scores considering the variable gender (entire sample, or within countries). The Chinese students scored consistently higher than Flemish students for each of the five numerical facility skills. *T*-test results comparing the scores for the two countries, show a significant difference in scores for addition (t=-52.48, p<.001), subtraction (t=-39.88, p<.001), multiplication (t=-59.58, p<.001), division (t=-34.03, p<.001) and mixed operation (t=-41.61, p<.001). Also within each grade, t-test results indicate that Chinese learners attain significantly higher scores than Flemish learners for each of the five numerical facility skills. This is also the case when comparing students with different mathematical skills. Additional t-test analysis results can be obtained from the authors.

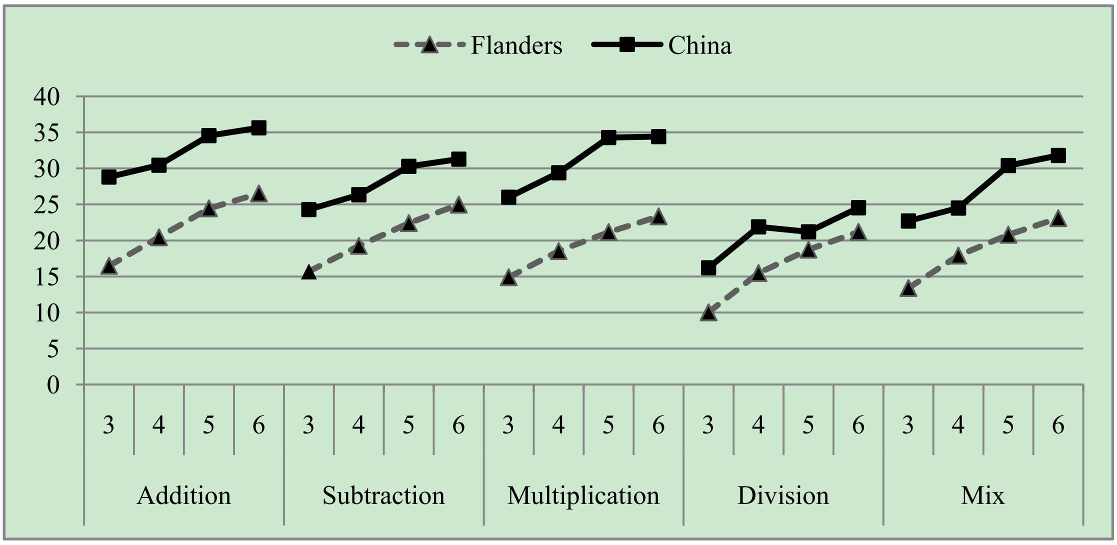


Figure 1
Number of correctly solved items for each specific numerical facility, at each grade level in Flanders and China.

3.2 Comparative analysis of the mastery and development of numerical facility skills

3.2.1 Numerical facility skills in relation to country, grade and mathematical ability level

In a first step, an univariate analysis of variance (ANOVA) was carried out on the sum scores of the five mental calculation tasks, with mathematics level and country as fixed factors. The main effects of mathematics ability level and country are significant ($F_{(2,\ 8157)}$=85.10, p<.001, η^2 =.02; $F_{(1,\ 8158)}$=1068.65, p<.001, η^2 =.12). Also, the interaction between country and mathematics ability level is significant ($F_{(2,\ 8158)}$=3.65, p<.05, η^2 =.001). The differences between Flanders and China were larger for high-skill than for low-skill students. The effect size related to differences between countries is small, while the effect size related to mathematics ability level is of medium size (Green, Salkind, & Akey, 2000).

3.2.2 In-depth analysis of the mastery of the five particular numerical facility skills

In a second step, the five numerical facility skills are studied separately. Table 2 shows how mastery of all five skills is strongly correlated in both countries (see Table 2). Yet, there are some interesting differences in the correlation parameters, both in Flanders and China.

Table 2
Correlations between the mastery of the five numerical facility

	Addition	Subtraction	Multiplication	Division	Mixed
Addition		.77**	.70**	.66**	.78**
Subtraction	.75**		.72**	.72**	.81**
Multiplication	.61**	.78**		.82**	.82**
Division	.60**	.76**	.79**		.83**
Mixed	.56**	.73**	.78**	.85**	

Note: Correlation for Flemish students are presented above the diagonal, and correlation for Chinese student are presented below the diagonal.
**p*<.05 ** *p*<.01

First, a multivariate analysis of variance (MANOVA) was carried out in relation to the five numerical facility scores. Significant differences between countries, grades and mathematics abilities are found. The main effects of country ($F_{(5,\ 8154)}$=476.14, p<.01, η^2 =.23), grade ($F_{(15,\ 8144)}$=66.75, p<.01, η^2 =.04) and mathematics abilities ($F_{(10,\ 8149)}$=77.48, p<.01, η^2 =.05) are significant (Table 3). Also, the interactions between country and grade ($F_{(15,\ 8144)}$=10.95, p<.01, η^2 =.01), and between grade and ability ($F_{(30,\ 8129)}$=9.98, p<.01, η^2 =.01) are significant.[1] Different mastery patterns are observed considering the grades in both countries and levels in mathematics abilities within in each grade. The ANOVA results reveal that the countries, grades and mathematics abilities and country X Grade interaction are significant for addition, subtraction, multiplication, division and mixed operations (p<.01). For country X grade, China-Flanders differences for addition, multiplication decrease with grade; China-Flanders difference for division increase with age; and, China-Flanders differences for multiplication and mixed operation between Flanders and China fluctuate up and down. The interaction of country X grade is significant for addition,

[1] A MANOVA analysis confirmed that no main effect of gender or any interaction effects were found (Country X Gender, Math X Gender, Grade X Gender for the addition, subtraction, multiplication, division and mixed calculation). Also, the MANOVA reveals that there were no interaction effects of mathematics ability and country or grade for the five numerical facility skills.

subtraction, division and mixed operation (*p*<.05), and marginally significant for multiplication *p*=..06). For ability X grade, the difference between average level and lower level abilities for became smaller with grade for addition, subtraction, multiplication and mixed except division.

3.2.3 Interaction effects of country, grade and specific mathematics ability level

In a third step, three MANOVA analyses with the country and grade as the fixed factors were carried out considering the five numerical facility scores of low, average, and high achiever. The latter is an interesting perspective since it helps to detect whether countries attain different achievement scores for these three particular groups of students. Differences could suggest that curriculum or didactical approaches in a particular country could cater to a better extent for the needs of particular ability group.

The MANOVA results show that the main effect of country and grade, and the interaction effect of country and grade, are significant for the group of

- low achievers ($F_{(5, 2018)}$=99.09, p<.01, η^2 =.20; $F_{(15, 2008)}$=, p<.01, η^2 =.06; $F_{(15, 2008)}$=3.162, p<.01, η^2 =.01),
- average achievers ($F_{(5, 4167)}$=272.48, p<.01, η^2 =.25; $F_{(15, 4167)}$=35.71, p<.01, η^2 =.04; $F_{(15, 4157)}$=6.66, p<.01, η^2 =.01), and,
- high achievers ($F_{(5, 1957)}$=476.14, p<.01, η^2 =.23; $F_{(15, 1947)}$=476.14, p<.01, η^2 =.23; $F_{(15, 1947)}$=476.14, p<.01, η^2 =.23).

For the group of low achievers, the ANOVA results show that the interaction of country and grade is only significant for multiplication and division tasks. For the group of low achievers, the difference between Chinese-Flemish students increased across grades for multiplication and division only. Both the Flemish and Chinese students in higher grades perform better in multiplication and division tasks as compared to addition, subtraction and mixed problems. This indicates that for low achievers, numerical facility skills in relation to multiplication and division further develop during schooling. This is consistent with previous studies pointing out that multiplication is related to retrieval in working memory (Imbo & Vandierendonck, 2007b).

For the group of average achievers, the ANOVA results reveal similar results as for the entire sample: the interaction of country and grade is significant for addition, division and mixed calculation. The difference between Chinese-Flemish students decreases across grades for addition and increase for division while the difference is going up and down for the mixed calculation. Both the Flemish and Chinese students' skills in relation to addition, division and mixed calculation improve during schooling, while this is lesser the case for multiplication and subtraction.

For the group of high achievers, the analyses reveal that the interaction of country and grade is significant in relation to addition and subtraction problems. The difference between Chinese-Flemish students decreased across grades for addition and subtraction only.In this ability group, both Flemish and Chinese learners obtain better results in addition and subtraction tasks during subsequent school grades.

Table 3
MANOVA results reflecting differences in numerical facility skills, considering country, grade and mathematics ability level

		Low			Average			High		
Source	Dependent variable	F	Sig.	η^2	F	Sig.	η^2	F	Sig.	η^2
Country	Add	303.7	.00**	.13	784.02	.00**	.16	422.27	.00**	.18
	Sub	102.3	.00**	.05	280.95	.00**	.06	210.53	.00**	.10
	Mul	269.55	.00**	.12	760.05	.00**	.15	395.27	.00**	.17
	Div	105.44	.00**	.05	280.28	.00**	.06	164.62	.00**	.08
	Mix	132.85	.00**	.06	371.85	.00**	.08	165.25	.00**	.08
Grade	Add	71.34	.00**	.10	100.42	.00**	.07	35.02	.00**	.05
	Sub	45.73	.00**	.06	67.19	.00**	.05	24.85	.00**	.04
	Mul	55.39	.00**	.08	74.25	.00**	.05	28.34	.00**	.04
	Div	93.95	.00**	.12	142.37	.00**	.09	51.65	.00**	.07
	Mix	64.09	.00**	.09	86.24	.00**	.06	23.81	.00**	.04
Country X Grade	Add	0.25	.862	.00	4.29	.005**	.00	9.07	.00**	.01
	Sub	0.71	.544	.00	2.04	.107	.00	3.94	.008*	.01
	Mul	2.77	.040*	.00	2.03	.107	.00	.83	.477	.00
	Div	6.26	.000**	.01	4.68	.003**	.00	.79	.500	.00
	Mix	2.13	.094	.00	4.48	.004**	.00	2.62	.050	.00

Note: * p<.05 ** p<.001

Add: Addition; Sub: subtraction; Mul: multiplication; Div: division; Mix: Mixed operation. Low: low achiever; Ave.: average achiever; High: high achiever.

3.3 Differences in the interrelated mastery of the five numerical facility skills

In this second section, we explore the potential differences in the complex interplay between the five numerical facility skills (See Table 4).

Table 4

Testing differences in mastery of different types of numerical facility skills in different countries and different grades

	Flanders			China			Grade3			Grade4			Grade5			Grade6		
	M	SD	t(912)	M	SD	t(7246)	M	SD	t(2073)	M	SD	t(2131)	M	SD	t(1923)	M	SD	t(2029)
+&-	1.30	3.75	10.46	4.24	6.61	54.59	4.03	7.05	26.06	3.69	5.73	29.72	3.8	5.12	32.58	4.12	7.45	24.89
+&x	2.31	4.24	16.45	1.34	8.21	13.85	2.64	8.43	14.28	1.16	8.06	6.62	0.53	6.01	3.89	1.39	8.47	7.39
+&:	5.61	5.19	32.66	7.82	9.73	68.39	11.80	9.30	57.74	8.04	9.44	39.33	6.22	8.01	34.04	4.04	8.75	20.79
+&M	3.05	3.74	24.64	5.02	9.2	46.42	5.73	9.04	28.87	5.44	9.34	26.88	4.14	7.5	24.18	3.8	8.89	19.26
-&x	1.01	3.82	8.01	-2.90	6.61	-37.39	-1.39	6.27	-10.10	-2.53	7.52	-15.56	-3.27	6.06	-23.63	-2.73	5.68	-21.63
- &:	4.31	4.65	28.01	3.58	7.98	38.18	7.76	7.52	47.03	4.35	8.11	24.77	2.42	6.93	15.32	-0.08	5.65	-0.63
-&M	1.75	3.35	15.81	0.78	7.62	8.73	1.70	7.33	10.55	1.75	8.39	9.61	0.34	6.92	2.14	-0.31	5.94	-2.37
x&:	3.30	3.84	26.00	6.48	7.41	74.45	9.15	7.35	56.68	6.89	7.39	43.01	5.69	6.66	37.43	2.65	5.46	21.86
x&M	.74	3.19	7.03	3.68	6.84	45.84	3.09	6.92	20.32	4.28	7.08	27.93	3.6	6.17	25.61	2.41	5.96	18.25
:&M	-2.56	3.78	-20.43	-2.8	6.38	-37.35	-6.06	6.80	-40.61	-2.61	5.75	-20.93	-2.08	5.35	-17.08	-0.23	4.97	-2.13

Note:

[a] + refers to addition; - refers to subtraction; x refers to multiplication; : refers to division;

[b] All of the paired sample *t*-tests are significant at the .001 level.

A paired-samples *t*-test was applied to compare - pairwise - the mastery of particular the five numerical facility skills in Flemish and Chinese learners (See Table 4). The results show how particular numerical facility skills are easier or more difficult than others. The differences found, are mostly in line with the sequencing order in which topics are introduced to learners in the curricula, except for the pair of subtraction-multiplication and the pair of division-mixed operation. In general, both Chinese and Flemish learners achieved significantly higher results in addition as compared to other calculation tasks. Learners attained significantly higher results in subtraction tasks as compared to division and mixed calculation tasks. And learners performed significantly better in multiplication tasks than division and mixed calculation tasks.

Next, a multivariate analysis of variance (MANOVA) was performed with country, grade and mathematics ability as the independent variables and the differences between pairs of numerical facility skills as dependent variables (See Table 5). Results reveal that the main effect of country ($F_{(10, 8149)}$=114.88, p<.01, η^2 =.05), grade ($F_{(30, 8128)}$=27.87, p<.01, η^2 =.01) and mathematics abilities ($F_{(20, 8139)}$=39.93, p<.01, η^2 =.02), and the interaction between country and grade ($F_{(40, 8118)}$=12.55, p<.01, η^2 =.006) and between grade and mathematics ability ($F_{(60, 8098)}$=6.81, p<.01, η^2 =.005) are significant.

Table 5
Differences in the mastery of contrasting types of numerical facility skills

Dependent variable[a]	**C**			**G**			**Ma**			**C*G**			**G*Ma**		
	F	Sig.	η^2	F	Sig.	η^2	F	Sig.	η^2	F	Sig.	η^2	F	Sig.	η^2
+&-	153.04	.00	.018	1.06	.36	.000	49.11	.00	.013	2.61	.05	.001	3.2	.00	.002
+&x	16.96	.00	.002	2.07	.10	.000	1.18	.31	.000	10.92	.00	.004	4.22	.00	.003
+&:	47.08	.00	.006	43.24	.00	.016	56.75	.00	.013	27.59	.00	.010	11.13	.00	.008
+&M	34.27	.00	.004	1.93	.12	.001	11.5	.00	.002	6.49	.00	.002	12.24	.00	.009
-&x	308.91	.00	.037	3.81	.01	.001	53.96	.00	.013	6.65	.00	.002	6.4	.00	.005
- &:	6.58	.01	.001	75.05	.00	.027	10.04	.00	.002	3.36	.00	.006	13.82	.00	.010
-&M	14.49	.00	.002	4.47	.00	.002	6.94	.00	.001	5.42	.00	.002	13.66	.00	.010
x&:	196.85	.00	.024	69.87	.00	.025	102.74	.00	.024	15.26	.00	.009	6.46	.00	.005
x&M	164.69	.00	.020	4.58	.00	.002	20.06	.00	.004	6.06	.00	.002	6.51	.00	.005
:&M	2.62	.11	.000	65.23	.00	.023	43.74	.00	.011	24.94	.00	.009	5.41	.00	.004

[a]+ refers to addition; - refers to subtraction; x refers to multiplication; : refers to division; M refers to mixed operation. C refers to country; G refers to grade; Ma refers to mathematics ability level.

In table 5, it is interesting to observe – whereas most differences in mastery seem to be equal in China and Flanders - certain differences between the mastery of particular pairs of operations. Chinese learners achieve 2.90 points higher in multiplication versus subtraction, while Flemish learners score 1.01 points higher on subtraction versus multiplication. Multiplication tasks seem to have a different position in the context of both curricula. This finding is in line with the previous studies (LeFevre & Liu, 1997).

4. Discussion, Conclusions and Limitations

The present study provides evidence about cross-cultural differences in the mastery of basic numerical facility skills.

4.1 Difference in five numerical abilities

4.1.1 Is there a differences in children's numerical facility skills across countries?

The results of the present study indicate that Chinese primary school students outperform Flemish students in relation to all numerical facility skills. This finding is in line with previous research that indicates that East Asian students outperform their Western peers (Imbo & LeFevre, 2009, in press). Since the didactical approaches in Flanders and China are mainly comparable, (See Appendix 1a, b and c), it is argued that differences in numerical facility skills can be explained on the base of differences in the language used in both countries. Previous studies stressed the importance of the linguistic transparency in learning the counting system (Dowker, Bala, & Lioyd, 2008), it is therefore acceptable that Chinese students perform better in mathematics because the Chinese language for reading numbers is consistent with Arabic number writing. For example, in Chinese, the number 72 is read as qi shi er (seven ten two); in which *qi* represents the number 7, *er* refers to the number two and *shi* refers to the place value ten. In Dutch, on the other hand, the reading of numbers is inconsistent with the Arabic number representation: 72 is read as *tweeënzeventig* (two and seventy). This is more demanding for the working memory system, that plays an important role in mental calculation (DeStefano & LeFevre, 2004; Lau & Hoosain, 2009). It is thus argued that the consistency between Arabic number representation and the Chinese language for reading numbers results in higher mental calculation performance in Chinese students.

4.1.2 Is performance in numerical facility skills influenced by curriculum arrangement and instructional strategies?

The results of the current study indicate that the curriculum content partly has an impact on learner performance. The analyses show a significant effect of grade on children's performance in addition, subtraction and division. But, these differences between Chinese and Flemish learners decrease in higher grades. The difference in performance between Chinese and Flemish learners is easily explained by the curriculum content: in China, first grade children already become familiar with mental calculation tasks up until 100, which is a higher requirement as compared to the Flemish curriculum. But from grade three on, this difference is gone and therefore differences in performance on addition, subtraction and division tasks disappear.

An interesting result can be found with regard to multiplication tasks: no significant interaction effect of country and grade is observed. In Flemish and Chinese students of grade 5 and grade 6. The result indicates that the difference between Flemish and Chinese students does not decrease across grade while the differences on other operation between Flemish and Chinese student do decrease for the other operations. Though Zhou et al., (2006) state that two types of strategies can be used to complete simple arithmetical tasks: procedural strategies and rote verbal memory strategies, the current findings suggest that both Chinese and Flemish learners rather apply rote memory strategies to memorize multiplication facts. This is confirmed by the curriculum analysis (appendix 1b) that show how in both countries, rote verbal strategies are applied to teach the multiplication table.

4.1.3 Does the cultural dimension play a different role considering learners of different ability levels?

In the current study, we distinguished between low, average and high performing learners. The question was whether learners with different ability levels, perform differently in China and Flanders.

In low achievers, the interaction effect of country X grade was not significant with regard to addition, subtraction and mixed operation tasks. But, in both countries, low achievers reflect a different mastery level in multiplication and division tasks. Lower achievers seem to meet consistently the ceiling in their ability when dealing with the more complex operations. This is consistent with findings of previous studies that indicate that children with learning difficulties easily meet difficulties in working memory that will affect performing operations (Geary & Hoard, 2002).

For the group of average achievers, the interaction of country and grade is significant for addition, division and mixed calculation. Both Flemish and Chinese learners' skills in relation to addition, division and mixed calculation improve during schooling, while this is lesser the case for multiplication and subtraction.

For the group of high achievers, the interaction of country and grade is significant in relation to addition and subtraction problems. In this ability group, only Flemish learners obtain better results in addition and subtraction tasks during subsequent school grades. This can be explained by the Chinese learners who already attain in the early grades a ceiling level that hardly changes during subsequent grades.

Previous studies showed that the mental representations to solve the elementary number combinations are different for high and low achieving learners (Desoete et al., 2009; Pitta & Gray, 1997). Low achievers' mental representations are strongly associated with procedural aspect of numerical processing while the high achievers' mental representations rather focus on abstractions (Pitta & Gray, 1997).

4.2 Mastery level for each two of the five numerical facility abilities

4.2.1 The mastery of the five numerical facilities in China and Flanders

The results shows that learners perform best addition tasks, followed by subtraction and multiplication tasks. Children are significantly less accurate in solving division tasks and series of tasks with mixed operations. This order in masterly levels obviously reflects the curricular arrangement (see appendix 1a). In both countries children are first taught addition skills, to continue with subtraction and multiplication, followed by division skills.

The sequence in the development of the five numerical facility skills can also be influenced by the way they differ in posing cognitive demands. In the past 20 years, studies from a cognitive arithmetic perspective have revealed that performance in these different operations involves the use of a variety of strategies (Fayol, 1990; Siegler, 1996): addition skills build upon children's counting strategies, whereas multiplication is acquired by the memorization of associations between pairs of digits and answers (Roussel, Fayol, & Barrouillet, 2002). Moreover, brain studies reveal that addition skills are closely linked to visuo-spatial processing while multiplication skills are connected to verbal processing (Zhou, 2007).

If the curriculum is the crucial factor, then we expect the differences between Chinese and Flemish to decrease with increasing age. However, if language is the crucial factor, then we expect no age-related decrease in the differences between Chinese and Flemish. And the results show a decrease between Chinese and Flemish students for addition, subtraction, and division but not for multiplication. Because multiplication is more language-related than the other operations! Thus, conclusion: curriculum explains differences in addition, subtraction and division but language explains differences in multiplication.

4.2.2 Cultural differences and the relationship between the numerical facilities

An interesting result appears when studying the performance in subtraction and multiplication tasks in Chinese and Flemish children: whereas Chinese children perform better in multiplication tasks than in subtraction tasks, the opposite was found in Flemish children. This finding confirms the hypothesis that language plays an important part in explaining the differences in mathematics performance between both populations. Thanks to their more straightforward linguistic system, Chinese children seem to apply retrieval strategies quicker than Flemish students (Imbo & Lefevre, 2009). For example, it is easier for the Chinese students to remember “nine multiplied by nine is eighty-one” (jiu jiu ba shi yi) while it is a little difficult for Flemish students to remember the result of nine multiplied by nine is one-and-eighty (“eenentachtig”).

When comparing the mastery of pairs of numerical facility skills, interesting differences between both countries are found. First, the differences between addition-subtraction, addition-division, addition-mixed operation, subtraction-mixed operation and multiplication-mixed operation are smaller in Flemish students than in Chinese students. And the differences between addition-multiplication, subtraction-multiplication, subtraction-division, multiplication-division, division-mixed are in an opposite direction, except for the pair division-mixed operation.

Additionally, the results show that division tasks are easier than series of mixed operation tasks for children in both countries. However, Chinese children seem to have an advantage in multiplication and addition tasks, more than in all of the other operations. This finding is in line with previous studies that indicate that Chinese students rely almost completely on direct retrieval when solving simple addition (Geary, 1996) and multiplication tasks (LeFevre & Liu, 1997).

4.3 Limitations, implications and directions for future research

The results of the present study go beyond the available evidenced about the superior mastery of mathematics by Asian learners. A more fine-grained analysis of differences has been presented and discussed. But the present study reflects a number of limitations that should be considered in future research.

First, the future research should go beyond the “product” level of mathematics learning and map the differences in actual processes and strategies being adopted by learners in the different cultural settings.

Second, the “cultural” dimension was approached in a rather general way. Next to a focus on language differences, only a number of curriculum related differences were considered (curriculum content, basic didactical approaches, timing, ...). Other factors should be considered in an explicit way to study cultural differences; e.g., the actual teaching approach, homework, shadow education, the impact of parents, etc.

Thirdly, both Flanders and China are highly ranked in the PISA list. It could be interesting to involve other countries in the cross-cultural studies that reflect critical performance differences.

Despite the limitations, mentioned above, the present study also introduces a number of new directions. Among others, a focus on learners with different mathematics abilities could be the base for more in-depth cross-cultural studies that – at the same time – incorporate the study of present didactical strategies to foster the needs of these different groups and cater for the development of the five different numerical facility skills.

References

Ashcraft, M. H., Donley, R. D., Halas, M. A., & Vakali, M. (1992). Working memory, automaticity, and problem difficulty. In J. I. D. Campbell (Eds), The nature and origins of mathematics skills. (pp. 301-329.). Amsterdam: Elsevier.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4,* 417-423.

Baudonck, M., Debusschere, A., Dewulf, B., Samyn, F., Vercaemst, V., & Desoete, A. (2006). *De Kortrijkse Rekentest Revision KRT-R* [The Kortrijk Arithmetic Test Revised, KRT-R]. Kortrijk, Belgium: CAR Overleie.

Brownell, W. A. (1928). *The development of children's number ideas in the primary grades*. Chicago: University of Chicago.

Campbell, J. I. D. (1997). On the relation between skilled performance of simple division and multiplication. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 1140-1159.

Campbell, J. I. D. (1999). Division by multiplication. *Memory & Cognition, 27*, 791-802.

Campbell, J. I. D., & Xue, Q. L. (2001). Cognitive arithmetic across cultures. *Journal of Experimental psychology: General. 130* (2), 299-315.

Chan, B. M., & Ho., C. S-H. (2010). The cognitive profile of Chinese children with mathematics difficulties. *Journal of Experimental Child Psychology, 107*, 260-279.

Chein, I. (1939). Factors in mental organization. *Psychological Record, 3*, 71–94.

Colome, A., Laka, I., & Sebastian-Galles, N. (2010). Language effects in addition: How you say it counts. *The Quarterly Journal of Experimental Psychology, 63*, 965-983.

De Corte, E., Greer, B., & Verschaffel, L. (1996). *Psychology of mathematics teaching and learning*. In D. C. Berliner, & R. Calfee (Eds). Handbook of educational psychology (pp. 491-549). New York: MacMillan.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1*, 83-120.

Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation : double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex, 33*, 219-250.

Dehaene, S. (1992). Varieties of numerical abilities, *Cognition, 44,* 1-42.

Desoete, A., & Grégoire, J. (2007). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences, 16,* 351–367.

Desoete, A., Roeyers, H., & De Clercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities, 37,* 50-61.

Desoete, A., Stock, P., Schepens, A., Baeyens, & D. Roeyers, H., (2009). Classification, seriation, and counting in grade 1, 2 and 3 as two-year longitudinal predictors for low achieving in numerical facility and arithmetical achievement? *Journal of Psychoeducational Assessment*, 26, 252-264.

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353-386.

Dowker, A. (2005). *Individual differences in arithmetic: implications for psychology, neuroscience and education*. Hove, UK: Psychology Press.

Dowker, A., Bala, S., & Liloyd, D. (2008). Linguistic Influences on Mathematical Development: How Important Is the Transparency of the Counting System? *Philosophical Psychology, 21*, 523-538.

Fayol, M. (1990). *L'enfant et le nombre [Children and numbers]*. Neuchatel, Switzerland: Delachaux & Niestle.

Geary, D. C. (1996). The problem-size effect in mental addition: developmental and cross-national trends. *Mathematical Cognition, 2*, 63-93.

Geary, D. C. (2006). Development of mathematical understanding. In D. Kuhl & R. S. Siegler (Vol. Eds.), *Cognition, perception, and language, Vol 2* (pp. 777–810). W. Damon (Gen. Ed.), *Handbook of child psychology (6 th Ed.)*. New York: John Wiley & Sons.

Geary, D. C., & Hoard, M. K. (2005). Learning disabilities in arithmetic and mathematics: theoretical and empirical perspectives. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 253-268). New York: Psychology Press.

Geary, D. C., & Hoard, M. K. (2002). Learning disabilities in basic mathematics: deficits in memory and cognition. In J. M. Royer (Ed). Mathematics cognition (pp.93-115). Greenwich, CT: information Age.

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development, 67,* 2022-2044.

Geary, D. C., Salthouse, T. A., Chen, G.P., & Fan, L. (1996). Are East Asian Versus American Differences in Arithmetical Ability a Recent Phenomenon? *Developmental Psychology, 32* , 254-262.

Greay, D. C., Bow-Thomas, C. C., Fan, L., & Siegler, R. S. (1993). Even before formal instruction, Chinese children outperform American children in mental addition. *Cognitive Development, 8*, 517-529.

Greay, D. C., Bow-Thomas, C. C., Fan, L., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influences of age, language, and schooling. *Child Development, 67*, 2022-2044.

Green, S., Salkind, N., & Akey, T. (2000). *Using SPSS for Windows* (2nd Ed). Upper Saddle River, NJ: Prentice Hall, USA.

Husen, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries* (Vols. I & II). New York: Wiley.

Imbo, I., & LeFevre, J-A. (2009). Cultural differences in complex addition: Efficient Chinese versus adaptive Belgians and Canadians. *Journal of Experimental Psychology, 35* (6), 1465-1476.

Imbo, I., & LeFevre, J.-A. (in press) Cultural differences in strategic behavior: A study in computational estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Imbo, I., & Vandierendonck, A. (2007a). The development of strategy use in elementary school children: working memory and individual differences. *Journal of Experimental Child Psychology. 96*, 284-309.

Imbo, I., & Vandierendonck, A. (2007b). Do multiplication and division strategies rely on executive and phonological working memory resources? *Memory & Cognition, 35*, 1759-1771.

Koshmider, J. W., & Ashcraft, M. H. (1991). The development of children's mental multiplication skills. *Journal of Experimental Child Psychology, 51,* 53-89.

Lau C.W., & Hoosain R.(2009). Working memory and language difference in sound duration: A comparison of mental arithmetic in Chinese, Japanese, and English. *Psychologia, 42*, 139-144.

Lau C.W., Hoosain R.(2009Siegler, R. S. & Mu, Y. (2008).

LeFevre, J-A, & Liu, J. (1997). The role of experience in numerical skill: Multiplication Performance in Adult from Canada and China. *Mathematical Cognition, 3* , 31-62.

Pitta, D., &Gray, E. (1997). In the Mind. What can imagery tell us about success and failure in arithmetic?, In G.A. Makrides (ed.), Proceedings of the First Mediterranean Conference on Mathematics, Nicosia, Cyprus, pp. 29-41.

Prais, S. J. (2003). Cautions on OECD's recent educational survey (PISA). *Oxford Review of Education, 29*, 139-163.

Raghubar, K. P., Barnes, M. A.,& Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110-122.

Robitaille, D. F., & Travers, K. J. (1992). *International studies of mathematics II: Contexts and outcomes of school mathematics*. Oxford: Pergamon Press.

Roussel, J., Fayol, M., & Barrouillet, P. (2002). Procedural vs. direct retrieval strategies in arithmetic: A comparison between additive and multiplicative problem solving. *European Journal of Cognitive Psychology*, *14*, 61–104.

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology, 12,* 281-300.

Siegler, R. S., & Mu, Y. (2008). Chinese children Excel on novel mathematics problems even before elementary school. *Psychological Science, 19* , 759-763.

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skills. *Journal of Experimental Psychology: General*, 106, 250-264.

Siegler, R.S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Stigler, J. W. Lee, S. Y., & Stevenson, H. W. (1986). Digit memory in Chinese and English: Evidence for a temporally limited store. *Cognition, 23*, 1-20.

Thevenot, C., Castel, C., Fanget, M., & Fayol, M. (2010). Mental subtraction in high and lower skilled arithmetic problem solvers: verbal report versus operand-recognition paradigms. Journal of Experimental psychology: learning, memory and cognition, 36 (5), 1242-1255.

Thorndike, E. L. (1922). *The psychology of arithmetic*. New York: Macmillan.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: Vol. 2. Estimation of magnitudes. *Psychological Review, 8*, 384–394.

Thurstone, L. L., & Thurstone, T. G. (1941). *Factorial studies of intelligence. Psychometric Monographs (2).* Chicago: University of Chicago Press.

Wang, J., & Lin, E.(2009). A meta-analysis of comparative studies on Chinese and US students' mathematics performance: Implication for mathematics education reform and research. *Educational Research Review*, *4*, 177-195.

Whorf, B. L. (1956). *Language*, *Thought and Reality.* In J. B. Carroll (Ed.), Language, thought, and reality: Selected writings of Benjamin Lee Whorf. (pp.213-214) Cambridge, MA: Technology Press of Massachusetts Institute of Technology.

Xin, Y. P. (2007). Word Problem Solving Tasks in Textbooks and Their Relation to Student Performance, *The Journal of Educational Research, 100* , 347-359.

Zhang, H., & Zhou, Y. (2003). The teaching of mathematics in Chinese elementary schools. *International Journal of Psychology, 38*, 286-298.

Zhao, N. N., Valcke, M., Verhaeghe, J., Desoete, A. & Xu, K. (resubmitted). A standardized instrument to diagnose mathematics performance in Chinese primary education: Application of item response theory. *Asia Pacific Journal of Education.*

Zhou, X. L. , Chen, C. S., Zang, Y. F. , Dong, Q., Chen, C. H., Qiao, S. B., & Gong, Q. Y. (2007). Dissociated brain organization for single-digit addition and multiplication. *Neuroimage, 35*, 871-880.

Zhou, X. L., Chen, C. S. Qi, D., Zhang, H. C., Zhou, R. L., Zhao, H., Chen, C. H., Qiao, S. B., Jiang, T., & Guo, Y. (2006). Event-related potentials of single-digit addition, subtraction and multiplication*.* *Neuropsychologia,44*, 2500-2507.

Zhou, X. L., Chen, C. S. Qiao, S.B., Chen, C. H., Chen, L. Lu, N., & Dong, Q., (2009). Event-related potentials for simple arithmetic in Arabic digits and Chinese number words: a study of the mental representation of arithmetic facts through notation and operation effects. *Brain Research,1302*, 212-224.

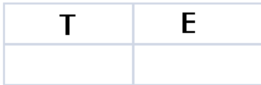
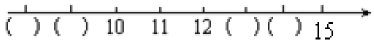
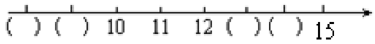
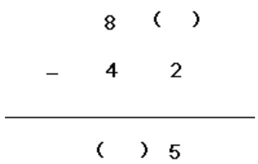
Appendix 1a

Comparison of the mathematics curriculum of Flanders and Mainland China

Domain	Number		Addition		Subtraction		Multiplication		Division	
Country	Flanders	China	Flanders	China	Flanders	China	Flanders	China	Flanders	China
Grade 1	1-20	<100	Sum<10	sum<100	minuend <10	minuend<20	-	-	-	-
Grade 2	1-100	<10000	Sum<100	sum <10000	minuend <20	minuend <10000	Multiplication table	Multiplication table	-	Multiplication table
Grade 3	1-1000	Fraction+ Decimal	Sum <1000 Decimals	sum <10000	minuend <100 decimals	minuend <10000	Multiplying by 10 and 100 1 digit * 2 digit 1 digit * 3 digit	Multi-digit * 1-digit 2-digit * 2-digit	Multiplication table; Division with remainder (divider <10)	Division With Remainder; Divisor is 1-digit
Grade 4	1-100 000	100 000 000	Sum <100 000		minuend <1000		Multiplying by 5 and 50	3-digit * 2-digit	Division with remainder (all)	Divisor is 2-digit
Grade 5	1-10 000 000	Multiple and factor	Sum <1 000 000 000 fractions	Decimal	<100 000 fractions	Decimal	Multiplying by 1 000 and 10 000 Fraction * whole number Decimal * whole number	Decimal	Fraction : whole number Decimal : whole number	Decimal
Grade 6	1-100 000 000	Negative number	Sum <1 000 000 000	Fraction	<1 000 000 000	Fraction	Fraction * fraction Decimals	Fraction	Fraction : fraction Decimal : decimal	Fraction

Appendix 1b

Similarities and differences in didactical approaches in Flanders and China

	Flanders	China
Number	Place-value concept	Place-value concept
		
		
Addition	Eg. 8=4+4=2+6=1+7…. Eg. 17+5=17+3+2	Eg. 8=4+4=2+6=1+7… Eg. 17+5=17+3+2
Subtraction	Eg. TE-E=TE TE-T=E Eg. 5+6=11->11-6=5 Eg. 14-2=12 14-12=2	
Multiplication	Recite 5 X 2 is five times two 5 X 2 = 2+2+2+2+2	Recite 5 X 2 is five times two 5 X 2 = 2+2+2+2+2
Division		

Appendix 1c

Teaching time in Flanders and China weekly in spent in mathematics education

Country	Flanders	China
Lessons duration	50 minutes	40 minutes
Grade 1	5 lessons	3-4 lessons
Grade 2	5 lessons	3-4 lessons
Grade 3	5 lessons	4-5 lessons
Grade 4	5 lessons	4-5 lessons
Grade 5	5 lessons	4-5 lessons
Grade 6	5 lessons	4-5 lessons

Chapter 10

Mathematics learning performance and mathematics learning difficulties in China: what did we learn?

Abstract

The overall aim of the research presented in this doctoral dissertation was to analyse the factors affecting mathematics performance of learners in Chinese primary schools. Models were discussed to map the complex interplay between individual level, classroom level and school level variables that affect mathematics performance in Chinese learners. In particular, the further impact of socioeconomic status (SES), classroom teaching processes (interaction and questioning) and homework were investigated. In a second part of the PhD study, variables related to being at risk for mathematics difficulties (MD) were studied. In groups with a high versus low performance level, the development of basic numerical facilities was studied. Figure 1 presents a graphical overview of the PhD study and how the subsequent chapters are positioned within the overall PhD dissertation.

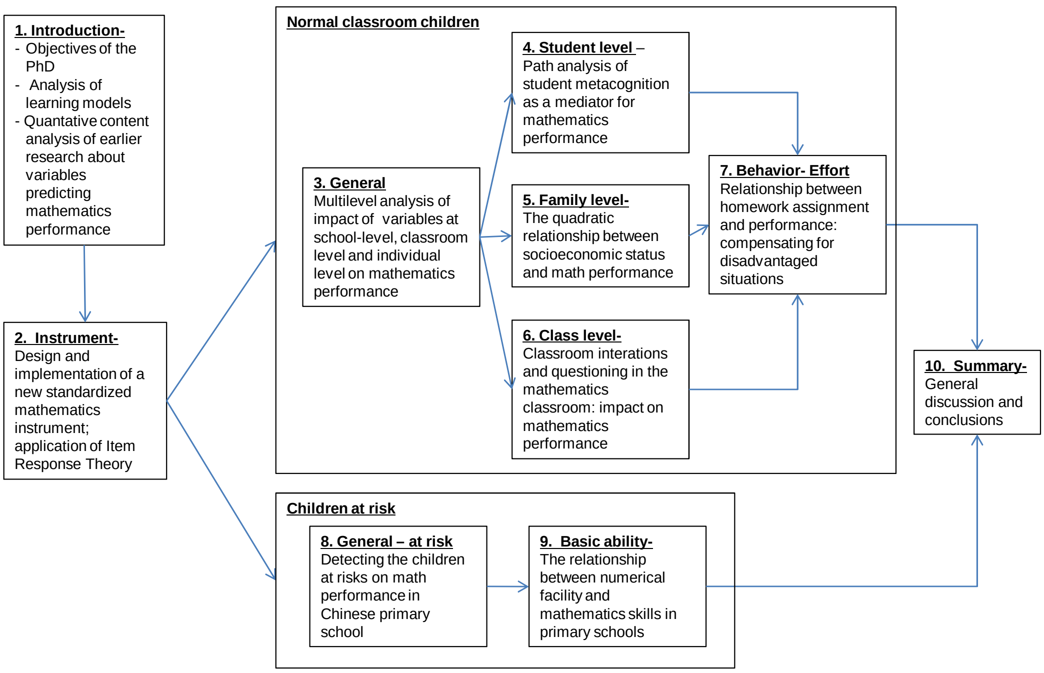


Figure 1. Overview of the chapters in the PhD dissertation.

In this final chapter, the main findings of the eight studies, presented in the PhD, are summarized. We integrate the research findings and theoretical and practical implications are discussed. From this point, we formulate limitations of our research and present directions for future research.

1. Introduction and conceptual framework

It is hard to overemphasize the importance of mathematical literacy in our society. In everyday life situations we need to be in time, pay bills, follow directions or use maps, look at bus or train timetables or comprehend instructions and consult expiry dates. Differences in the mastery of mathematics between and within individuals are normal. Teachers are expected to cope with learning differences and to adjust their teaching approaches to the needs of all students. However, in some cases these differences appear to be so large that they have to be considered as 'mathematics difficulties' (Grégoire & Desoete, 2009). In this first section of the concluding chapter, we briefly recapitulate the studies presented earlier, and this in view of making clear the decision made in each chapter that helped to shape the conceptual framework adopted in the different studies to describe and explain mathematics learning and performance in Chinese primary schools.

Tests are needed that are sufficiently sensitive and specific to map the full range of mathematics abilities and to differentiate children between with mathematics learning difficulties and children not at risk for developing mathematics difficulties. Moreover, a comprehensive assessment is needed in order to offer a solid remediation based on the strengths and weaknesses of every child (Grégoire, 1997; Stock, Desoete & Roeyers, 2006).

In this dissertation, first of all, we focused on the development of a standardized instrument to assess mathematics performance in China (see Chapter 2). The new test was developed covering the new mathematics curriculum of the primary school, as well as reflecting a validated conceptual model reflecting twelve mathematics building blocks already emphasized in previous studies (see Desoete & Roeyers, 2005; Nunes et al, 2007; Zimmermann & Cunningham, 1991).

According to the twelve-building blocks model, mathematics depends on adequate number-naming or reading (NR) skills that help link numbers to other types of representation (e.g., the Arabic number presentation '9' is linked to an oral representation of the word 'nine'). Children need to know that 'nine' is not written as '6' and that '47' is not read as 'seventy four'. The second mathematics building block centers on the mastery of the mathematics lexicon. To solve mathematics problems, children have to deal with operation symbols (S) (e.g., x, +, <, >) without making mistakes of a perceptual (e.g., x or +, - or =, < or >) or phonetic type (e.g.' min' or minus, 'maal' or times).Furthermore, the building blocks incorporate insight in the number structure, the knowledge (K) of the positional system in writing down numbers, and the mastery of the base-ten structural relationships. K skills are required to understand that 47 is composed of 4 decades and 7 units and that 47 is 1 more than 46 on the number line. A next mathematics building block indicates that mathematics depends on procedural (P) knowledge to calculate and to solve number problems (e.g., 47-9=_). Children have to know how to make subtractions to solve 47-9 as 38 and not as 42. The next building block stresses linguistic skills (L) that are skills enabling children to understand and to solve mathematics problems presented in phrases; as word-problems (e.g., 9 less than 47 is_). A key building block when solving word problems and other mathematics tasks is the reliance on mental representation (M). A simple 'translation' of keywords in a word problem (e.g., ‘less’) into calculation procedures (e.g., ‘addition’), without the usage of adequate representations, might lead to ‘blind calculation’ or ‘number crunching’. This superficial solution approach easily leads to errors, such as answering '38 to tasks as '47 is 9 less than _', ’29 is 9 more than _' and ‘76 is half of _’. Contextual skills (C) refer to a building block that comprises goes beyond the direct mathematical numbers, operations, etc. Some children have problems with this building block due to the limited capacity of their working memory (‘cognitive overload’) or due to an insufficient knowledge base (or ‘expertise’) about the broader problem setting. This can easily be linked to the next building block that requires children to select relevant information (R) in order to create an adequate mental representation of the problem. Children have e.g., often difficulties in ignoring irrelevant information in a mathematics task. They believe that all numbers have to be ‘used’ in order to solve a problem. They answer ‘59’ (47+3+9) to the problem ‘Willy has 47 cards. Wanda has 3 books and 9 cards more than Ann. How many books had Wanda?’ . Number sense skills (N) comprise a ninth building block that enables learners to solve tasks without giving directly the exact answer. Some children are not able to estimate a solution (e.g., 250-49=_ will be around 200). G (“Geheugen” in Dutch and “memorising” in English) skills presents the tenth building block and refers to memorized information and/or automated skills. These are stored in long term memory and readily available in a or more conscious way. For example, when executing tasks related to multiplication, the tables are automatically generated to support the solution process. Visualization skills (VS) represent a building block helpful for solving spatial problems. VS skills enable students to produce and use geometric or develop and understand graphical representations of mathematical concepts, principles, or problems. (Fennema & Sherman, 1977; Zimmermann & Cunningham, 1991). The last building block refers to ‘logical thinking’ (LT). Logical thinking skills imply the understanding of logical relations between quantities in order to represent numbers and arithmetical tasks (Nunes et al, 2007; Piaget, 1952).

In a pilot study, and in relation to the twelve mathematics building block, test items were developed and tested by experts and teachers for each primary school grade. In a main study, these items were evaluated, involving 10,959 students from schools in five different Chinese regions. The Item Response Theory scaling of the items helped to establish a valid and reliable test, covering the six primary school grades with a sufficient overlap between grades to map the full scale of mathematics abilities in primary school children.

Second, based on a review of the literature, a holistic model was developed and evaluated in the context of mathematics learning in China. Since variables are nested, different levels have to be considered. Therefore, we looked at variables at the school level, the class level, and the student level (see figure 2). In typical educational setting, these variables are nested in a hierarchical way. Individual learner variables are nested within class variables that are nested with the school level (Raudenbush & Bryk, 2002).

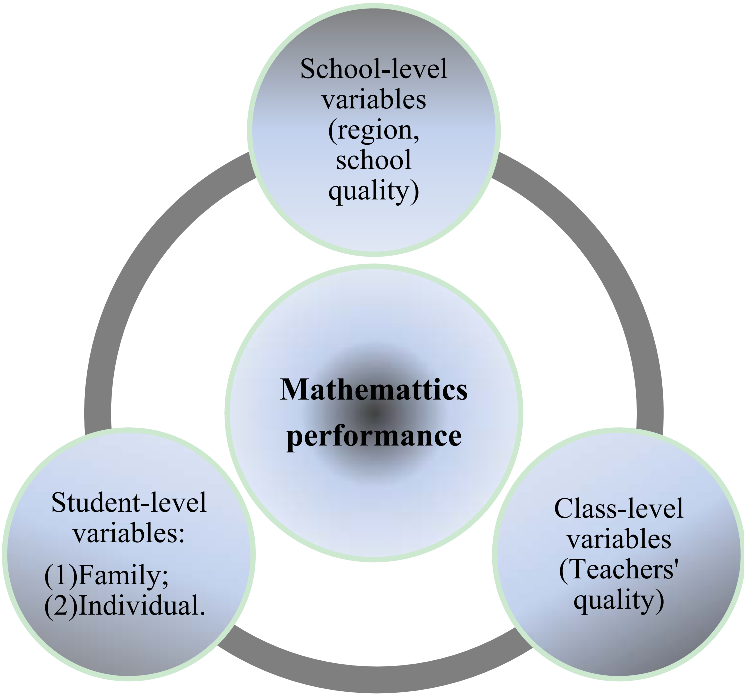


Figure 2 Model of the individual and contextual variables on mathematics performance.

Our study considered the findings of earlier studies to test the particular impact of individual, class, and school level variables. A meta-analysis of Bosker and Witziers (1996) revealed that up to 18% the variance in academic performance is to be attributed to school level variables. In addition, because that children live in different regions in China with very different Gross Domestic Product levels (GDP), family SES was added as a potential predictor of mathematics outcomes. In addition, we distinguished between the geographical setting fo schools in urban and rural areas. Also, the school mean of the students' socioeconomic status (reflected in parents' job and family wealth level) were considered at the school level. At the class level, grade and "type" of teacher were considered. In addition, at the individual student level, family background variables (SES) were studied, next to gender, birth order, mathematics self-efficacy (MSS; Marat, 2005), metacognition and mastery of the Chinese language. Model testing was based on the data, obtained from 10,959 students enrolled in Chinese primary schools in rural/urban within five provinces with different developmental levels (see Chapter 3).

In chapter 4, the role of individual variables as mediators between contextual variables and mathematics performance was studied in more detail. We distinguished internal variables such as grade, gender, mathematics self-efficacy, and metacognitive experiences. As external variables, the Gross Domestic Product of the region was considered in the analysis. Father's and mother's educational level were selected as variables within the family context. Further, in the school context, teacher's educational level and the teacher beliefs about Student Learning, Stage of Learning and Teaching Practices were added to the model. Figure 3 represents that resulting multi-level model in a graphical way.

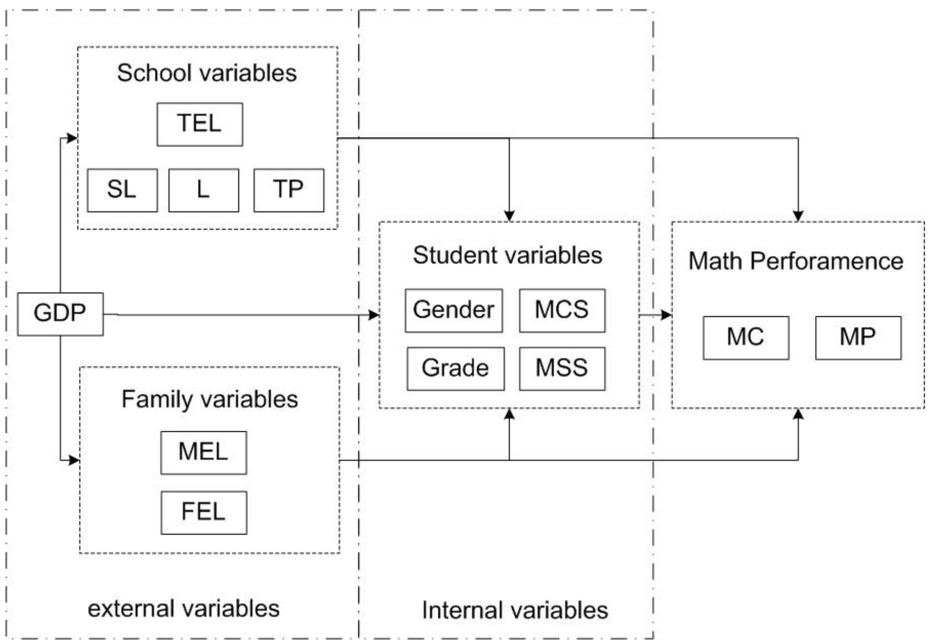


Figure 3
Levels in internal and external variables as they are related to mathematics performance.

Note: GDP-Gross domestic product; MEL-Mother Educational Level; FEL-Father Educational Level; MP-Mathematics Performance; MSS-Mathematics Self-efficacy Score; MC- Veracity of Mental Calculation; Gender-Student's Gender; TEL-Teachers' Educational Level; SL-Teacher's belief on Student Learning; L-Teacher's belief on Stage of Learning; TP-Teacher's belief on Teacher Practice; MCS-Metacognition Calibration Score.

In chapter 5, the relationship between socioeconomic status and learning performance was studied in more detail. This was based on the analysis of the mathematics performance of 10,959 students, enrolled in grade one to grade six, from twenty Chinese primary schools. A multistage stratification sampling procedure was followed. These twenty schools are located in five Chinese regions reflecting different development levels; and are located in either rural or urban settings.

In chapter 6, classroom teaching practices were analyzed. A sample of 601 students and 9 teachers participated in this study.

In chapter 7, we focused on the nature and impact of types of homework assigned by teachers, parents and learners. The analysis of data from 10,959 students, enrolled in Chinese elementary schools, was analyzed.

The purpose of chapter 8 was to explore the interrelated effect of individual (gender, metacognition and self-efficacy) and environmental variables (family SES, teacher certification level, GDP of school region) on being at risk to develop mathematics learning difficulties in Chinese primary schools. In this study we analyzed in particular the performance of 2738 children reflecting mathematics learning difficulties in China.

The aim of chapter 9 was to investigate the impact of cultural differences in the mastery of numerical facility skills. A total of 7247 Chinese students and 913 Flemish students, enrolled in grade 3 to grade 6 of primary education, were involved in this study.

In this final chapter, we summarize and discuss the main findings of the studies. Furthermore, limitations of this doctoral research project will be formulated and future research challenges will be proposed. Finally, we conclude with the implications of our work.

2. Main findings

2.1 Research objective 1: Instrument development

In chapter 2, a new and comprehensive test was developed to cover twelve mathematics building blocks: number reading skills, mathematics 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. In view of considering the content and structure of the Chinese mathematics curriculum, the three key mathematics domain were clearly represented: number and algebra, shape and space, statistics and probability (see MOE, 2001). To develop and implement the instrument, an IRT approach was adopted to estimate item difficulty and participants' ability (trait). This helped to position individual learners along an underlying latent trait continuum.

The instrument proved to be reliable with reliability indices (comparable to Cronbach's α) ranging from .94 to .96 for the different primary school grade tests. The average mean item and test correlation was .39 (SD = .14). An average mean bi-serial correlation of .54 (SD = .19) was observed. Secondly, the test helped to identify participants' mathematics ability, ranging from -4.89 to 4.14 and covered item traits ranging from -5.30 to 9.26 ($-5.30 < -4.89 < 4.14 < 9.26$). Test information curves provided information about the mathematics abilities between theta -1.65 and 2.22. Finally – and this is a critical quality indicator - only about 9.1 % of the respondents (n=998), were located in the area where test information is lower than the standard error. As to the validity of the instrument, analysis results pointed at a good validity. First, the internal content validity was checked by comparing the difficulty estimation, based on experts' expectations, and available IRT data; resulting in a .82 correlation. Second, the construct validity of each item is based on the number of students who answered correctly or incorrectly, reflected a good result. The fact that this first research objective was attained was the starting point of pursuing the subsequent research aims.

2.2 Research objective 2: variables related to mathematics performance

2.2.1 General model on variables contributing to the mathematics performance

In chapter 3, data were obtained from 10,959 students with the newly developed test and on the base of additional questionnaires and tests, to capture the relevant predictors of mathematics learning and performance in Chinese primary schools. The results – in general – confirm the complex nature of indicators that influence mathematics performance (Walberg 1981; 1982; Creemers, 2007). At the school level, the province development level of a school and the rural/urban location of the school were found important to be included in the prediction model. But, since individual student – in whatever province the school was located - still were able to attain high performance levels in mathematics, differences between schools had to be studied within each province based on the previous studies in China (Zhao, Valcke, Desoete, Verhaeghe, in press). In addition, the aggregated average socioeconomic status of learners in a school was also a significant predictor. This result is in line with the previous studies that the students are not randomly assigned in the schools (Hanushek, Kain & Rivkin, 2004; Perry and McConney, 2010). The latter remained significant until individual student level variables were included in the model (e.g., reading level) .

At the class level, grade was found to be a significant predictor for students' mathematics achievement. Teacher's level of graduation did significantly help to predict student performance, until individual students level variable of metacognitive experiences was added to the model (see below). Teachers graduating with a diploma from a formal higher education institute were found to be related to higher achievement levels, compared to teachers who graduated from open teacher training systems (e.g., Open University teacher education). The result is the same as the previous studies which claimed that the quality of the teachers work for the students' performance (Goldhaber & Brewer, 2000; Smith, Desimone, & Ueno, 2005). At the student level, reading performance, mathematics self-efficacy and metacognitive experiences revealed to play a significant role in the prediction of mathematics performance. And these findings which is in line with the previous studies: high level of metacongitive experience (Efklides, 2006; Efklides, 2008; Foong, 1993), mathematics self-efficacy (Pajares & Graham, 1999) and reading performance (Grimm, 2008) have a positive effect on the mathematics performance. The results indicate that remedial or intervention programs have to be proactive and screen for these three characteristics of Chinese students; especially in students at risk of developing mathematics difficulties. Finally, the socioeconomic status of family was a weak and a polynomial predictor. This finding is explored in more detail in a subsequent study. And, this study revealed that individual variables explain up to 46.67% of the total variance in mathematics performance. However, the result did not means that the school level variables and class level variables had no contribution to the mathematics performance. This results also revealed the same as in the previous studies that the average of the school level variance is for 18% of the total variance (Bosker & Witziers, 1996). But, after controlling for student characteristics, school and class level variables disappear. At the school level, the aggregated socioeconomic status of school was a significant predictor. At the class level, grade was a significant predictor. And teacher's graduation level only predicted mathematics outcomes when – at the the individual student level - metacognition was added. At the student level, reading performance, mathematics self-efficacy and metacognition were – as already suggested - important. Socioeconomic status of family remained a weak and polynomial predictor.

2.2.2 Students' variables plays an important role for mathematics performance

In chapter 4, the role of individual variables as mediators between contextual variables and mathematics performance was studied in detail. Path analysis was used to test the direct and indirect relations between predictors of mathematics performance in 1749 students. Information was also obtained from 91 teachers. The final path model is presented below, reporting the standardized path coefficients.

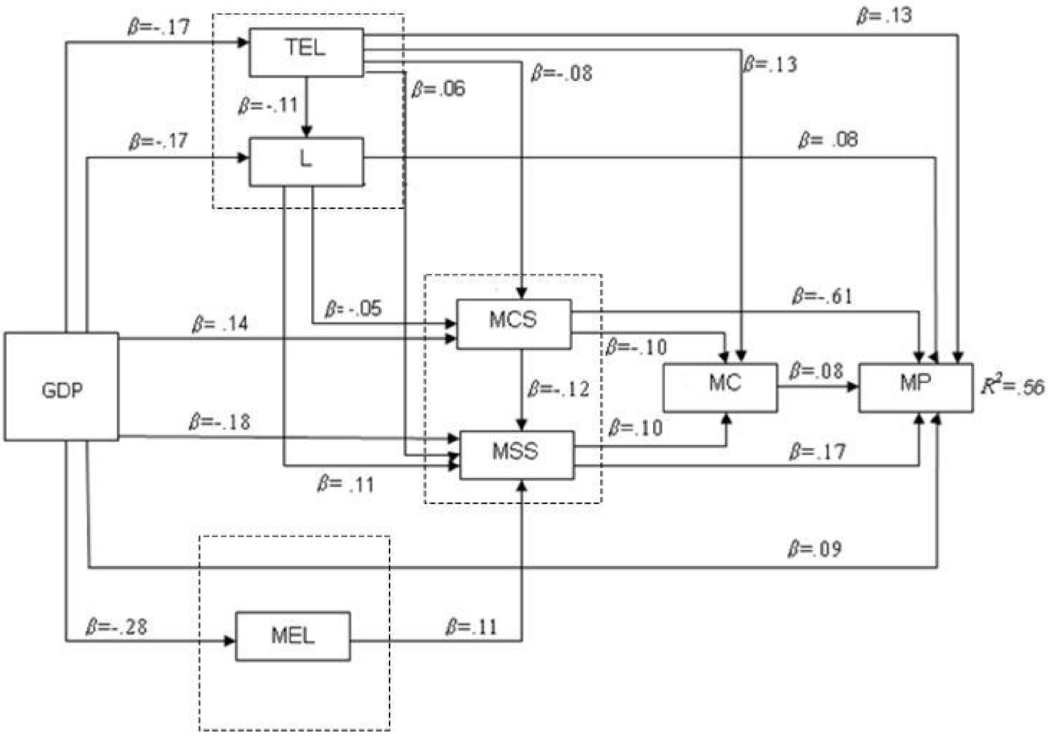


Fig 2.

Result of the path analysis

Note: GDP-Gross domestic product; L-Teacher's belief on Stage of Learning; MCS-Metacognition Calibration Score; MEL-Mother Educational Level; MP-Mathematics Ability; MSS-Mathematics Self-efficacy Score; TEL-Teachers' Educational Level; MC- Mental Calculation.

In the Chapter 3, the results revealed that when the students variables were added to the multilevel regression level, the class level and the school level variables disappeared. This means that there might be some indirect ways from the contextual variables to the mathematics performance. In Chapter 4, this hypothesis has been tested. By using the path analysis, the results of this study stress the importance of the economic development level of the region where a school is located and the role of teachers' beliefs and teachers' educational level. This is line with the previous studies that the development level of regions where the school located has indirect effect on the students' performance (Perry, & McConney, 2010), and if the teachers' beliefs on the mathematics tend to be more constructivism, the students performance will be better (Staub & Stern, 2002) works. In addition, the educational level of the mothers has a positive effect on the students' mathematics self-efficacy. Also, teacher's educational level has a positive impact on metacognitive experiences which is in line with the previous study, while the teachers' beliefs about the stage of learning rather have a negative impact on mathematics self-efficacy. These results have not been found before. The relationship between the self-efficacy and the contextual variables is not been explored in the previous studies in the primary school. These results imply that, although Chinese teachers follow the strict sequence in the Chinese mathematics curriculum, this does not restrict the development of students' metacognitive experiences and mathematics self-efficacy. This finding seems to be in conflict with the student-centred learning style that is promoted everywhere as the best approach to teach problem solving (Marton, Dall'Alba, & Lai, 1993; Huang & Leung, 2005; Watkins & Biggs, 2001). However, this typical "Chinese" finding will be confirmed on the base of the results of the video analysis study, reported in Chapter 5. Finally, we conclude that students from schools located in provinces with a higher economic development level, tend to have attain a higher level of metacognitive experiences and reflect lower levels of mathematics self-efficacy.

2.2.3 Family socioeconomic status has a weak impact on mathematics performance

Chapter 5 investigated the relationship between family socioeconomic status and mathematics performance of Chinese children. The relationship between SES and performance is a recurrent theme in mathematics education research. Meta-analysis of research about the relationship between SES and mathematics performance reveals divergent results. First, we observe varying ways to measure socioeconomic status. After a review of key studies, we decided to integrate father's and mother's educational level, father's and mother's job level and family wealth indicators in our operational definition of SES. To study job levels, we could build on available Chinese research (Li, 2005a; 2005b). In the subsequent analyses, the general factor score for parents' job level, family wealth indicators and variables about parents' educational level were step-by-step added to the model.

The analysis results show how the relationship between SES and mathematics performance is U-shaped. This is clearly different from the mainly linear relationships, dominantly reported in e.g., international performance indicator studies (PISA, TIMSS). The U-shape in this particular relationship implies that students from a disadvantaged family and higher socioeconomic background have a higher probability to attain higher mathematics scores. For example, within the lowest 25% achievers (n=2920 students), 45.34% of their parents are both working and 28.42% of these parents are both peasants. These persons represent up to 80% of the population in China (Li, 2005). The U-shape can be seen as a positive sign about the way the Chinese educational system gives opportunities to lower SES groups. It is suggested that this shows how National Examination Policies in China increase opportunities for and "openness" of the society (Blau and Ruan, 1990; Kracke, 1947; Parish, 1981; Wu, 2007). Nevertheless, since we speak here about 80% of the population, educational authorities have to remain alert. Also, the actual situation of the lowest 25% achievers from this 80% of the population should especially worry educational authorities.

2.2.4 Teacher's influence on mathematics performance

In chapter 6, our dataset consisted of videotaped lessons. Basic analysis of the video material showed how "one teacher-whole class" interaction (about 15 minutes/per 40 minutes lesson) and interaction of "one teacher-individual student" (around 13 minutes/per 40 minutes lesson) dominated the interaction patterns in Chinese mathematics classroom at primary school level which is in line with the previous studies (Stevenson & Lee, 1995; Lim, 2007). The results of multiple regression and multilevel analysis confirmed that the interaction of "teacher-individual student" and the "student-student" interaction in public, guided by the teacher and/or other colleagues, play a positive role to attain better mathematics performance. Larger proportions of "student-student" interactions, rather result in a negative effect on outcomes. In classroom teaching, teachers dominate the Chinese classroom interaction which causes the paradox of the Chinese teaching (Marton, Dall'Alba, & Lai, 1993; Watkins & Biggs, 2001). In the other studies in other countries, the class teaching should be dominated by the students, for example, the students will get more opportunities in discussion by the small group, and it is assumed that this can improve the performance (Yackel, Cobb, & Wood, 1991). This type of interaction is found to be suitable for large classroom sizes, typically found in China. The interaction is expected to go beyond teacher centered approaches and also to support a students' zone of proximate development, since it aims at helping students to reflect on their learning process. In addition, the research focused on the questioning strategies of teachers and how this is related to learning outcomes. The previous studies show that the quality of the questioning contributes to the mathematics performance (Nathan & Knuth, 2003). Questions stressing problem solving strategies (30.16%), identification (26.61%) and evaluation (11.53%) were found to dominate Chinese classrooms. The results showed that higher proportions of questions focusing on problem solving strategies and evaluation result in higher performance (Garza, 2009). In addition, it was stressed that problem solving questions and evaluation questions support student reflections on errors and on the construction of math concepts as the same in the previous study (Swanson & Lussier, 2001).

2.2.5 Homework to compensate for a disadvantaged situation

Chapter 7 made an effort to combine contextual and individual student variables. Homework was selected as the key study object. In the particular Chinese context, homework is a complex phenomenon, since three key actors assign homework:

teachers, parents and students themselves; this makes this study unique in the broader discussion about the educational benefits related to homework assignment. This in particular refreshes considering the debates in the homework literature (see e.g., Kohn, 2006 or Cooper, 1994). In the previous studies, there are more debates on the homework and the self-regulation (Hong, Peng, & Rowell, 2009; Zimmerman & Kitsantas, 2005) which implies that the homework is an effort or behavior from the person by self-regulation. The study results show how homework can compensate for a disadvantaged situation and as such improve mathematics performance which is in line with the previous studies that in Asia, the homework is considered as investing additional “effort” as is as such a key to academic success (Chen & Stevenson, 2008; Sue & Okazaki, 1990). The study – in particular - revealed that parents and students assign homework as a way to meet a disadvantaged background under the society of the “openness” to the mobility (Biau, & Ruan, 2990). Our dataset shows how parents and students from disadvantaged families can use homework as a way to compensate for their underprivileged background. Also, students seem to develop homework assignments, considering the extent to which their parents assign homework. Learning performance of students of parents with low level jobs assigning them a moderate amount of homework, results in a significant performance improvement. In contrast, student achievement of learners was significantly lower when their parents - with high level jobs - did not assign homework. Students from disadvantaged families benefitted largely from getting extra homework if the students from white collar families did not assign homework for them. The result also means that the effect should be considered in recent China. This result is specific in Chinese context because in the other countries the homework studies focus on the teachers’ homework behavior and not on the parents and the students. But, we can still find some related results from the international studies, such as, it is believed that parents’ beliefs about homework affect the students’ beliefs on homework (Cooper, Lindsay & Greathouse, 1998).

2.3 Research objective 3: variables related to mathematics difficulties

2.3.1 A model to study variables related to mathematics difficulties

In chapter 8, the performance of students with mathematics learning difficulties (students of the lower 25 percentile in each grade) are compared with the performance of students without learning difficulties.

The study revealed that a low self-efficacy level and below average metacognitive experiences increased the likelihood to reflect mathematics difficulties. Second, a lower socio-economic status of the family, a restricted teacher certification level and the lower GDP level of the school region also increase the likelihood of being a student at-risk. Furthermore, adding school’s regional GDP and the teacher certificate level to the model, decreases the unexplained random effect variance from 20.65% and 34.7% to 9.20% and 35.3%, respectively.

Next to individual characteristics, context variables also affect the risk of developing mathematics difficulties. This findings is comparable to previous studies. In developing countries, about 200 million primary school children run the risk to fail attaining adequate performance levels that truly reflect their developmental and cognitive potential, and this mainly due to a disadvantageous context (Grantham-McGregor, Cheung, Cueto, Glewwe, Richter & Strupp, 2007).

2.3.2 Difference in the development of basic numerical facilities

The findings of the study, reported in chapter 9 reveal that Chinese students outperform Flemish students. This is consistent with the findings or earlier research (Imbo, & LeFevre, 2009). The present study goes beyond the available “typical” cross-national comparative studies since it also focuses on differences in the different grades of primary school. Although the performance difference between Chinese and Flemish students decreases with increasing grades, the difference remains significant even at the end of elementary school. Moreover low achievers especially meet problems with multiplication tasks; and this in both countries which is in line with the previous studies (Desoete et al., 2009). In addition, students in both countries perform better on addition tasks as compared to subtraction or multiplication tasks, the latter also being easier than division or mixed exercises. As shown in the previous studies, the multiplication tasks relies more on the phonological working memory resources (Imbo, & Vandierendonck, 2007). But, there is little difference for the Chinese students and Flemish students in that the difference between multiplication and other calculations for Flemish students is little smaller difference between multiplication and other calculations for Chinese is larger. The results of a 4(Grade) x 2 (country) x 3 (level of skills) MANOVA on five calculation of addition, subtraction, multiplication, division and mixed operations reveal that, although the difference between Flemish and Chinese students are significant for addition, subtraction, multiplication, division and mixed calculation at each grade, with the increasing of the school year, the differences between two countries for addition, subtraction, division and mixed become smaller except for multiplication. Another result is for the lower achiever meet the problem of solving multiplication operation.

The results of a 4 (Grade) x 2 (country) x 3 (level of skills) MANOVA reveal that students perform better on addition than on subtraction or multiplication tasks; the latter being easier than division or mixed exercises. Flemish learners reflect consistently smaller differences in the mastery of types of numerical facility skills (e.g., addition versus subtraction, subtraction version mixed operation) as compared to Chinese learners. In both countries high achievers perform stable (across the four grades) on all numerical facility tasks, while low achievers do not reflect a stable achievement level when comparing their mastery in the four different grades.

3. Overall conclusions and further discussion of the findings

3.1 Assessing mathematics in China

As mentioned in Chapter 1, in Chinese education in general and in Chinese mathematics education in particular, we observe a debate about the importance of “Suzhi” (meaning ability, quality, literacy, etc.) versus “zhishi” (meaning knowledge, information, etc.). This debate is not raised by educational psychologists, but rather by educational researchers. Although, the debate seems to reflect a theoretical debate about the nature of curricula and education, the debate has already influenced researchers in the way they create suitable mathematics instruments. In parallel with the start up of the current PhD in 2007, at least two research projects started in China that centered on this discussion: the first is a project about the “educational output of the elementary education” started in 2007 by the Chinese Ministry Of Education (see, Xin, Tian & Zou, 2010); the second is a project about “Characteristics of Chinese children and adolescent psychological development survey” supported in 2008 by the Chinese Ministry of Science and Technology (see, Dong & Lin, 2011). The previous project focused on the curriculum and specific subjects; the latter project focused on cognitive ability, scholastic ability and psychological health. In these projects, a recurrent theme centered on “what to develop and measure” in education.

The debate about knowledge and ability affected our strategy to develop a new test, suitable to be used in Chinese primary schools. To respect the “ability” orientation, analysis of the literature helped to identify relevant mathematics “building blocks”. To respect the “knowledge” orientation, the new instrument also considered the three main domains considered in the Chinese primary school curriculum. A large scale implementation of the test made it clear that this test could be considered as a valid and reliable way to test mathematical performance of Chinese children. Compared to available tests to be used in Chinese primary education, for example, Children Developmental Scale of China (CDCC, Hang, zhou, & Zhang, 1992-1994), Diagnosing Scale of Cognitive Ability for Children (DSCAC, Lv, 1991), we can state that our instrument (1) constructed the items based on the most recent curriculum syllabus (MOE, 2001) which can be seen to be a standardization instrument for the scholastic test for teachers to use for diagnosis in the primary school; (2) made the instrument by using the sample from more than 5 levels of the provinces and estimated the item parameters by item reponse theory which provide the developing space for the instrument, for example, if the syllabus is changed, the overlapped items can be used to change to calibrate the parameters of new item and old items, or for example, the difficulties of items can't cover the students' ability in the future, the new items can be added the instrument by the overlapped items also; (3) the overview of continuum scale of the items and the students from six grades can provide us more information for the developing studies, for example, the longitude studies can be done based on our standardization instrument. The test is clearly a step forward from available instruments.

3.2 Mathematics performance in China: three levels are needed to develop a comprehensive picture

The development of mathematical literacy in primary school is a complex process that is influenced by a large set of variables. Three levels (the school level, the class level, and the student level) were needed to develop a picture that is sufficiently rich and varied to predict mathematics performances of primary school children in China. The results are in line with the findings of research in Flemish primary (Ministerie van de Vlaamse Gemeenschap, 2004) and secondary education (Opdenakker & Van Damme, 2006). These studies found that between 15 percent and 20 percent of the variance in achievement scores could be attributed to school variance. In addition, the results reveal that there are large differences between schools, so research and related theories about the development of mathematics abilities should not only take into account individual processes of children, but should also incorporate the context in which these children are expected to develop. This is in line with the basic models already discussed in chapter 1 (Creemers, Geary, etc.).

A multilevel model was needed to identify significant predictors of mathematics performance in Chinese primary schools (n=10,959), explaining 46.67% of the total variance. At the individual level, a clear set of variables contributed in a significant way to mathematics outcomes of children; especially metacognitive experiences and the self-efficacy played an important role to attain proficient mathematics performance. This is in line with earlier research pointing at the value of these variables (Efklides, 2006; 2008; Pajares & Graham, 1999). At the classroom level, teacher's beliefs had only limited predictive value in view of mathematics performance. This is not in line with earlier research that teachers' beliefs contribute a lot to the performance (Staub & Stern, 2002). However, our findings could be explained by the use of the specific survey instrument to determine teacher beliefs. The findings also differ from what we learned in Chapter 4 where teacher's beliefs were significantly related to mathematics performance of their students. From Chapter 6 we already learned that the type of questions asked by teachers can influence performance; especially evaluation and problem-solving related questions. The role of the teacher was also confirmed by the video analysis study. The nature of student-teacher interaction has to be qualified as "teacher-centered". This could be linked to the size of Chinese classrooms. Also, we suggest that the paradox reflected in these results can be solved. The teacher-centered approach does not restrain teacher to strongly support individual students and as such move in the zone of proximate development of learners (Daniels, 2001; Vygotsky,1978). Next to the individual and classroom level; also the school level is needed to explain mathematics performance. The (aggregated) average socioeconomic status of a school, the grade and the teacher's level of graduation were found to be important predictors of mathematics outcomes in Chinese primary education.

Compared to the previous studies, the multilevel analysis carried out the studies in a different way. The contextual variables were added to the model at the first step and deleted later. The coefficients of these dropping help us to find out the indirect relationship between different level variables. There are more interactions based on these studies can be carried out.

3.2.1 Metacognition and mathematics

Metacognition presents a promising contemporary research field in psychology and education (Efklides, 2001, 2008; Veenman, Van Hout-Wolters, & Afflerbach, 2006). The concept has been introduced to describe and explain how people gain control over their learning and thinking, particularly in the case of errors and difficulties they meet during information processing in general and problem solving in particular (Efklides & Sideridis, 2009; Flavell, 1976). Metacognition has been described as having three facets, namely metacognitive knowledge, metacognitive experiences and metacognitive skills (Efklides, 2001, 2008; Flavell, 1979). 'Metacognitive knowledge' has further been described as the knowledge and deeper understanding of cognitive processes and products (Flavell, 1976). In addition, *metacognitive experiences* are what the person is aware of and what she or he feels when coming across a task and processing the information related to it (Efklides, 2008). Metacognitive experiences, in essence, make the person aware of his or her cognition and trigger control processes that serve the pursued goal of the self-regulation process (Efklides, 2008; Koriat, 2007). To assess metacognition, we adopted a calibration approach. The theoretical expectations as to the potential impact of metacognition could be supported by our analysis results. Path analysis revealed that a large proportion of mathematics performance can directly be predicted from the level of students' metacognitive experiences. In addition, other student characteristics and contextual variables also influenced mathematics performance in direct or indirect ways; as already suggested by our earlier research findings.

3.2.2 SES and the impact of homework

The U-shape relationship between SES and mathematics performance seems to be typical for the Chinese situation which is not been found before (see meta-analysis studies of Sirin, 2005; White, 1982). This presents a big problem since a large proportion of students is from blue collar families. And as stated earlier, their performance is lower than students from parents with a lower job level or from white collar families. Though we observe in China a level of “openness” in society, making it possible that students can grab opportunities to change their life by working hard, the SES related findings have to be taken into account by educational authorities. National or school-based programs could consider the SES background of learners and center the attention on this blue collar target group.

Homework in China is not only assigned by teachers. Also parents and learners themselves develop homework tasks. Our study revealed that parents and students from disadvantaged families approach homework as a way to compensate for an unprivileged background. This research finding has yet not been reported in the literature (See, synthesis of research in Cooper, Robinson, & Patall, 2006. Learning performance of students with parents with a low level job and assigning a moderate amount of homework help to improve significantly mathematics performance. Students from disadvantaged families benefit largely from homework involvement.

The relationship of SES and mathematics performance is a result embedded in the local culture. Compared to the other studies, the studies on the relationship between SES and performance here provide more information about the Asian culture, such as the quadratic relation between the SES and performance in primary school, homework as the effort to compensate for disadvantaged situation. In Chapter 1, we mentioned that the parents’ and student’s effort contributed to the performance by many studies (Huang, & O;Neil, 1997; Ho, & Hau, 2008). But the previous studies focus on the relationship between motivation, strategies, effort and performance. Here, the two studies explored the relationship between the effort from different persons and their combination results.

3.3 Mathematics learning difficulties in China

During the past two decades, more and more studies focused on mathematics learning difficulties (Desoete, Roeyers, & De Clercq, 2004). Our study resulted in three conclusions:

- (a) Individual variables no doubt play a very important role in being at risk for learning difficulties. A low self-efficacy level and low metacognitive experiences increase the likelihood of experiencing mathematics difficulties.
- (b) Family SES and teacher certification level should also be considered as potential influences on learning opportunities and related performance. A lower

family SES levels, a restricted teacher certificate level, and a lower GDP level of the school region increases the likelihood of being a student at risk.

- (c) Individual variables and environmental variables interact and help to explain a large proportion of the unobserved class variability of being positioned in a student group with learning difficulties. A school's regional GDP causes the school-level unexplained variance of the random intercept to drop to .582. Adding school's regional GDP and a teacher's certificate level to the model, decreases the unexplained random effect variance from 20.65% and 34.7% to 9.20% and 35.3%, respectively.

4. Limitations and directions for future research

Already when presenting the various studies in the previous chapters, particular limitations were discussed in relation to the specific assessment procedures and about the research designs actually adopted. However, some more general limitations can be presented about our research project in general. A main limitation is linked to the particular set of predictors focused upon to describe and explain mathematics performance. As already stated in the introductory chapter, table 3 presents additional processes and variables that could be considered in future research. This observation can be repeated at each "level" in our model. At school level, we neglected - for example - parental involvement in school, school autonomy, and governmental initiative. At classroom level, attention could be paid to teacher's classroom preparation, actual didactical strategies being adopted. At the individual level, factors that have been neglected are intelligence, mathematics attitudes, mathematics anxiety, etc. Also at the individual level, but linked to the family setting, we could incorporate parents' investment in their children, family rules about school. Nest to the specific potential impact of isolated processes and variables, also the interaction between these predictors have not been studied.

In relation to test design and our focus on tests to be used in subsequent grades, we had to include a relevant subset of overlapping test items at each grade level. This overlap marred the ability to develop subtest scores about particular mathematics building blocks or about the three Chinese curriculum domains. In the present study, we continuously considered the complete test score. However, based on the available item bank, and after adding new items, it can become possible to develop a test that – next to a general mathematics score – also provides researchers and teachers with sub-scores in relation to the twelve building blocks and curriculum domains. This new test could open entire new research possibilities in view of performance analysis or the study of the impact of particular instructional interventions..

Third, although our large sample was selected on the base of a multi-stage stratification sampling approach, we still have to point out that the number of schools (n=20) is not very large in view of carrying out multilevel analysis. In order to counter this limitation, the item response theory helped to estimate and construct a continuum scale to position items and participants.

Fourthly, we acknowledge that all our studies were cross-sectional in nature. Longitudinal studies are needed to replicate the results of the current studies, but also to further analyses the complex interaction effects between predictor variables over time. As suggested in the opportunity-propensity model (Byrnes & Miller, 2007), past achievement is input to predict later achievement and reacts with predictor variables; e.g., high performance results boost expectations, motivation, parent expectations, etc. In this way, the feedback loop in models can be studied.

Since our focus reflects a strong interest in the influence of contextual variables, a number of individual variables have been neglected. Future studies might address additional individual variables, and the interaction between contextual variables and these individual variables.

We also see two main challenges for setting up future research. First of all, it remains necessary that pay more attention to the development of an empirical, unifying model that explains the development of mathematics abilities in a particular context. Groups of researchers could set up related research about this model and enlighten the importance of specific constellations of variables. This could result in a programmatic strand of fundamental research that – step-by-step – presents evidence for a comprehensive model. The second challenge, and perhaps the most important, is the need to link the model to intervention studies. As already suggested in some of our studies, the available evidence shows how early intervention can play a decisive role in countering mathematics difficulties and to redirect children's mathematical development in a productive direction.

5. Implications

5.1 Theoretical implications

At the theoretical level, the findings presented in this dissertation add to our understanding of culturally contextualized variables in mathematics education and mathematics learning.

First, our results confirm the findings of previous studies about the relationship between socioeconomic status and mathematics performance (Sirin, 2005; White, 1982). However, our study adds insights about the U-shaped relationship between these variables in Chinese primary schools. In addition, our studies confirm that Chinese children attain a higher mathematics performance level as compared to

Belgian children. But, the findings are now more fine-grained by centering on particular basic numerical skills. The differences between mastery of multiplication do not decrease with age. This suggests that differences in the mastery of this and other operations might depend on e.g;, the the language system when solving particular number problems or dealing with number facts (Imbo & Vandierendonck, 2007).

Second, our results hint at particular interactions between context variables and individual variables. In Chapter 3, random effects of metacognitive experiences in a school reduced the model variance significantly. In Chapter 5, the interaction between the aggregated school level SES and the individual family SES showed a differential effect of student enrollment in schools of a different level. In Chapter 6, the interaction between the school level, parents' job and assignment of homework by parents and students, showed how the school level influences student performance, but that student nevertheless can counter the negative impact of this context (see also the findings in Chapter 5). Finally, in Chapter 8, we perceive an interaction between a teacher's certificate level and a students' self-efficacy. To conclude, our results show that it is to be recommended not to assess one specific mathematical ability, but to look for a set of markers at the individual, class and school level and for the interaction between those markers when we aim to model variance in mathematics proficiency.

5.2 Practical implications

From a practical perspective, each chapter provides teachers with information to take charge of learners' mathematics learning.

From Chapter 3, 5 and 7, we learn that students in the disadvantaged setting can attain higher achievement if they make an effort to compensate for their situation. However, also derived from Chapter 3 and 8, students enrolled in schools from lower economical regions are more likely to become students-at-risk. Educational authorities should therefore pay additional attention to students from disadvantaged families in these particular schools and regions.

A further implication is linked to initial teacher training. In this context, one of our studies pointed at the negative impact of a particular teacher certificate level. In-service teachers could be trained to identify risk-situations, to continuously diagnose learners and to be focused on risk situations and contexts. This also requires that educational administration should provide more opportunities to teachers to be involved in in-service teacher training.

At the family level, we learned that (see Chapter 5) it remains worthwhile to put pressure on student's pressure to work hard or to make more efforts. Homework assignments – when a moderate amount is given - seem to contribute to students' mathematics learning and to compensate for being at risk.

At the student level, metacognition and self-efficacy were found to be two important predictors of mathematics performance. In Chapter 4, teacher and mother's educational level had an impact on the metacognition and the self-efficacy. Since available research shows that metacognition can be affeced by the special tasks(Akama, & Yamauchi, 2004; Efklides, 2006), instructional interventions could focus on this approach to foster mathematics learning and performance.

6. Conclusions

This dissertation studied the individual and interrelated role of a large set of variables that help to predict a proportion in the variance in mathematics performances in a large group of Chinese children. Our studies confirmed that education reflects and is dependent on a complex interplay of variables at the school, class, family and student level.

It was not a surprise to find again a confirmation for the key impact of individual factors that explain a large proportion of the variance in learning performance. Next to metacognitive experiences, mathematics self-efficacy, Chinese reading skills, …play a role in the mathematics learning and resulting performance. Moreover, individual characteristics mediated between contextual variables and outcomes. From our multilevel analyses we also learned that school level and class level variables become non- significant when individual variables are added to models.

Nevertheless, also family, class, and school variables - and the interaction between contextual and individual variables - contribute to student learning. Family SES, and the economical level and urbanization level of the region where schools are located, have a significant, but weak impact on mathematics performance. Differences within regions seem to be considered. Also, the U-shaped relationship family SES the mathematics performance reflects a different situation from what we observe in China versus other countries where a rather linear relationship is consistently observed. Third, classroom interaction practices clearly contribute to mathematics performance. In particular teacher questioning approaches can play a role in evoking particular mathematics outcomes.

During instructional processes, many actors affect student learning and outcomes: teachers, principals, parents, and students. The processes invoked by these actors also interact with one another. Mathematics performance therefore depends on this complex interaction between processes and variables related to different actors in different contexts. This results in unique constellations in the interaction between processes and variables. For instance, it explains why parents and students from disadvantaged family backgrounds can adopt homework assignment as a way to compensate for an underprivileged background.

In our final sections of our dissertation, more attention was paid to children at risk. This research was descriptive in nature and helped to compare children at risk with children reflecting "normal" performance levels. It helped to track how the economic development level of a region does affect performance of children at risk. It also helped to compare Flemish and Chinese learners at risk. The results indicate that children at risk in both cultural settings meet comparable problems. This implies that educators can start working together to exchange solutions to support children at risk.

To conclude this thesis, we return to the initial task, tackled when setting up our research. We started with the development of a new instrument to study mathematics performance. The nature and quality of this instrument was decisive to direct subsequent studies about models to describe and explain learning and performance. The instrument will continue to play a role in future studies; in particular to trace the impact of intervention studies. We hope that other researchers will adopt our instrument and continue to refine the models developed and tested thus far. We expect this will further theory development and future practices and ultimately benefit the teaching an learning of primary school children.

References

Akama, K., & Yamauchi, H. (2004). Task performance and metacognitive experiences in problem-solving. *Psychological Reports, 94*, 715-722

Biggs, J. & Collis, K. (1982). Evaluating the Quality of Learning: the SOLO taxonomy New York: Academic Press

Blau, P. M. and Ruan, D. C. (1990). In-equality of Opportunity in Urban China and America. In Kalleberg, A. L. (ed). *Research in Stratification and Mobility*, vol 9, CT: JAI Press.Hong, E., Peng, Y. & Rowell, L. L. (2009). Homework self-regulation: Grade, gender, and achievement-level difference. *Learning and Individual Difference, 19*, 269-276.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain*. New York: David McKay Co Inc.

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.

Chen, C. & Stevenson, H. W. (2008). Motivation and mathematics achievement: a comparative study of Asian-American, Caucasian-American, and East Asian high school students. *Child Development, 66* , 1215-1234.

Cooper, H., Lindsay, J. J., Nye, B., & Greathouse, S. (1998). Relationships among attitudes about homework assigned and completed, and student achievement. *Journal of Educational Psychology, 90*, 70–83.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research, 76*, 1–62.

Cooper, H.M. (1994). *The Battle over Homework.* Thousand Oaks, Calif.: Corwin Press, 1994.

Creemers, B.P.M. (2007) Educational effectiveness and improvement: the development of the field in mainland Europe. *In:* Townsend, T. (ed) *International Handbook of School Effectiveness and Improvement*, (pp. 223-242), New York, Springer.

Daniels, H. (2001) *Vygotsky and Pedagogy*. London: Routledge.

Dehaene, S. (1997). The number sense. New York: Oxford University Press, 1997; Cambridge (UK): Penguin press.

Deng, Z., & Treiman, D. J. (1997). The Impact of the Cultural Revolution on Trends in Educational Attainment in the People's Republic of China. *American Journal of Sociology, 103* (2), 391-428.

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., Stock, P., Schepens, A., Baeyens, & D. Roeyers, H., (2009). Classification, seriation, and counting in grade 1, 2 and 3 as two-year longitudinal predictors for low achieving in numerical facility and arithmetical achievement? *Journal of Psychoeducational Assessment*, 26, 252-264.

Dong, Q., & Lin, C. D. (2011). *Zhongguo ertong qingshaonian xinli fayu tezheng diaocha xiangmu jishu baogao* [*Characteristics of Chinese children and adolescent psychological development survey: Technical Report*]. Beijing: Science publisher.

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

Efklides, A. (2008). Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist, 13*, 277-287.

Foong, P. Y. (1993). Development of a framework for analyzing mathematical problem solving behaviours. *Singapore Journal of Education, 13*(1), 61-75.

Garza, R. (2009). Latino and white high school students' perceptions of caring behaviors: Are we culturally responsive to our students? *Urban Education, 44*, 297-321.

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.

Grantham-McGregor S, Cheung Y, Cueto S, Glewwe P, Richter L,& Strupp B. (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet,* 369: 60–70.

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.

Ho, I. T. & Hau, K. T. (2008). Academic achievement in the Chinese context, The role of goals, strategies, and effort. International Journal of Psychology 43(5), 892-897.

Huang, D. & O'Neil, Jr., H. F. (1997). The Role of Parental Expectation, Effort, and Self-Efficacy in the Achievement of High and Low Track High School Students in Taiwan. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).

Huang, R. J. & Leung, K. S. F. (2005). Cracking the paradox of Chinese learners: looking into the mathematics classrooms in Hong Kong and Shanghai. In Fan, L.H. (Eds). How Chinese learn mathematics: perspectives from insiders.(pp. 348-381) World Scientific Publishing Company.

Imbo, I., & LeFevre, J-A. (2009). Cultural differences in complex addition: Efficient Chinese versus adaptive Belgians and Canadians. *Journal of Experimental Psychology, 35* (6), 1465-1476.

Imbo, I., & Vandierendonck, A. (2007). Do multiplication and division strategies rely on executive and phonological working memory resources? *Memory & Cognition, 35*, 1759-1771.

Kohn, A. (2006). *The Homework Myth: why our kids get too much of a bad thing*. Cambridge, MA.: Da Capo Press.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*, 147-166.

Li, Chun Lin. 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, Chun Lin. 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. (2007). Characteristics of Mathematics Teaching in Shanghai, China: Through the lens of a Malaysian. *Mathematics Education Research Journal, 19*(1), 77-89.

Marton, F., Dall’Alba, G., & Lai, K. T. (1993). The paradox of the Chinese learner (Occasional Paper 93.1). Melbourne, Australia: Educational Research and Development Unit, Royal Melbourne Institute of Technology.

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

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

Perry, L. B. and McConney, A. (2010). Does the SES of the School Matter? An Examination of Socioeconomic Status and Student Achievement Using PISA 2003.” *Teachers College Record, 112(4)*: 1137-1162. Retrieved from http://www.tcrecord.org ID Number: 15662, Date Accessed: 6/1/2010 11:49:34 AM.

Scheerens, J., & Creemers, B. P. M. (1989). Conceptualizing school effectiveness. *International Journal of Educational Research, 13(7),* 691-706.

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.

Staub, F. C., & Stern, E. (2002). The nature of teachers’ pedagogical content beliefs matters for students’ achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology, 2:* 344-355.

Stevenson, H. W., & Lee, S. Y. (1995). The East Asian version of whole-class teaching. *Education Policy*, *9*, 152-168.

Sue, S. & Okazaki, S. (1990). Asian-American educational achievements: a phenomenon in search of an explanation. *American Psychologist, 43*, 913-920.

Swanson, H. L., & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research, 71*, 321–363.

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.

Vygotsky, L.S. (1978). *Mind in society*, Cambridge: Harvard Univ. Press, MA

Walberg, H. J., Haertel, GD, Pascarella, E., Junker, LK, & Boulanger, FD (1981). Probing a model of educational productivity in science with national assessment samples of early adolescents. *American Educational Research Journal*, *18*, 233-249.

Walberg, H. J., Pascarella, E., Haertel, G. D., Junker, L. K., & Boulanger, F. D. (1982). Probing a model of educational productivity in high-school science with national assessment samples. *Journal of educational psychology, 74*(3), 295-307.

Watkins, D. A., & Biggs, J. B. (2001). The paradox of the Chinese learner and beyond. In D. A. Watkins & J. B. Biggs (Eds). The Chinese Learner: Cultural, psychological, and contextual influences (pp.3-26). Hongkong: Comparative Education Research Center, The University of Hong Kong; Melbourne, Australia: Australian Council for Education Research.

Watkins, D. A., & Biggs, J. B. (2001). The paradox of the Chinese learner and beyond. In D. A. Watkins & J. B. Biggs (Eds). The Chinese Learner: Cultural, psychological, and contextual influences (pp.3-26). Hongkong: Comparative Education Research Center, The University of Hong Kong; Melbourne, Australia: Australian Council for Education Research.

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

Xin, T. (2010). International Conference on Educational Monitoring and Evaluation, Newsletter 2, 7, 6-7.

Xin, T., Tian, W., & Zou, Z. (2010). Jiaoyujieguo gongping de celiang jiqi dui jichu jiaoyu fazhan de qishi [The measurement of educational output equity and its implications to the development of elementary education]. *TsingHua Journal of Education, 31*, 21-26, 83.

Yackel, E., Cobb, P., & Wood, T. (1991) Small-group interactions as a source of learning opportunities in second-grade mathematics. *Journal for Research in Mathematics Education, 22* (5), 390-408.

Zimmerman, B.J., & Kitsantas, A. (2005). Students' perceived responsibility and completion of homework: The role of self-regulatory beliefs and processes. *Contemporary Educational Psychology, 397-417.*

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

Appendix - For Chapter 1

International reference (120)

Au, T. K., & Harackiewicz, J.M. (1986). The Effects of Perceived Parental Expectations on Chinese Childrens Mathematics Performance. *Merrill-Palmer Quarterly-Journal of Developmental Psychology, 32*(4),, 383-392.

Aunio, P., Aubrey, C., Godfrey, R., Pan, Y. & Liu, Y.. (2008). Children's Early Numeracy in England, Finland and People's Republic of China. *International Journal of Early Years Education*,*16*(3), 203-221.

Behr, G. (1983). Instructional Accomplishment Patterns of Asian/Pacific Islander Elementary Students. ERIC: ED251500.

Bempechat, J. , & Drago-Severson, E. (1999). Cross-national differences in academic achievement, Beyond etic conceptions of children's understandings. *Review of Educational Research, 69*,(3), 287-314.

Bempechat, J., Graham, S. E., Jimenez, N. V. (1999). The socialization of achievement in poor and minority students - A comparative study. *Journal of Cross-Cultural Psychology 30,*, 139-158.

Benjamin, J. (2006).Cross Cultural Comparison of Rural Education Practice in China, Taiwan, and the United States. Paper presented at the Annual Meeting of the American Educational Research Association (Apr 2006). ED493744.

Bodovski, K., &. Durham, R.E. (2010). Parental Practices and Achievement of Mexican and Chinese Immigrant Children in the USA, Assimilation Patterns? *Research in Comparative and International Education.,5*(2), 156- 175 .

Brenner, M. E., Herman, S., Ho, H-Z, & Zimmer, J. M.(1999). Cross-national comparison of representational competence. Journal for Research in Mathematics Education, 30,, 541-557.

Cai, J. (1998). An Investigation of U.S. and Chinese Students' Mathematical Problem Posing and Problem Solving. *Mathematics Education Research Journal, 10*,(1),37-50.

Campbell, J. R. E. (1994). Differential Socialization in Mathematics Achievement, Cross-National and Cross-Cultural Perspectives. *International Journal of Educational Research. 21*(7), 667-747.

Cantwell, Z. M. (1966). An exploratory study of the development of more effective testing programs for students in differing cultural backgrounds. ERIC: ED010246.

Cao, Z., Forgasz, H., & Bishop, A.(2007). Perceived Parental Influence on Mathematics Learning: A Comparison among Students in China and Australia. *Educational Studies in Mathematics. 64*, 85-106.

Cao, Z., H. Forgasz, et al. (2005). A Comparison of Perceived Parental Influence on Mathematics Learning among Students in China and Australia. In Chick, H. L. & Vincent, J. L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education*, Vol(2), (pp. 209-216). Melbourne: PME.

Chan, B. M.-y. and Ho, C. S.-h. (2010) The Cognitive Profile of Chinese Children with Mathematics Difficulties. *Journal of Experimental Child Psychology, 107*(3), 260-79.

Chan, D. W. (2001). Assessing giftedness of Chinese secondary students in Hong Kong: a multiple intelligences perspective. *High Ability Studies, 12*,(2), 215-234.

Chang, P.T. (1984). A Comparative Study of Mathematics Education between the Province of Taiwan, Republic of China and the United States. ED248142.

Chao, S. J., Stigler, J. W. & Woodward, J. A. (2000). The effects of physical materials on kindergartners' learning of number concepts. *Cognition and Instruction, 18*(3), 285-316.

Chen, C. and Stevenson, H. W. (1995). Motivation and Mathematics Achievement, A Comparative Study of Asian-American, Caucasian-American, and East Asian High School Students. *Child Development. 66 (4)*, 1215-1234.

Chen, C. S., Lee, S. Y., Stevenson, H. W.(1996). Long-term prediction of academic achievement of American, Chinese, and Japanese adolescents. *Journal of Educational Psychology 88*(4), 750-759.

Chen, J. J. L. (2005). Relation of academic support from parents, teachers, and peers to Hong Kong adolescents' academic achievement, The mediating role of academic engagement. Genetic Social and General Psychology Monographs 131(2), 77-127.

Chen, X. Y., Rubin, K. H., Li, D. (1997). Relation between academic achievement and social adjustment:, Evidence from Chinese children. Developmental Psychology 33(3), 518-525.

Cheung, C. K. & Tang, K. L. (2010). Socialization Factors Conducive to Social Work Students' Gain in Competence Experience in the Hong Kong SAR, China. *Journal of Social Work, 10,*(1), 42-58.

Cheung, C. S. & McBride-Chang, C. (2008). Relations of perceived maternal parenting style, practices, and learning motivation to academic competence in Chinese children. *Merrill-Palmer Quarterly-Journal of Developmental Psychology, 54(*1), 1-22.

Chung, K. K. H. and Tam, Y. H. (2005). Effects of cognitive-based instruction on mathematical problem solving by learners with mild intellectual disabilities. *Journal of Intellectual & Developmental Disability, 30*(4), 207-216.

Crystal, D. S., Chen, C., Fuligni, A. J., Stevenson, H. W., Hsu, C.-C., Ko, H.- J., Kitamura, S., & Kimura, S. (1994). Psychological maladjustment and academic achievement: A cross-cultural study of Japanese, Chinese, and American high school students. *Child Development, 65*, 738-753.

D'Ailly, H. (2003). Children's autonomy and perceived control in learning, A model of motivation and achievement in Taiwan. *Journal of Educational Psychology 95*(1), 84-96.

D'Ailly, H. (2004). The role of choice in children's learning, A distinctive cultural and gender difference in efficacy, interest, and effort. *Canadian Journal of Behavioural Science-Revue Canadienne Des Sciences Du Comportement 36*(1), 17-29.

Dailly, H. H. (1992). Asian Mathematics Superiority - a Search for Explanations. *Educational Psychologist, 27*(2), 243-261.

Daniel, L. M., S. B. Lim, et al. (2003). Eight-year outcome of very-low-birth-weight infants born in KK hospital. *Annals Academy of Medicine Singapore 32*(3), 354-361.

Derevensky, J. L., Hart, S. & Farrel, M. (1981). An Examination of Achievement Related Behavior of High and Low Achieving Inner City Pupils. Psychology in the Schools, 20 (3), 328-336.

Dong, Q., Zhou,X. L.， Chen, C. S., Jiang, T., Zhang, H. C. & Dong, Q. (2007). Chinese kindergartners' automatic processing on numerical magnitude in Stroop-like tasks. *Memory & Cognition 35*(3), 464-470.

Fuligni, A. J. & Stevenson, H. W. (1995). Time Use and Mathematics Achievement among American, Chinese, and Japanese High-School-Students. *Child Development 66*(3), 830-842.

Geary, D. C., Bowthomas, C. C., Fan, L., & Siegler, R. S.(1993). Even before Formal Instruction, Chinese Children Outperform American Children in Mental Addition. *Cognitive Development 8*(4), 517-529.

Geary, D. C., BowThomas, C. C., Fan, L., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children, Influence of age, language, and schooling. *Child Development 67*(5), 2022-2044.

Geary, D. C., Fan, L., Bow-Thomas, C. C. (1992). Numerical Cognition - Loci of Ability Differences Comparing Children from China and the United-States. *Psychological Science, 3*(3), 180-185.

Geary, D. C., Salthouse, T. A., Chen, G. P. & Fan, L. (1996). Are East Asian versus American differences in arithmetical ability a recent phenomenon? *Developmental Psychology, 32*(2), 254-262.

Ginsburg, H. P., Lin, C-L., Ness, D., & Seo, K-H. (2003). Young American and Chinese Children's Everyday Mathematical Activity. *Mathematical Thinking and Learning. 5* (4), 235-258.

Hao, L. X. and Bonstead-Bruns, M. (1998). Parent-child differences in educational expectations and the academic achievement of immigrant and native students. *Sociology of Education, 71*(3), 175-198.

Hess, R. D., Chang, C. M. & McDevitt, T. (1987). Cultural Variations in Family Beliefs About Childrens Performance in Mathematics - Comparisons among Peoples-Republic-of-China, Chinese-American, and Caucasian-American Families. *Journal of Educational Psychology, 79* (2), 179-188.

Ho, C. S. H. & Fuson, K. C. (1998). Children's knowledge of teen quantities as tens and ones: Comparisons of Chinese, British, and American kindergartners. *Journal of Educational Psychology, 90*,(3), 536-544.

Ho, C. S. H. &Cheng, F. S. F. (1997). Training in place-value concepts improves children's addition skills. *Contemporary Educational Psychology, 22*(4), 495-506.

Ho, H. Z., Senturk, D., Lam, A. G., Zimmer, J. M., Hong, S. Okamoto, Y., Chiu, S.Y., Nakazawa, Y. Wang, C. P. (2000). The affective and cognitive dimensions of math anxiety, A cross-national study. *Journal for Research in Mathematics Education 31*(3), 362-379.

Ho, I. T. & Hau, K. T. (2008). Academic achievement in the Chinese context, The role of goals, strategies, and effort. International Journal of Psychology 43(5), 892-897.

Hong, E. & Lee., K.-h. (2000). Preferred Homework Style and Homework Environment in High- versus Low-Achieving Chinese Students. Educational Psychology, An International Journal of Experimental Educational Psychology. 20(2),125-137.

Hong, E. S., Pend, Y. , Rowell, L. L. (2009). Homework self-regulation, Grade, gender, and achievement-level differences. Learning and Individual Differences 19(2), 269-276.

House, J. D. (2002). Instructional Practices and Mathematics Achievement of Adolescent Students in Chinese Taipei, Results from the TIMSS 1999 Assessment. Child Study Journal.32(3),157-178.

Hsi, V. & Lim, V. (1977). A Summary of Selected Research on Cognitive and Perceptual Variables. Asian American Bilingual Center, 2168 Shattuck Avenue, 3rd Floor, Berkeley, California 94704

Huang, D. & O'Neil, Jr., H. F. (1997). The Role of Parental Expectation, Effort, and Self-Efficacy in the Achievement of High and Low Track High School Students in Taiwan. Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, March 24-28, 1997).

Huang, W. (1999). Reasoning about conventional time as a function of conventional time systems. Memory & Cognition 27(6), 1080-1086.

Huntsinger, C. S. & Jose, P. E. (1993). Ethnic Differences in Mathematics Teaching Styles, Chinese-American and Caucasian-American Mother-Father-Daughter Triads. Paper presented at the Annual Meeting of the Society for Research in Child Development (60th, New Orleans, LA, March 25-28, 1993). ED365544.

Huntsinger, C. S. & Jose, P. E. (1995). Chinese-American and Caucasian American Family-Interaction Patterns in Spatial Rotation Puzzle Solutions. *Merrill-Palmer Quarterly-Journal of Developmental Psychology, 41*(4), 471-496.

Huntsinger, C. S. and P. E. Jose (2006). A longitudinal investigation of personality and social adjustment among Chinese American and European American adolescents. Child Development 77(5), 1309-1324.

Huntsinger, C. S., Jose, P. E., & Ching, W. D. (1994). Ethnic differences in early math learning: A comparison of Chinese-American and Caucasian-American families. Presented at the Annual Symposium of the Jean Piaget Society, Chicago, IL. (ERIC Document Reproduction Service No. 375940).

Huntsinger, C. S., Jose, P. E., Larson, S. & Ching, W-D. (1996). Different Cultures, Different Competencies, A Comparison of Chinese-American and Euro-American First and Second Grade Children. Paper presented at the Biennial Conference on Human Development (14th, Birmingham, AL, March 29-31, 1996. (ED399029).

Huntsinger, C. S., Jose, P. E., Larson, S. L., Krieg, D. B., & Shaligram, C. (2000). Mathematics, vocabulary, and reading development in Chinese American and European American children over the primary school years. *Journal of Educational Psychology, 92*, 745–760.

Huntsinger, C. S., Jose, P.E., Liaw, F.R., & Ching, W. (1997). Cultural differences in early mathematics learning, A comparison of Euro-American, Chinese-American, and Taiwan-Chinese families. International Journal of Behavioral Development 21(2), 371-388.

Huntsinger, C. S., Larson, S. L., Krieg, D. B. (1998). Mathematics and Vocabulary Development in Chinese American and European American Children over the Primary School Years. Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998). ED422445.

Hutsinger, C. S. & Jose, P. E. (1995). Chinese American and Caucasian American Family Interaction Patterns in Spatial Rotation Puzzle Solutions. Merrill-Palmer Quarterly.

Imbo, I. and J. A. LeFevre (2009). Cultural Differences in Complex Addition, Efficient Chinese Versus Adaptive Belgians and Canadians. *Journal of Experimental Psychology-Learning Memory and Cognition*, 35(6), 1465-1476.

Kinlaw, C. R., B. Kurtz-Costes, et al. (2001). Mothers' achievement beliefs and behaviors and their children's school readiness, A cultural comparison. *Journal of Applied Developmental Psychology, 22*(5), 493-506.

Kit-fong Au, T. & Harackiewicz, J. M. (1986). The Effects of Perceived Parental Expectations on Chinese Children's Mathematics Performance. *Merrill-Palmer Quarterly: Journal of Developmental Psychology, 41* (4), 471-496.,

Kong, C.-K., Hau, K.-T., Cheng, Z-J. (1998). Chinese Students' Self-Concept and Academic Performance, Big-Fish-Little-Pond Effects and the Role of Perceived School Status.Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998). ED419213.

Kwok, D. C. &Lai, D. W. (1993). The Self-Perception of Competence by Canadian and Chinese Children. Paper presented at the Annual Convention of the Canadian Psychological Association (Montreal, Quebec, Canada, May, 1993). ED358963.

Kwok, D. C. &Lytton, H. (1996). Perceptions of mathematics ability versus actual mathematics performance, Canadian and Hong Kong Chinese children. *British Journal of Educational Psychology, 66*, 209-222.

Lee, E. W. (1989). Chinese American Fluent English Proficient Students and School Achievement. *NABE, The Journal for the National Association for Bilingual Education. 13*(2), 95-111.

Leu, Y.-C. & Wu, C.-J. (2004). The Mathematics Pedagogical Values Delivered by an Elementary Teacher in Her Mathematics Instruction: Attainment of Higher Education and Achievement. Proceedings of the 28th conference of the International Group for the Psychology of Mathematics Education, 3, 225-232.

Lin, E. & Wang, J. A. (2009). A meta-analysis of comparative studies on Chinese and US students' mathematics performance: Implications for mathematics education reform and research. Educational Research Review 4(3), 177-195.

Liu, C. F., Zhang, L. X., Luo, R., Rozelle, S., & Loyalka, P.(2010). The effect of primary school mergers on academic performance of students in rural China. *International Journal of Educational Development 30*(6), 570-585.

Luo, Z., Jose, P. E., Huntsinger, C, S. & Pigott, T. D. (2007). Fine motor skills and mathematics achievement in East Asian American and European American kindergartners and first graders. *British Journal of Developmental Psychology, 25*, 595-614.

Ly, T. M. (2008). Asian American parents' attributions of children with down syndrome, Connections with child characteristics and culture. *Intellectual and Developmental Disabilities, 46*(2), 129-140.

Lynn, R. (2008). IQ and mathematics ability of Tibetans and Han Chinese. *Mankind Quarterly, 48*,(4), 505-510.

Marjoribanks, K. & Kwok, Y. (1998). Family capital and Hong Kong adolescents academic achievement. *Psychological Reports, 83*(1), 99-105.

Marsh, H. W., Hau, K. T. & Kong, C-K(2002). Multilevel causal ordering of academic self-concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students. American Educational Research Journal 39(3), 727-776.

Mau, W. C. (1997). Parental influences on the high school students' academic achievement: A comparison of Asian immigrants, Asian Americans, and white Americans. *Psychology in the Schools 34*(3), 267-277.

Miller, K. F., Smith, C. M., Zhu, J. J. & Zhang, H. C. (1995). Preschool Origins of Cross-National Differences in Mathematical Competence - the Role of Number-Naming Systems. *Psychological Science, 6*,(1), 56-60.

Miura, I. T., Okamoto, Y., Kim, C. C., Steere, M., & Fayol, M. (1993). First -Graders Cognitive Representation of Number and Understanding of Place Value - Cross-National Comparisons - France, Japan, Korea, Sweden, and the United-States. *Journal of Educational Psychology, 85,* (1), 24-30.

Miura, I. T., Okamoto, Y., Kim, C. C., Steere, M., & Fayol, M. (1994). Comparisons of Children Cognitive Representation of Number - China, France, Japan, Korea, Sweden, and the United-States. *International Journal of Behavioral Development, 17*(3), 401-411.

Muldoon, K., Simms, V., Towse, J., Menzies, V. & Yue, G. A.(2011). Cross-Cultural Comparisons of 5-Year-Olds' Estimating and Mathematical Ability. *Journal of Cross-Cultural Psychology, 42*, (4), 669-681.

Ng, F. F. Y., Pomerantz, E. M., & Lam, S-F (2007). European American and chinese parents' responses to children's success and failure: Implications for children's responses. *Developmental Psychology, 43*(5), 1239-1255.

Ng, S. S. N. & Rao, N. (2010). Chinese Number Words, Culture, and Mathematics Learning. *Review of Educational Research, 80*(2), 180-206.

Ng, W. H. A., Chan, Y. L., Au, K. S. A., Yeung, K. W. D., Kwan, T. F. & To, C. Y.(2005). Morphometry of the corpus callosum in Chinese children: relationship with gender and academic performance. *Pediatric Radiology, 35*(6), 565-571.

Ngwudike, B. C. (2009). Competitive Edge: A Cross-National Examination of Mathematics Achievement in 53 Jurisdictions. ED504433 http://ghaa.jsums.edu/journals/articles/CompetitiveEdge.pdf

O'Connor, C., Hill, L. D., Robinson, S. R.(2009). Who's at Risk in School and What's Race Got to Do With It? Review *of Research in Education, 33*,(1), 1-34.

Pan, Y. Q., Gauvain, M., Liu, Z. K. & Cheng, L. (2006). American and Chinese parental involvement in young children's mathematics learning. *Cognitive Development, 21*(1), 17-35.

Parmar, P., Harkness, S., Super, C. M.(2004). Asian and Euro-American parents' ethnotheories of play and learning: Effects on preschool children's home routines and school behaviour. *International Journal of Behavioral Development, 28*(2), 97-104.

Parmar, P., Harkness, S., Super, C. M.(2008). Teacher or playmate? Asian immigrant and Euro-American parents' participation in their young children's daily activities. *Social Behavior and Personality, 36*(2), 163-175.

Pearce, R. R. & Lin, Z. (2007). Chinese American post-secondary achievement and attainment: a cultural and structural analysis. *Educational Review, 59*(1), 19-36.

Perry, M., Vanderstoep, S. W., Yu, S. L.(1993). Asking Questions in 1st-Grade Mathematics Classes - Potential Influences on Mathematical Thought. *Journal of Educational Psychology, 85*(1), 31-40.

Peverly, S. T. (2005). Moving past cultural homogeneity: Suggestions for comparisons of students' educational outcomes in the United States and China. *Psychology in the Schools, 42*(3), 241-249.

Pezdek, K., Berry, T. & Renno, P. A.(2002). Children's mathematics achievement: The role of parents' perceptions and their involvement in homework. *Journal of Educational Psychology, 94*(4), 771-777.

Phillipson, S. & Phillipson, S. N. (2007). Academic Expectations, Belief of Ability, and Involvement by Parents as Predictors of Child Achievement, A Cross-Cultural Comparison. *Educational Psychology, 27* (3), 329-348.

Pong, S.-L. (2009). Grade Level and Achievement of Immigrants' Children, Academic Redshirting in Hong Kong. *Educational Research and Evaluation. 15*, (4), 405-425.

Post, D. & Pong, S. L. (2009). Student labour and academic proficiency in international perspective. *International Labour Review, 148,(*1-2): 93-122.

Salili, F. & Hau, K. T.(1994). The Effect of Teachers Evaluative Feedback on Chinese Students Perception of Ability - a Cultural and Situational Analysis. Educational Studies, 20(2), 223-236.

Schleppenbach, M., Flevares, L. M., Sims, L. M. & Perry, M. & (2007). Teachers' responses to student mistakes in Chinese and US mathematics classrooms. *Elementary School Journal, 108*(2), 131-147.

Schleppenbach, M., Perry, M., Miller, K. F., Sims, L. & Fang, G.(2007). The answer is only the beginning, Extended discourse in Chinese and US mathematics classrooms. *Journal of Educational Psychology, 99*(2), 380-396.

Siegler, R. S. & Mu, Y. (2008). Chinese children excel on novel mathematics problems even before elementary school. *Psychological Science, 19*(8), 759-763.

Stevenson, H. W. & Lee, S. Y. (1995). The East-Asian Version of Whole-Class Teaching. Educational Policy, 9(2), 152-168.

Stevenson, H. W., Chen, C. S., Lee, S.(1993). Motivation and Achievement of Gifted-Children in East-Asia and the United-States. *Journal for the Education of the Gifted, 16*(3), 223-250.

Stevenson, H. W., Lee, S. Chen, C.S., Kato, K. & Londo, W.(1994). Education of Gifted and Talented Students in China, Taiwan, and Japan. In: Ross, Patricia O' Connell, Ed. National Excellence: A Case for Developing America's Talent. An Anthology of Readings; see EC 303 213. ED372582.

Stevenson, H. W., Stigler, J. W., Lee, S-Y, Lucker, W. G. Kitamura, S. & Hsu, C-C. (1985). Cognitive Performance and Academic Achievement of Japanese, Chinese, and American Children. *Child Development, 56*(3), 718-734.

Towse, J. N. & Saxton, M. (1997). Linguistic influences on children's number concepts, Methodological and theoretical considerations. *Journal of Experimental Child Psychology, 66*, (3), 362-375.

Tsang, S.-L. (1976). The Effects of the Language Factor and the Cultural Content Factor of Mathematics Achievement Test on Chinese and Chicano Students. University Microfilms, Dissertation Copies, P.O. Box 1764, Ann Arbor, Michigan 48106 (Order No. 76-18,808). ED146255.

Tsui, M. & Rich, L. (2002). The only child and educational opportunity for girls in urban China. *Gender & Society, 16*,(1), 74-92.

Tsui, M. (2005). Family income, home environment, parenting, and mathematics achievement of children in China and the United States. *Education and Urban Society, 37*(3), 336-355.

Uttal, D. H., Lummis, M., Stevenson, H. W.(1988). Low and High Mathematics Achievement in Japanese, Chinese, and American Elementary-School-Children. *Developmental Psychology, 24*(3), 335-342.

Wang, C. L., Chuang, H. Y., Ho, C-K, Yang, C-Y, Tsai, J-L., Wu, T-S., & Wu, T-N. (2002). Relationship between blood lead concentrations and learning achievement among primary school children in Taiwan. *Environmental Research, 89*(1), 12-18.

Wang, J. & Lin, E. (2005). Comparative Studies on U. S. and Chinese Mathematics Learning and the Implications for Standards-Based Mathematics Teaching Reform. Educational Researcher.34(5), 3-13.

Wang, J. & Lin, E. (2008). An Alternative Interpretation of the Relationship between Self-Concept and Mathematics Achievement, Comparison of Chinese and US Students as a Context. Evaluation and Research in Education, 21 (3), 154-174.

Wang, J. (2006). An Empirical Study of Gender Difference in the Relationship between Self-Concept and Mathematics Achievement. *Educational Psychology, 26* (5), 689-706.

Wang, J. (2007). A Trend Study of Self-Concept and Mathematics Achievement in a Cross-Cultural Context. *Mathematics Education Research Journal, 19* (3), 33-47.

Xu, J. & Farrell, E. (1992). Mathematics Performance of Shanghai High School Students: A Preliminary Look at Gender Difference in Another Culture. School Science and Mathematics.92 (8), 442–444.,

Yang, M. T.-L. & Cobb, P. (1995). A Cross-Cultural Investigation into the Development of Place-Value Concepts of Children in Taiwan and the United States. *Educational Studies in Mathematics, 28* (1), 1-33.

Zang, Z. X., Han, Z. Z. & Zang, Y. F.(2008). Effects of hand clasping and arm folding on academic performance. *Neural Regeneration Research, 3*(8), 914-917.

Zhang, L.-F. (2007). Intellectual Styles and Academic Achievement among Senior Secondary School Students in Rural China. *Educational Psychology, 27* (5), 675 – 692.

Zheng, Z., Peverly, S. T., Boehm, A. E, & Lin, C. D. et al. (2000). American and Chinese children's understanding of distance, time, and speed interrelations. Cognitive Development, 15(2), 215-240.

Zhou, M., Peverly, S. T., & Lin, J. S.(2005). Understanding early mathematical competencies in American and Chinese children. *School Psychology International, 26*(4), 413-427.

Zhou, Z., Cheng, C., Mottram, L. & Rosenblum, S. (1999). Understanding Informal and Formal Mathematical Abilities in Mainland Chinese and Chinese-American Children. Paper presented at the Biennial Meeting of the Society for Research in Child Development (Albuquerque, NM, April 15-18, 1999). ED429846.

Zhou, Z., Peverly, S. T., Lin, C. D. (2004). Cross- and within-cultural variations in children's understanding of distance, time, and speed interrelationships, A follow-up study. *Journal of Genetic Psychology, 165*(1), 5-27.

Zhou, Z., Peverly, S. T., Lin, J. S. (2005). Understanding Early Mathematical Competencies in American and Chinese Children. School Psychology International, 26 (4), 413-427.

National reference (110)

An, H. Y. (2006). gaoyi xuesheng ziwo xiaonenggan, fangyuxing beiguan,, ziwo fang'ai yu shuxue chengji de guanxi [Relationship on the self-efficacy, defensive pessimist, self-handicapping and mathematics performance for grade 10 students]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Bao, M., Zhang, H. W., & Guo, H. Y. (2008). Haerbin sili zhongxue xuesheng shuxue chengjiu dongji diaocha yanjiu [Investigating Research on Math Achievement

Motivation of Private High School Students]. *Shuxue jiaoyu xuebao, 17*(1), 51-55.

Chen, X. B. (2007). daxuesheng ziwo jiankong nengli yu gaodeng shuxue xueye chengji de guanxi yanjiu [Relationship between the self-control ability and the mathematics performance in the university]. *Journal of Yuxi Normal University Journal*, 12, 83-86.

Chen, Z. G. (2009). putong gaozhong yousheng shuxue xuexi chengji fenhua chengyin fenxi [Reasons for the higher performance students who start to become lower achiever for the secondary school students](Unpublished Master’s thesis). Hebei Normal University, Shijiazhuang, P.R. China.

Diao, C. H. (2010). daxuesheng shuxueguan yu gaodeng shuxue xueye chengji de xiangguan yanjiu [Relationship between mathematic beliefs and the learning performance for the university students]. *Journal of Yichun college. 32*, (8), 157-158.

Ding, Y. Y., & Shen, Z. J. (2008). shuxue yuanrenzhi, xueye chengjiu yu shuxue wenti tichu sanzhe guanxi de yanjiu [Relationship between math metacognition, learning performance and problem asking]. *Anhui guangbo dianshi daxue xuebao, 1,* 55-59

Ding, Z. B. (1988). shilun bian kewai buke wei kenei buke de kexingxing - tigao zhongchasheng shuxue chengji de shiyanbaogao [The experimental report for the project of "improving the students with middle and lower achievement"-the possiblity of the tranforming from the exact courses into normal courses]. *Jiangsu jiaoyu, 16,* 28-30.

Duan, Y. W. (2006). nannvsheng shuxue chengji bodong chayi de guiyin fenxi [The attribution of the frustration of the mathematics performance for the boys and girls]. *Shuxue jiaoxue tongxun, 12*, 7-9.

Fang, G., Tong, L. Q., Liu, F. (1988). sanguosichengshi xiaoxue ertong shuxue chengji de bijiao yanjiu: kuawenhua yanjiuzhiyi [Comparing the mathematics performance for four cities in three countries: a cross -cuture study]. *Acta Psychologica Sinica, 3,* 268-278.

Fang, X. W.(2010). yingxiang xuesheng daxue shuxue chengji yinsu de tongji fenxi: yi huzhou shifan xueyuan wei yangben [The statistics analysis on the variables which affect the mathematics performance for the university students of Huzhou Teacher college]. *Gaodeng hanshou xuebao (ziran kexueban). 23*, (5), 49-51.

Fei, Y. T.(2011). xizang chuzhongxuesheng shuxue xuexi chengji de diaocha yu sikao [Tibet's middle school students' mathematics performance and reflection]. *xizhang keji, 2,* 44-45.

Gao, R. (2002). chuer xuesheng shuxue jiaolv yu shuxue xuexii chengji guanxi de diaocha fenxi [Survey on the impact of math anxiety on the mathematics learning performance for grade 8 students]. *zhongxue jiaoyan: shuxueban, 11*, 19-20.

Gou, Z. Q.(2010). gaozhongsheng xuexi taidu, fumu shoujiaoyu chengdu yu shuxue chengji de yingxiang yanjiu [Relationship of learning attitude, parental educational level and mathematics learning performance]. *Journal of Cangzhou Teacher college, 26,* (3), 76-77.

Gu, Z. G.. (2007). zhongxuesheng shuxue xueke zhi chengjiu mubiao, ziwo xiaoneng ,xuexi celue ji xueye chengji guanxi yanjiu [The Relationships Study of Achievement Goals, Self-Efficacy, Learning Strategies, and Academic Achievement in Mathematics Course] (Unpublished Master's thesis). Suzhou University, Suzhou, P.R.China.

Guo, Y. Y., Zheng, G. P., He, S. Q, & Yang, X. J. (2006). Shuxue chengji de yingxiang yinsu diaocha fenxi ji jiaoxue duice tantao [The Diagnoses and Teaching Methods on the Influential Factors of the Mathematical Grade]. Journal of Zhaotong Teacher College, 28 ,(5), 50-52.

Hao, J. Y. (2004). shuxue renzhi jiegou yu shuxue xuexi chengji xiangguanxing de yanjiu [Relationship between the mathematics cognitive structure and mathematics learning performance]. (Unpublished Master's thesis). Nanjing Normal University, Nanjing, P.R. China.

He, X. Y.(1998), xiaoxuesheng shuxue ziwo xiaonenggan, ziwo gainian yu shuxue chengji guanxi de yanjiu [Research on the relationship between students' self-efficacy, self-concept and mathematics performance for the primary students*]. xinli fazhan yu jiaoyu, 1*, 45-48.

Huang, D. Q. (2006). gaozhongsheng de shuxue jiaolv yu shuxue ziwo xiaonenggan dui shuxuechengji de yingxiang [Research on the Effects of Mathematics Anxiety and Its Self-efficacy on Mathematics Achievement of Senior High School Students]. (Unpublished Master's thesis). Hebei University, Baoding, P.R.China.

Ji, X. Y. (2009). gaozhongxuesheng shuxue jianmo nengli yu shuxue xueye chengji guanxi de ciaocha fenxi [Survey on and Analysis of "Relationship of Highschool Student's Modeling Ability and Math Learning Grades]. *shanghai jiaoyu keyan, 7*, 57-58.

Jiang, B. (2006). jingzheng-hezuo xuexi dui xiaoxuesheng shuxue chengji de shiyanyanjiu [An experimental research on the competitive-cooperative learning's effects on students' mathematics achievements in elementary schools *]. jiaoyu daokan, 11*, 32-33.

Kong, L. Y. (2002). Research on Instruct-reading Teaching Model in the New Period [shuxue jiaolv yu shuxuechengji guanxi de yanjiu]. (Unpublished Master's thesis). Capital Normal University, Beijing, P.R.China.

Lai, R. S..(2005). yangli chengxian fangshi yu renzhi fangshi dui gaozhongsheng shuxue xuexi chengji de yingxiang yanjiu [Research on Effects of Worked-example Presentation Form and Cognitive Style on Mathematics

Learning Performance of Senior High School Students] (Unpublished Master's thesis). Jiangxi Normal University, Nanchang, P.R. China.

Li, Hua. (2006). Junior High Pupil Mathematics Result Polarization Origin and Countermeasure Research [chuzhongsheng shuxue chengji liangji fenhua de chengyin ji duice yanjiu]. (Unpublished Master's thesis). Central China Normal University, Wuhan, P.R. China.

Li, J. W. (2005). gaoyi xuesheng shuxue xueke zizun shuiping ,chengjiu mubiao yu xueye chengji de guanxi yanjiu [A Relationship Study about Level of Math Subject Self-esteem, Achievement Goal and Academic Achievement of Students in Grade One Senior Middle School]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Li, J. Y.. (1989). fumu wenhua chengdu dui xuesheng shuxue xuexii chengji de yingxiang chutan [Impact of parental literacy level on the mathematics performance]. *Journal of Guizhou Normal University (Natural Science) ,2*, 92-95.

Li, L.(2007). gaozhi xuesheng shuxue xuexi xingqu jiqi dui shuxue chengji de yingxiang [The impact of learning interests on the learning performance for the vocational college students]. (Unpublished Master's thesis). Suzhou University, Suzhou, P.R.China.

Li, M. (1998).yige shuxue moxing de qishi: dui dianda puzhuanban xuesheng kaoshi chengji yu zuoye, daoke qingkuang xiangguanxing de yanjiu[An implication for the math model: the research on the relationship between students' performance, homework and attending for the course in the open university]. *yuancheng jiaoyu zazhi, 3*, 30-32.

Li, M. Z. (1994). yizhi pinzhi yu xuesheng shuxue xueye chengji de guanxi yanjiu [Relationship between spiritual quality of mathematics performance]. *shuxue jiaoyu xuebao, 1*, 65-69.

Li, M. Z.. (1994). chengjiu dongji, shuxue xuexi taidu yu xuesheng shuxue xueye chengji de guanxi [Relationship between achievement motivation, math learning attitudes, and mathematics learning performance]. *Journal of Guizhou Normal University (Natural Science), 12* (2), 45-51.

Li, Q., & Ni, Y. Q.(2006). Multilevel Analysis of Effects of Teachers' Variables on Primary School Students' Mathematics Achievements [jiaoshi bianliang dui xiaoxuesheng shuxue xuexi chengji yingxiang de duoshuiping fenxi]. *Jiaoshi jiaoyu yanjiu, 18*(3), 74-80.

Li, S. l. (2008). banji huanjing ,shuxue jiaolv yu yingyongti chengji de guanxi [The Relation between Class Environment and Mathematic Anxiety and Word-problem Score](Unpublished Master's thesis). Qufu Normal University, Qufu, P.R. China.

Li, W. C. (2008). chengjiu dongji, duoyuan hudong yu shuxuexuexi chengji guanxi de shizhengyanjiu [Empirical studies on the achievement motivation, interaction and mathematics performance*]. jiaoyu shijian yu yanjiu (zhongxueban), 6*, 36-40.

Li, W. C. (2008). chengjiu dongji, duoyuan hudong yu shuxuexuexi chengji guanxi de shizhengyanjiu [Emprical studies on the achievement motivation, interaction and mathematics performance]. *shuxue jiaoxue yanjiu, 7*, 2-7.

Li, X. Y., & Li, W. J. (2004). chuzhongsheng shuxue chengjiu dongji yu ziwo tiaojie xuexi ji xueye chengjiu zhiguanxi yanjiu [Relative Study on Middle School Students' Mathematical Achievement Motivation Self-Regulated Learning and Mathematical Achievement]. *Shuxue jiaoyu xuebao, 13*, (2), 52-54.

Li, Y. Z., (2007). daxuesheng shuxuechengjiu dongji ,xuexi celue yu xuexi chengji de guanxi [Relationship between the math achievement goal, learning strategies and learning performance for the undergraduate students]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Li, Z. M.(1996). zhili shuiping, shuxue xuexi dongyin, celue yu xueye chengji de guanxi [Intelligent quality, learning motivation, strategies and learning performance]. *Journal of Guizhou Normal University (Natural science version), 2*, 78-78.

Liang, H. C. (1999). shuxue xuexi taidu, xuexi celue yu shuxue xueye chengji guanxi de yanjiu [Relationship between math learning interests, learning strategies and mathematics learning performance]. *Journal of Guangxi Normal University(Natural Science version), 16*,(4), 111-116.

Liang, H. C. (2001). shuxue qingang, yizhi yu shuxue xuexi chengji guanxi de yanjiu [Relationship between the mathematics emotion, spiritual and learning performance]. *Journal of Guangxi Education college, 3*, 81-84.

Liang, H. C., & Huang, Y. J. (2011). shuxue ziwo jiankong ji xuexi dongji dui shuxue chengji yingxiang jili de yanjiu [Research of the Mechanism for Effects of Mathematics Self-monitoring and Mathematics Academic Motivation on Mathematics Grade]. *Shuxue jiaoyu xuebao, 1*, 58-60.

Liu, J. (2008). qingang yinsu he siwei dui gaozhongsheng shuxue chengji yingxiang de fenxi ji jiejue celue [Effect of emotional factors and thinking style on the mathematics performance for secondary school students]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Liu, L. M., Lin, X. L., & Sang, Y.(2009). xuesheng chengji de kekong yingxiang yinzi fenxi:yi gaodeng shuxue weili [The analysis on the factors which can be controlled and impact on the mathematics performance for high education]. *zhongguo kejiao chuangxin daokan, 19,* 143-143.

Liu, S. R., & Wen, J. (2009). shuxue xueye buliang yu xueye zhengchang xuesheng de chengjiu dongji yanjiu :yi hunansheng hengshanxian gaoshan xuesheng weili [The motivation study on the students with learning difficulties and the Normal students : based on the sample of grade 12 students in Hengshan district in Hunan Province]. *jiaoyu celiang yu pingjia, 1,(*1), 41-43.

Liu, S. X. (2006). gaozhongsheng shuxue xuexi guiyin, xuexi celue yu shuxue xuexi chengji guanxi de yanjiu [Relationship between learning attribution, learning strategies and mathematics performance for the secondary school students]. (Unpublished Master's thesis). Shangdong Normal University, Jinan, P.R. China.

Liu, W. H., (2004). gaozhong shuxue xuexi chengji liangji fenhua xianxiang chansheng de yuanyin ji duice yanjiu [The Research on the Existence of Periodic Solution and the Permanence for a Class of Discrete Population Models]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Liu, X. M., & Zhou, L. (2007). xueye fankui, shuxue ziwo xiaonenggan yu chuyi xuesheng shuxue chengji de guanxi [Relationship bewteeen learning feedback, math self-efficacy and mathematics performance]. *Chinese clinical psychology Journal, 15*,(1), 53-55.

Lu, Z. J. (1992). nongcun chuer xuesheng shuxue xuexi chengjii fenhua yuanyin de duibi diaocha [Survey on the differential performance for the grade 8 students for the rural side]. *shanghai jiaoyu keyan, 4,* 60-61.

Luo, H. T., Zheng, X. M., & Pu, D. Q. (2002). yingxiang zhongxuesheng shuxue chengji yinsu de tongjing fenxi [A path analysis of factors that influence the student mathematics achievement]. *shuli tongji yu guanli. 21* (6), 15-20.

Luo, H., & Lv, Z. G. (2009). chuzhong xuesheng shuxue zizhu xuexi jiqi yu shuxue chengji de guanxi yanjiu[Relationship between the students' self-regulation and their performance for the middle school]. *NanYang shifan xueyuan xuebao, 8*,(12),108-111.

Luo, J. T. (2004). duomeiti jiaoxu dui xiaoxuesheng shuxue chengji, shuxue ziwo xiaonenggan ji shuxe xingqu de yingxiang [The Influence of Multi-Media on Students Academic Performance, Self-Efficacy and Interest in Primary School Math Teaching](Unpublished Master's thesis).South China Normal University, Guangzhou, P.R. China.

Luo, R. S., Shen, J. L., Wang, M. C. (2006). xuexi taidu, fangfa yu zhongxuesheng shuxue chengji de xiangguan yanjiu [Relationsihp of learning attitude, learning methods on the mathematics performance of middle school students].*Baoding Teacher College Journal, 15,*(2), 57-60.

Mei, H. X. (2008). yingxiang nongchun chuzhongsheng shuxue xuexi chengji de xiangguan yinsu de fenxi ji jiaoxue duice yanjiu [Explore the factors which contribute to the mathematics performance and related teaching strategies for rural middle school students]. (Unpublished Master's thesis). Central China Normal University, Wuhan, P.R.China.

Peng, S. G., & Liao, M. (2009). Relationship between the mathematics analysis course scores and the mathematics internships scores for college students who major in mathematics education[shuxue jiaoyu zhuanye shuxue fenxi kecheng chengji yu shixi chengji de xiangguangxing fenxi]. gaodeng hanshou xueyuan xuebao, 22, (1), 49-50.

Qiao, F. Q. (2010). shugan, shuxue xiaonenggan yu shuxue chengji de guanxi [Relationship of Number sense, mathematics self-efficacy and mathematics performance](Unpublished master's thesis). Jinan University, Jinan, P. R. China.

Qiu, X. L., & Chen, L. Q. (2010). xuesheng shuxue chengji xiangguanxing fenxi [Relationship on the mathematics performance]. *Journal of Fujian Normal University(natural science version), 26*, (4), 115-119.

Shen, D. L., Li, H. Y., Zhuang, S. F., Du, H., Hu ,J. J., Zhang, W.(2002). zhongxiaoxuesheng de zhili, xueyetaidu yuqi shuxue xueye chengjiu de xiangguan yanjiu [Relationship between intelligent quality, learning attitude and mathematics learning performance]. *Journal of Tianjing Normal University, 2*, 1-5.

Song, G. W., & Li, S. L. (2009). shuxue jiaolv dui chuzhongsheng shuxue yingyongti chengji de yingxiang [The Influence of Mathematic Anxiety on Middle School Students' Word- problem Score]. *xinlixue tanxin, 29*, (1), 45-49.

Su, J. H. (2005). yunyong xiaozu hezuo xuexi tigao xiaoxuesheng shuxue chengji de shiyan yanjiu [Experimental Research on Enhancing Primary Students' Maths Marks by Group Cooperative Learning]. (Unpublished Master's thesis). Liaoning Normal University, Shenyang, P.R.China.

Sun, G. Z. (2007). zhongxuesheng yuanrenzhi, cuoti guanli yu xuexi chengji de guanxi [Relationship between metacognition, incorrect problems management and learning performance]. (Unpublished Master's thesis). Nanjing Normal University, Nanjing, P.R. China.

Sun, J. L. (2007). shuxue jiaolv, gongzuo jiyi he shuxue chengji zhijian de guanxi [The Relationship among Mathematics Anxiety, Working Memory and Mathematics Achievement]. (Unpublished Master's thesis). Henan University, Kaifeng, P.R.China.

Sun, X. H., & Huang, Y. Y.. (2004). yingxiang xianggang nannvsheng shuxue chengji de qianbianliang chayi [Gender Difference of Mathematics Learning among Hong Kong Students]. *Shuxue jiaoyu xuebao, 13* (4), 72-74.

Sun, Z. Y. (2007). shuxue mei de jiaoyu dui xuesheng shuxue xuexi taidu ,chengji yingxiang de shiyan yanjiu [Effect of teaching mathematics beauty on learning attitudes and learning performance: experiment study]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Tan, R. B. (1996). shaoshu minzu chuerxuesheng feizhili yinsu yu shuxuechengji guanxi de yanjiu [Relationship between non-IQ variables and mathematics performance for the grade 8 minority students]. *Journal of Liuzhou Teachers, 11*(1), 75-79.

Tong, F. Y. (1985). xiaoxue nannvsheng yuwen, shuxue chengji chayi de diaocha[The survey on the difference of the primary school Chinese and mathematics performance]. *jiaoyu lilun he shijian,5*, 32-36..

Wang, C. Y. (2003). shuxue xuexi taidu, xuexi celue dui gaozhongsheng shuxue xuexi chengji de yingxiang yanjiu [Impact of math learning attitude, learning strategies on mathematics learning performance for the secondary school students]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Wang, G. Y. (2006). gaozhixuesheng celue ,ziwo tiaojie xuexi he shuxue chengjiu dongji zhijian de guanxiyanjiu [Relationship between learning strategies, self-regulation learning and learning achievement motivation for the vocational college students]. *Journal of zhejiang transport and vocation college, 7(*2), 51-55.

Wang, G. Y.(2006). gaozhisheng xuexi celue ,ziwo tiaojie xuexi he shuxue chengjiu dongji zhijian de guanxi yanjiu [A Survey on Between Relational about the Study Strategies, the Self- adjustment Study and Mathematics Incentives to Achievement of Higher Vocational Students], *Journal of Zhejiang jiaotong zhiye jishu xueyuan xuebao, 7*(2), 51-55.

Wang, J. Y. (2006). gaoyi xuesheng shuxue chengjiu mubiao dingxiang yu shuxue xuexi jiaolv, zongti manyigan yiji shuxuechengji de guanxi yanjiu [The Study of Relationship with Mathematics Achievement Goal Orientation,Mathematics Learning Anxiety, General Well-being and Mathematics Achievement of Student in Grade One of High School]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Wang, M. L. (2005). shuxue renzhi jiegou bianliang dui chuzhong xuesheng shuxue chengji yingxiang de shiyan yanjiu [Experiment study on the math cognitive structure on the mathematics performance]. (Unpublished Master's thesis). Nanjing Normal University, Nanjing, P.R.China.

Wang, N. Q.(2009). yuanrenzhi jiankong yu shuxue xuexi chengji de xiangguanxing yanjiu [Relationship between metacongition control and mathematics learning performance]. *gaodeng hanshou xuebao: ziran kexue ban, 22,* 57-58.

Wang, P. (2010). zhongxuesheng shuxue xingbie keban yingxiang ji xingbie rentong shuiping yu shuxue chengji de guanxi [The Development of Middle School Students' Mathematical Gender Stereotype, Identity of Sex Role and Their Impact on Mathematic Achievements]. (Unpublished Master's thesis). East China Normal University, Shanghai, P.R. China.

Wei, X.(2008). Chuzhongsheng shuxue ketang mubiao jiegou, dongji xinnian yu xueye chengji de guanxi [The relationship between mathematics classroom teaching goal, motivation belief and learning performance]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Wu, J. (2006). gaoshi shuxue zhuanye xuesheng qinggang yinsu yu chengji de xiangguanxing yanjiu [Relationship between the emotional variables and the mathematics performance for the college students on teacher education]. Journal of Qujing Teacher education college, 3, 86-89.

Wu, X. H., Hou, X. P., Zhu, S. S. (2007). yingxiang xuesheng shuxue chengji de renge yinsu [Personality impact on the mathematics performance].*shuxue jiaoyu xuebao, 16*,(2), 68-70.

Wu, Zhonghe. (2008). guanyu "sujiaoban" xiaoxue shuxue shiyan jiaokeshu xuesheng xueye chengjiu de yanjiubaogao [Research report for the effect of “Su zhou version” experimental textbok on the mathematics performance for primary school].*Jiangsu jiaoyu yanjiu, 8*, 35-40.

Xiao, Z. W.(2010). "zizhutanjiu"dui xuesheng shuxue xueye chengji de yingxiang de shiyan yanjiu [Experiment study on the "self -export"on the mathematics performance]. *shuxue jiaoxue tongxun, 33*, 7-10.

Xie, M., Xin, T., & Li, D. W.(2008). jiaoshi zige he zhiye fazhan yinsu dui xuesheng shuxue chengji de yingxiang: yige kuawenhua bijiao [Effects of teachers' qualification and professional development factors on their students' mathematical achievements: a cross-culture comparison]. *Xinli yu xingwei yanjiu, 6*(2), 124-129.

Xie, Q., & Chen, Y.L.(2006). chuzhongsheng de kongjian nengli yu shuxue chengji ,xingbie de guanxi [Relationship between the space ability, gender and the mathematics performance]. *shanghai zhongxue shuxue, 9,* 12-13.

Xue, H. P. (2010). jiaoyu fenquan guanli zhidu dui nongcun zhongxiaoxuesheng shuxue chengji yingxiang shizheng yanjiu [The empirical research on the impact of decentralized education management system on the student's math achievement in rural compulsory education in China]. *jiaoyu kexue, 26*(4), 26-36.

Xue, S. B. (2005). chu'er xuesheng shuxue ziwo xiaoneng ,ziwo gainian yu shuxue chengji guanxi de yanjiu [Relationship of mathematics self-efficacy, self-concept and mathematics performance for grade 8 students]. (Unpublished Master’s thesis). Northeast Normal University, Changchun , P.R.China.

Yan, P. L. (2009). keban yinxiangweixie yu nvxing shuxue chengji de yanjiu [A Research of Stereotype Threat and Women s Math Performance]. (Unpublished Master’s thesis). North Western University, Xi’an, P.R. China.

Yan, X. N. (2002). Yingxianggaozhongnvsheng shuxue chengji de xinli yinsu su ji duice tantao[Psychological variables which impact on the mathematics performance for female secondary school students and related strategies]. (Unpublished Master’s thesis). Fujian Normal University, Fuzhou, P.R.China.

Yang, Q. (2007). zhongxuesheng chengjiu mubiao ,ziwo tiaojie xuexi celue yu shuxue chengji de guanxi yanjiu [Achievement goal, self-regulation learning strategies and learning performance for the middle school students]. (Unpublished Master's thesis). Shanghai Normal University, Shanghai, P.R.China.

Yang, Y.(2009). youshisheng xuexi zoiwo xiaonenggan, xuexi shiyingxing yu shuxue xueye chengji de guanxi yanjiu Relationship on the self-efficacy, learning adaption and mathematics performance for the preservice teachers who major in preschool[(Unpublished master's thesis). Capital Normal University, Beijing, P.R.China.

Yin, G. E., & Feng, H., (2006). suanshu yingyongti de fenlei jieguo yu shuxue chengji guanxi yanjiu [A Research on the Relationship of Arithmetic Word Problem Categorization with Mathematics Score]. *xinli kexue, 29*(4), 778-781.

Yu, N., & Feng, C. H. (2009). wenlike shuxue chengji chayi de tongji fenxi [Statistical analysis of the difference in mathematics scores between arts and science]. *Journal of Jinggangshan college, 30* (2), 38-49.

Yu, W. H., & Yu, P. (2011). geti ziwo jiankong nengli, siwei pinzhi yu shuxue xueychengji de guanxi yanjiu [Relationship between self-control, thinking style and mathematics performance for individual students]. *xinli kexue, 12*, 141-144.

Yue, R. F.(2006). yingxiang gaozhongsheng shuxue xueye chengji de zhu yinsu fenxi [The main factors which affect the mathematics learning performance of the secondary school students]. *Shuxue jiaoyu xuebao, 16,* (2), 111-116.

Zeng, L. B.(2009). chiyixuesheng xueye zizun,chengjiu mubiao he chengji de xiangguanxing yanjiu [Relationship between mathematics self-esteem, achievement goal and learning performance for grade 7 students]. *shidai jiaoyu (jiaoyu jiaoxue ban), 1,* 157-159.

Zeng, Y., Duan, C. S., & Xiang, L. (2011). yingxiang minhang daxuesheng shuxue xuexi chengji de yinsu [The factors which affect the mathematics learning performance for the university students of the airport]. *Journal of the Neijiang teacher college, 26,(*4), 79-81.

Zhang, D. M..(2005). Gaozhong xuesheng ziwo tiaojie yu shuxue siwei pinzhi ji shuxue xueye chengji de guanxi yanjiu [Relationship between the self-regulation learning, thinking style and mathematics performance for high school studnets]. (Unpublished Master's thesis). Capital Normal university, Beijing , P.R.China.

Zhang, L. M., Li, H. Y., & Xu, L. S.(2007). xuexi celue zai siwei fengge yu shuxue xueye chengjiu guanxi zhong de zhongjie xiaoying [Research on the learning strategies as the mediator of the thinking style and mathematics performance].*xinli yu xingwei yanjiu, 6*, (2), 117-123.

Zhang, M. H., & Yao, Y. J. (2005). "xiaoxue shuxue xuexi celue"jiaoxue dui tigao minzu diqu xiaoxuesheng shuxue chengji de shiyan yanjiu [Experimental Study Of Mathematics Learning Strategies
On Improving Primary School Students' Study Achievement in Minority Regions]. *Journal of Southwest Normal University, 31*(4), 39-44.

Zhang, N. (2006). gaozhongsheng shuxue jiaolv, shuxue ziwo xiaonenggan yu shuxue chengji de guanxi yanjiu [The Study on the Relations of Mathematics Anxiety and Mathematics Achievement of Senior High School Students] (Unpublished Master's thesis). Hebei Normal University, Shijiazhuang, P.R.China.

Zhang, S. X.. (1988). chuzhongxuesheng shuxue he lihua chengji zhijian xiangguan guanxi de yanzheng [Explore the relationship between the mathematics, physics and Chemistry for junior school students]. *Beijing jiaoyu, 11*, 23.

Zhang, S.(2006). gaoyi xuesheng shuxue xuexi xingqu,chengjiu mubiao dingxiang, xuexi celue yu xueye chengji de guanxi yanjiu [Relationship between math learning interests,, achievement goal oriented, learning strategies and learning performance for grade 10 students]. (Unpublished Master's thesis). Northeast Normal University, Changchun , P.R.China.

Zhang, W. J., Xin, T., & Kang, C. H. (2010). jiaoshi bianliang dui xiaoxue sinianji shuxue chengji de yingxiang : yige zengzhixing yanjiu [The Impact of Teacher Variables on the Fourth-Grade Math Achievement: A Value-Added Study]. *Jiaoyuxuebao, 6*(2), 69-76.

Zhang, W. Y., & Yu, T.(2007). daxuesheng shuxue chengji yingxiang yinsu de shizheng fenxi [Empirical study on the variables which contribute to the mathematics performance for university students]. *tongji yu xinxi luntan,22*,(4), 93-96.

Zhang, W. Y., Fan, W. G., & Zhang, S. B.(2008). zhongxuesheng shuxue xuexi xuanze nengli yu xuexi chengji xiangguanxing yanjiu [Relationship between the ability of learning to choose for mathematics and mathematics performance]. *shuxue jiaoyu xuebao, 17*,(1), 59-61.

Zhang, X. G., & Liu, X. Y. (2011). Duli xueyuan shuxue chengji yu daxue yingyu xuexi xiangguanxing diaocha [Relationship between mathematics performance and English learning performance in the university]. *Zhongguo jingji xuejia, 1,* 152-153.

Zhang, X. L. (2004). zhongxuesheng shuxue jiaolv,xinnian yu chengji de guanxi [The Research on Relation Among Mathematics Anxiety, Mathematical Beliefs and Mathematics Achievement]. (Unpublished Master's thesis). Hebei University, Baoding, P.R.China.

Zhang, X. L., Zhang, R. S., & Chen Y. H.(2007). zhongxuesheng shuxue jiaolv, xinnian yu chengji de guanxi [Relation of Mathematics Anxiety,Mathematical Beliefs and Mathematics Achievement]. *Chinese Journal of Clinical Psychology, 15* (3), 287-289.

Zhang, Y. J., & Sun, H. B. (2010). yingxiang xuesheng shuxue chengji de shuxue moxing fenxi [Model analysis on the variables which affect the mathematics performance]. *Kejiao daokan, 10*, 152-153.

Zhao, D. (2006). xiaoxuesheng shuxue wenti jijue zhong chengjiu dongji ,jieti celue he jieti chengji guanxi de yanjiu [Research on the Relationship among Achievement Motivation, Strategies and Results in Pupils' Maths Problems Solving Process] (Unpublished Master's thesis). Northeast Normal University, Changchun, P.R. China.

Zhou, D.,& Wang, X. C.(2007). xiaoxuesheng shuxue guan, shuxue xuexi celue yu xueye chengji de guanxi yanjiu [Relationship between learning beliefs, learning strategies and learning performance on mathematics for primary school students]. *Journal of Inner Mongolia Normal University(Education Science Edition). 20*, (6), 109-112.

Zhou, L. (2008). chuzhongsheng shuxue jiaolv, shuxue xueye ziwo gainian jiqi dui shuxue chengji de yingxiang [Math anxiety, math self-concept and their impact on the mathematics performance for the high school students]. (Unpublished Master's thesis). Henan Univeristy, Kaifeng, P.R.China.

Zhou, L.(2007). Chuyi shuxue xueye fankui, xueye ziwo zushe ji xueye chengji jian de guanxi [The Relationship between Academic Feedbacks, Academic Self-handicapping and Academic Achievement on Math of Grade Seven](Unpublished Master's thesis). Northeast Normal University, Changchun, P.R. China.

Zhou, L., & Liu,X. M.(2007). xueye fankui,xueyeziwoshezu yu chuyi xuesheng chengji de guanxi [Relationship between learning feedback, self-handicapping and mathematics performance for grade 7 students]. *zhongguo zuzhi gongcheng yanjiu yu linchuang kangfu, 2*, 3255-3259.

Zhou, Y. S.(2001). chuzohngsheng shuxuechengji guiyin yanjiu [Research on the attribution of the middle students for the mathematics performance]. *shuxue jiaoyu xuebao, 2*, 38-40.

Zhou, Y., & Tang, J. F. (2007). gaoshi da'er xuesheng shuxue jieti de yuanrenzhi dui jieti chengji de yingxiang [The impact of the metacognition on the problem-soving on the performance for Junior students in the normal university]. *shuxue jiaoyu xuebao, 2,* 60-63.

Zhou, Z., Lian, S. Q., & Zhou, C. L. (2005). zhongxuesheng kongjian tuxing renzhi nengli fazhan yu shuxxue chengji guanxi jiqi yu zhili de xiangguangxing yanjiu [Relationship on the space cognitive ability and the mathematics performance and its relationship with IQ for middle school students]. *Shuxue jiaoyu xuebao, 14* (1), 57-58

Summary

Mathematics learning performance and Mathematics learning difficulties in China

Introduction.
Mathematics is a key literacy in modern society (Engle, Grantham-McGregor, Black, Walker, & Wachs, 2007) and it is a key literacy being developed in the curriculum of the primary school in China (MOE, 2010). Many variables affect mathematics learning and related performance. A thorough understanding of these variables is important for the educators and policy makers. In addition, cross-national studies reveal that mathematics education is embedded in a particular culture (Tang, Zhang, Chen, Feng, Ji, Shen, Reiman,& Liu, 2006). Although a large number of research is available exploring the relationship between factors predicting mathematics performance (Geary, 2006), the former implies that the exploration of this particular relationship should be done in the context of a particular culture; in the present case in the Chinese cultural context. The present dissertation studies in detail the complex interplay of key factors at the level of students, the family and the school that are assumed to contribute to mathematics learning and performance in Chinese primary schools.

The theoretical background.
From the large set of models, already available to explain mathematics performance, three important heuristic models have been selected to direct our research: (a) Walberg's educational productivity model that stresses the need to consider the combined impact of student, family and school related factors; (b) Creemer's educational effectiveness model that stresses the need to re-structure predictor variables for performance into a hierarchical model; (c) Geary's evolutionary theory that implies that individual mechanisms make efforts to solve problems in their personal situation; thus combining an evolutionary perspective with an individual development perspective. These three models gave a first foundation for a framework to study mathematics learning and performance. In the introductory chapter, this first framework is also used in view of a quantitative content analysis of mathematics research literature. This approach was expected to further direct the identification and selection of predictor variables for the studies set up in the context of our dissertation. From international research databases (e.g. Web of Science), 1041 articles reporting on international or national studies were selected. After application of further selection criteria, 230 articles were analyzed and coded in more detail. The results show how –

since the year 2000 - more and more studies focus on mathematics performance and predictor variables. But, the results also show that to a lesser extent studies are set up at the primary school level in national (Chinese) articles as compared to international articles. This is partly explained by the differences in the aims of national or international studies. In the PhD studies, we therefore especially build on variables neglected in the international studies that are considered to be important predictors in the Chinese context.

General research problem and research aims.

Based on the previous studies, the general research problem is related to the identification of important predictors for mathematics performance in Chinese primary schools. The research focuses on three main aims: (1) the design, development and implementation of a standardized assessment for mathematics in Chinese elementary school children; (2) to test models that link predictors to mathematics learning performance in the normal students; (3) to examine models that link predictors to the performance of students at risk for mathematics difficulties.

The study.

The dissertation builds on eight studies, set up in line with the three research aims.

Standardized measurement instrument for mathematics performance.

The first chapter focuses on the design, development and evaluation of a standardized instrument to determine mathematics performance in all grades of Chinese primary schools. The instrument covers the curriculum dimensions in the Chinese primary school curriculum and covers twelve building blocks in relation to mathematics. Item Response Theory was applied in developing and testing the instrument. The evaluation of the final instrument – developed after a piloting phase - shows that (1) the item difficulty range covers a broad range in student ability; (2) the item information curves and test information curves reveal that the items cover a satisfactory range of abilities and provide reliable information about student abilities. Compared to existing instruments in China, this new instrument provides students with a tool to map the mastery and the development of primary school learners. The latter is guaranteed by the inclusion of a sufficient number of overlapping items in the subsequent grade level instruments.

Modeling the relationship between predictors and mathematics performance.

Chapter 3-6 report on studies that link variables at the level of the student, the class and the school to mathematics performance.

The findings in chapter 3 illustrate the multilevel structure in the school level variables, class level variables and student level variables when studying mathematics performance. Most of the variance can be attributed to variables at the student level. The interaction between school level variables, class level variables and student level variables suggests there are indirect effects from the class and school level on mathematics performance.

This is explored in more detail in Chapter 4 where the focus is on the contextual variables and how they interact with student level variables. The results in Chapter 4 provide a clear picture of the direct and indirect effects on mathematics performance. Teacher beliefs seem to affect student self-efficacy; mother's educational level influences students' metacognitive experiences. And mathematics self-efficacy and metacognitive experiences work as moderators for contextual variables.

In Chapter 5 we center on surprising findings from previous studies (Sirin, 2005) about the relationship between SES and mathematics performance. The results of chapter 5 indicate that – instead of a linear relationship – we observe a U-shaped relationship between SES and mathematics performance in China. This implies that students from lower SES families can also attain higher achievement levels and that school aggregated SES moderates the effect of family SES on performance.

Since we know from Chapter 3 that very few variables at the teacher level seem to contribute to mathematics performance, the study in Chapter 6 is nevertheless an attempt to study actual teaching behavior to be linked to mathematics performance. The study is based on the analysis of videotaped lessons. The results show that - in the Chinese context - teachers dominate the classroom interaction but that this particular interaction between one teacher and one student positively contributes to mathematics performance. Another interesting finding is that the nature of the questions teacher ask is also important. Especially questions that center on evaluation and on problem-solving strategies contribute to mathematics performance.

Another question roused by the findings in Chapter 5, is tackled in Chapter 7: is is possible to foster the achievement of students from lower SES families? In Chapter 5, we explore the effects of homework assignment by teachers, parents and students themselves. The research results are complex. First; it seems that parents and students from disadvantaged families make use of homework as a way to compensate for their underprivileged background. Students develop homework assignments to the extent their parents assign homework to them. Lastly, students from disadvantaged families benefit from homework assignments only when students from advantaged families do not get extra homework.

Models linking predictors to the performance of students at risk for mathematics difficulties

Chapter 7 and 8 present studies that examine a variety of variables that can be linked to the performance of children at risk of mathematics difficulties. Chapter 7 builds on the model studied in Chapter 3 and the results point at the impact of lower self-efficacy, lower level of metacognitive experiences, low family SES, and restricted teacher's quality on mathematics performance. But – and this is an additional finding - the development level of the region (GDP), also increases the probability of being at risk for learning difficulties. The results point at the large proportion of students at risk enrolled in schools located in low developmental level regions.

Chapter 8 rather focuses on student variables to study mathematics difficulties in learners at risk. We focus in this study on the mastery of a particular mathematics component: numerical facility skills. In this chapter, a cross-cultural approach is adopted by exploring and comparing the mastery of numerical facility skills in Flemish and Chinese students. The results show that up to the end of the primary school, Chinese learners outperform Flemish learners. But the difference decreases with increasing grade level for the four numerical facility skills except multiplication. In both countries high achievers perform good and stable on all numerical facility tasks, whereas low achievers do not attain a stable achievement level. This suggests the inference of cultural differences, such as critical differences in the mathematics language used in each cultural setting.

Implications.

At the theoretical level, the results of the eight studies reported in this dissertation provide some culture specific results about the nature and impact of mathematics education in China. For example, the individual characteristics play a moderating role between the impact of contextual variables (family, school) on mathematics performance. The interaction effect between individual level variables and school level variables is significant for the mathematics performance. This implies that the particular mathematics environment does not "work" for all students. From a theoretical perspective, an important findings is the U-shaped relationship between SES and mathematics performance in Chinese primary school; again to be attributed to culture differences. At the practical level, the results in this dissertation present implications for mathematics education. For educational administrators, extra attention should be paid to children with a disadvantaged background and especially to these children enrolled in schools in lower economical level regions. Though the results show that students from disadvantaged family can attain a higher performance, educational policies should be enhanced to support these students in their school setting. In relation to teacher training, our findings indicate that teachers should be made aware of what factors influence performance. In particular, teachers should be helped to develop didactical approaches that are aligned with the needs of children at risk and children with a disadvantaged background. The development of particular interaction styles and questioning approaches could be reinforced in the training of their didactical skills.

Limitations.

There are specific limitations that should be overcome in future studies. First, there is a limitation in the structure of the newly designed mathematics instrument. As explained above, the different grade level test partly consist of overlapping items with the former or subsequent grade level test. But the overlapping items do yet not cover all the 12 building blocks and the three curriculum domains. We can improve the item bank by adding new items in a future study. Second, in view of developing a model to explain mathematics performance, specific variables have been (see the content analysis). But some variables have been ignored and yet not included in the model; mainly because of time limitations. Additional variables at the school level and the individual level and their interaction should be focused upon in future studies. Thirdly, we only set up cross-sectional studies of mathematics performance. A longitudinal study, involving the same cohort of students that develops from grade 1 tot grade 6, could help to track in a detailed way particular performance changes over time and how over time particular variables at the student, class and school level do or do not play a particular influence.

Conclusions.

In summary, this dissertation explored the mathematics performance and related predictors in the Chinese primary school context. Firstly, a new and standardized instrument was developed, based on item response theory and embracing twelve mathematics building blocks and three mathematics curriculum domains. This instrument was found to be reliable to measure mathematics performance and helpful to differentiate students with different mathematics abilities. Second, on the base of model testing, we found that variables at the individual student level explain the largest proportion in the variance of mathematics performance. These variables work as a moderator between contextual variables and performance. Thirdly, in children at risk of mathematics learning difficulties, our cross-cultural study reveal that – in contrast to low achievers - high achievers perform very well and in a stable way on all numerical facility tasks.

Reference

Engle, P., Grantham-McGregor, S., Black, M., Walker, S., & Wachs, T. (2007). How to avoid the loss of potential in over 200 million young children in the developing world. *Child health and education, 1* (2), 67-87.

Geary, D. C. (2006). Development of mathematical understanding. In D. Kuhl & R. S. Siegler (Vol. Eds.), *Cognition, perception, and language, Vol 2* (pp. 777–810). W. Damon (Gen. Ed.), *Handbook of child psychology (6 th Ed.)*. New York: John Wiley & Sons.

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

Tang, Y., Zhang, W., Chen, K., Feng, S., Ji, Shen, Reiman E. M. & Liu, Y. (2006). Arithmetic processing in the brain shaped by cultures. *PNAS, 103*, 10775-10780.

Samenvatting

Wiskundeleerprestaties en wiskunde leermoeilijkheden in China

Inleiding.

Wiskunde is van groot belang in onze moderne maatschappij (Engle, Grantham-McGregor, Black, Walker, & Wachs, 2007). Gecijferdheid moet dan ook ontwikkeld worden bij lagere schoolkinderen om ze voor te bereiden op deze maatschappij. Dit is uiteraard ook van belang in China (MOE, 2010). Er zijn tal van variabelen die bepalen hoe vaardig kinderen worden op wiskundige taken. Voor leerkrachten is het relevant om na te gaan welke die variabelen zijn en hoe belangrijk ze zijn. Vergelijkende studies leren ons echter ook dat reken- en wiskunde onderwijs niet kan los gezien worden van de cultuur waarin deze plaats vindt (Tang, Zhang, Chen, Feng, Ji, Shen, Reiman,& Liu, 2006). Alhoewel er al tal van studies zijn die de relatie tussen deze variabelen onderzochten (o.a. Geary, 2006), ontbreekt onderzoek in de Chinese setting. In dit proefschrift willen we deze leemte vullen en nagaan welke kind-, familie- en schoolgerelateerde variabelen bijdragen tot de wiskundebeheersing van Chinese lagere schoolkinderen.

Theoretische achtergrond.

Uit de veelheid aan modellen die wiskundebeheersing verklaren, werden drie belangrijke heuristische modellen geselecteerd. (a) Het model van Walberg m.b.t. onderwijskundige productiviteit dat het belang van factoren benadrukt in de leerling, de familie en de school. (b) Het model van Creemers over onderwijskundige effectiviteit dat stelt dat we beïnvloedende variabelen moeten herordenen in een hiërarchisch model. (c) Het model van Geary dat een evolutionaire theorie is en individuele ontwikkeling van leerlingen bij het oplossen van problemen relateert aan evolutionaire ontwikkelingsmechanismes. Deze drie modellen gaven richting aan een kwantitatieve inhoudsanalyse van onderzoeksartikels m.b.t. wiskundebeheersing. Het doel was een verdere definiëring en explicitering van factoren die wiskundebeheersing beïnvloeden. Uit een aanvankelijke verzameling van 1041 internationale en nationale onderzoeksartikels werden 230 artikels geselecteerd die vervolgens geanalyseerd en gecodeerd werden. Hieruit bleek dat vanaf het jaar 2000, meer en meer en studies uitgevoerd werden rond wiskunde en al wat daarmee verband houdt. Toch bleek ook dat er heel wat minder onderzoek werd opgezet bij lagere schoolkinderen in China in vergelijking met international studies uit diezelfde periode. Aangezien de doelstellingen van nationale en international studies verschillen, werd voor dit proefschrift vooral die variabelen geselecteerd die minder aandacht kregen in de bestaande studies. Deze werden als predictoren naar voren geschoven voor onderzoek in de Chinese context om beter te begrijpen wat het leren van wiskunde in China richting geeft.

Algemene doelstellingen en onderzoeksvragen.

Het algemeen doel van dit proefschrift is het uittesten van een model dat predictoren expliciteert die wiskundebeheersing in het Chinese lagere onderwijs beschrijft en verklaart. Daarbij staan drie onderzoeksdoelen centraal: (1) het ontwikkelen en valideren van een gestandaardiseerd wiskundemeetinstrument voor het lagere onderwijs in China; (2) het testen van een model van predictoren die de wiskundebeheersing van van normale leerlingen in de lagere school beschrijven en verklaren; (3) het testen van dit predictorenmodel bij leerlingen die een risico lopen voor het ontwikkelen van wiskundeleermoeilijkheden.

Studies.

Het proefschrift bestaat uit acht studies die aansluiten op de drie vermelde onderzoeksdoelen.

Een gestandaardiseerd instrument om wiskundevaardigheden te onderzoeken.

Het eerste hoofdstuk bespreekt de ontwikkeling en validatie van een meetinstrument om de wiskundebeheersing van Chinese lagere schoolkinderen te bepalen. Het instrument bouwt verder op drie curriculumdomeinen in het Chinese wiskundecurriculum wiskunde en op twaalf cognitieve bouwstenen die in de literatuur onderscheiden worden m.b.t. rekenen-wiskunde. Bij de ontwikkeling van de items/test is Item Respons Theorie toegepast.

De onderzoeksresultaten tonen aan dat (1) de moeilijkheidsgraad van de items matchen met de vaardigheden van de leerlingen; (2) de item- en test informatiecurves tonen aan dat een brede variatie in vaardigheidsniveau kan gemeten worden en dit op een psychometrisch onderbouwde manier. In vergelijking met bestaande tests in China, laat dit nieuwe instrument leerkrachten toe dat leerkrachten om de actuele beheersing en de ontwikkeling van wiskunde te bepalen. Dit laatste is mogelijk doordat er in de opeenvolgende testen op klasniveau, telkens een deelverzameling van items overlapt.

Het testen van een model van predictoren die de wiskundebeheersing van normale leerlingen in de lagere school beschrijven en verklaren

Hoofdstuk 3-6 test een model dat predictoren op het niveau van de leerling, leerkracht en school betrekt op wiskundebeheersing in de lagere school in China.

De resultaten van het onderzoek beschreven in hoofdstuk 3 tonen het belang aan van variabelen op schoolniveau, klasniveau en leerlingniveau. Maar de grootste proportie van de variantie in wiskundebeheersing wordt verklaard door factoren op het leerlingniveau. De interactie tussen de variabele op het school, klas en leerlingniveau laat verder vermoeden dat er een indirect effect bestaat van de andere niveaus op wiskundebeheersing. Dit brengt ons naar hoofdstuk 4 waar vooral de rol van contextuele variabelen wordt onderzocht.

De resultaten gerapporteerd in hoofdstuk 4 geven ons een beeld van de directe en indirecte effecten van de variabelen die wiskundebeheersing beïnvloeden. De opvattingen (beliefs) van de leerkracht heeft een invloed op de self-efficacy van leerlingen. Verder heeft het opleidingsniveau van de moeder een directe invloed op het niveau van de metacognitieve ervaringen (metacognitive experiences) van leerlingen. Tenslotte blijken self-efficacy en metacognitieve ervaringen moderatoren te zijn tussen contextuele variabelen en wiskundebeheersing.

In hoofdstuk 5 wordt ingegaan op de speciale invloed van sociaal economische status (SES) op en wiskundebeheersing. Er wordt – in tegenstelling tot het meeste onderzoek waar een lineair verband wordt vastgesteld - een U-vormige relatie tussen SES en wiskundebeheersing vastgesteld in de Chinese context. Leerlingen uit een gezin met een lage SES doen het beter dan kinderen uit een gezin met een gemiddeld SES. De SES van de school (geaggregeerd) modereert het effect van SES van het gezin op wiskundebeheersing.

In hoofdstuk 3 werd aangegeven dat er slechts een beperkt aantal leerkrachtvariabelen een directe invloed hebben op wiskundebeheersing. Om dit nader te onderzoeken werden in het onderzoek gerapporteerd in hoofdstuk 6 video-opnames van lessen in Chinese lagere school klassen geanalyseerd. De resultaten in hoofdstuk 6 tonen dat in China de leerkracht de klasinteractie domineert. De directe interactie leerkracht-leerling is positief gerelateerd aan betere wiskundeprestaties. Bovendien blijken Chinese leerkrachten die veel evaluatieve vragen stellen en problem solving vragen, ook een hoger niveau in wiskundebeheersing te bereiken bij hun leerlingen.

Een vraag, voortbouwend op het onderzoek uit hoofdstuk 5 en bestudeerd in hoofdstuk 7 zoekt naar de manier waarop en mate waarin leerlingen van gezinnen met een lage SES er toch in slagen om goed te scoren op wiskunde. In dit onderzoek staat de impact van huiswerk centraal. In China stellen we vast dat drie actoren huiswerk geven: de leerkracht, de ouders en de leerlingen zelf. De onderzoeksresultaten geven een complex beeld. Eerst en vooral blijkt dat ouders en leerlingen uit gezinnen met een lage SES huiswerk gebruiken om te compenseren voor hun zwakkere familiecontext. Ten tweede blijken leerlingen zichzelf huiswerk te geven in de mate dat ouders hen taken geven. Leerlingen uit kansarme gezinnen (laag SES) blijken baat te hebben van huiswerk; maar enkel wanneer we dit vergelijken met leerlingen uit gezinnen met een hoge SES die geen extra huiswerk krijgen.

Het testen van een model van predictoren bij leerlingen die een risico lopen voor het ontwikkelen van wiskundemoeilijkheden

In hoofdstuk 7-8 gaan we in op factoren die een rol spelen bij een risicogroep voor wiskundeleermoeilijkheden in China. In hoofdstuk 7 wordt het predictorenmodel, reed onderzocht in hoofdstuk 3, getest bij risicokinderen. Dezelfde variabelen blijken naar voren te komen: een lage self-efficacy, gebrekkige metacognitieve ervaringen, een lage SES, en een lager diplomaniveau van de leerkracht zijn significante en negatief beïnvloedende indicatoren. Bovendien blijken leerlingen die school lopen in een regio met een lager economisch ontwikkelingspeil (lager GDP), ook een grotere kasn te lopen op wiskundemoeilijkheden. Dit is een belangrijke bevinding omdat juist heel veel leerlingen - met rekenproblemen – school lopen in economisch zwakker ontwikkelde regio's in China.

Hoofdstuk 8 gaat in op onderliggend e individuele processen bij wiskundebeheersing. Bij een aantal kinderen lukt rekenen wel, maar het gaat minder snel en minder geautomatiseerd. Het onderzoek in hoofdstuk 8 onderzoekt daarom de beheersing van het snel kunnen oproepen van rekenfeiten (numerical facility) en wiskundebeheersing. Daarbij werd een vergelijking gemaakt tussen de beheersing van Chinese en Vlaamse kinderen. De resultaten tonen aan dat Chinese kinderen het in alle leerjaren beter doen dan hun Vlaamse leeftijdsgenoten. Dit significant verschil tussen Chinese en Vlaamse kinderen neemt echter af met de leeftijd voor optellen, aftrekken en delen. Dit is niet het geval voor vermenigvuldigen, wat kan verklaard worden door culturele variabelen (bijv. de verschillen in de taal). Een tweede conclusie is dat - zowel in Vlaanderen als in China - kinderen die hoog scoren voor wiskunde het goed doen op de beheersing van alle rekenfeiten (dus zowel voor optellen, aftrekken, vermenigvuldigen als delen). Kinderen die niet vlot rekenen en dus rekenproblemen hebben, scoren veel ongelijkmatiger op de vier subdomeinen.

Implicaties.

Op het theoretisch niveau, ondersteunen de resultaten van de verschillende studies het belang van cultuurgebonden verschillen in factoren die de wiskundebeheersing bepalen. De dominante rol van factoren op het leerlingniveau wordt herbevestigd. Leerlinggebonden factoren blijken ook een modererende rol spelen tussen context variabelen (familie school) en wiskundebeheersing in China. Het interactie effect tussen individuele variabelen en schoolvariabelen toont ook aan dat dezelfde schoolomgeving niet voor alle leerlingen even geschikt is. Theoretisch van belang is ook de vaststelling dat in deze Chinese context er een U-vormige relatie bestaat tussen SES en wiskundebeheersing.

De resultaten van de onderzoeken gerapporteerd in dit proefschrift hebben ook een aantal beleids- en praktijkimplicaties voor het wiskundeonderwijs. De studies tonen bijvoorbeeld aan dat leerlingen uit kansarme gezinnen het toch nog behoorlijk kunnen doen voor wiskunde. Maar het blijft noodzakelijk om deze doelgroep voldoende en zelfs extra ondersteuning te geven. Vooral de vaststelling dat er in scholen in economisch minder ontwikkelde regio's meer kinderen met wiskundemoeilijkheden voorkomen is voor het beleid van belang. Wat betreft de lerarenopleiding bevestigen de onderzoeksresultaten het belang van de kwaliteit van "vragen stellen" in de klas en het aanpakken van een aangepaste interactiestijl met de leerlingen.

Beperkingen.

Aan elk onderzoek zijn er beperkingen. Voor een overzicht van specifieke beperkingen verwijzen we naar de discussie in de verschillende hoofdstukken. Een algemene beperking ligt ten eerste in de structuur va het nieuwe meetinstrument. Het aantal overlappende items in de verschillende testen per leerjaar, laat voorlopig niet toe om gedifferentieerde uitspraken te doen over alle twaalf cognitieve deeltaken of over de 3 curriculumdomeinen in China. In vervolgonderzoek kan deze beperking weggewerkt worden door nieuwe items toe te voegen aan de bestaande itembank. Een tweede beperking is dat wiskundevaardigheden beïnvloed worden door heel wat factoren school-, klas- en individueel niveau, maar dat we deze niet allemaal hebben kunnen meenemen in dit proefschrift. Vervolgonderzoek is dus aangewezen. Ten laatste is alle onderzoek in deze PhD opgezet op basis van een grootschalig maar cross-sectioneel onderzoeksdesign. Verdiepende longitudinale studies zijn aangewezen in vervolgonderzoek om de feitelijke ontwikkeling van bijv. een cohort leerlingen doorheen de jaren van de lagere school te kunnen volgen en daarbij de impact van specifieke predictoren op de wiskundebeheersing te kunnen traceren.

Conclusies.

Samenvattend kunnen we stellen dat dit proefschrift een aanzet heeft gegeven rond de modelvorming m.b.t. significante predictoren voor wiskundebeheersing in de context van het Chinese lagere onderwijs. Daarvoor werd ten eerste een nieuw meetinstrument ontwikkeld (doelstelling 1), gebaseerd op de item response theorie. Items in de opeenvolgende testen peilen naar de beheersing van twaalf wiskunde-bouwstenen en sluiten aan op drie domeinen in het Chinese wiskundecurriculum. Dit nieuwe meetinstrument bleek psychometrisch goed onderbouwd te zijn. Wat betreft het tweede doel van ons onderzoek, konden we vaststellen dat de variabelen op het leerlingniveau de grootste variantie verklaren in wiskundebeheersing. Deze variabelen werken ook als een moderator tussen de invloed van contextvariabelen (school en familie) en wiskundebeheersing . de resultaten tonen aan dat we ook rekening moeten houden met de interactie tussen de variabelen op de verschillende niveaus. Wat betreft het derde onderzoeksdoel, leerde het cross-culturele onderzoek dat – sterke - leerlingen zowel in China als in Vlaanderen een harmonieus profiel hebben van geautomatiseerde rekenvaardigheden en dat ze dus vlot zijn in het uiteenlopende rekenfeiten kunnen oproepen. Dit is niet het geval bij rekenzwakke leerlingen. De resultaten tonen ook aan hoe Chinese leerlingen over de ganse lijn beter presteren dan Vlaamse leerlingen en dit in alle leerjaren van de lagere school.

Referenties

Engle, P., Grantham-McGregor, S., Black, M., Walker, S., & Wachs, T. (2007). How to avoid the loss of potential in over 200 million young children in the developing world. *Child health and education, 1* (2), 67-87.

Geary, D. C. (2006). Development of mathematical understanding. In D. Kuhl & R. S. Siegler (Vol. Eds.), *Cognition, perception, and language, Vol 2* (pp. 777–810). W. Damon (Gen. Ed.), *Handbook of child psychology (6 th Ed.)*. New York: John Wiley & Sons.

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

Tang, Y., Zhang, W., Chen, K., Feng, S., Ji, Shen, Reiman E. M. & Liu, Y. (2006). Arithmetic processing in the brain shaped by cultures. *PNAS, 103*, 10775-10780.

Academic output

SSCI indexed journal (A1)

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

Zhao, N.N., Valcke, M., Desoete, A., & Verhaeghe, J. (in press). The quadratic relationship between socioeconomic status and learning performance in China by multilevel analysis: Implications for policies to foster education equity. *International Journal of Educational Development*.

Zhao, N.N., Valcke, M., Verhaeghe, J., & Desoete, A. (submitted). Modeling a standardized instrument to diagnose mathematics performance in China. *Asia Pacific Journal of Education.*

Zhao, N.N., Valcke, M., Desoete, A. Zhu, C. & Sang, G. Y.(submitted). A holistic model to predict mathematics performance: the interrelated impact of student, family and school context variables. *Scandinavian Journal of Educational Research*

Zhao, N.N., Valcke, M., & Desoete, A., Sang, G.Y. & Zhu, C. (submitted). Teacher's classroom teaching on students' performance in Chinese context: A video analysis. *Asia-Pacific Journal of Teacher Education*

Zhao, N.N., Valcke, M., Desoete, A. & Verhaeghe, J. (submitted) Can homework compensate for a disadvantaged socioeconomic background? *Educational Research*.

Zhao, N.N., Valcke, M., & Desoete, A. (submitted). Detecting the children at risks on academic achievement in Chinese primary school by hierarchical logistic regression model. *Educational Studies*.

Zhao, N.N., Burny, E., Valcke, M., Desoete, A. & Imbo, I. (submitted). The mastery of numerical facility skills in Chinese and Flemish primary school learners. *European Journal of Psychology of Education*

Zhao, N.N., Valcke, M., & Desoete, A. (submitted). Quantitative Content Analysis on the Studies of Mathematics Performance and related Predictors for Chinese Students From 1960s to 2010s. *Asia Pacific Education Review*

Conference contributions

Zhao, N.N., Valcke, M., & Desoete, A. (2009).Examining the factors influencing primary school students' metacognition in mathematics learning: A Chinese perspective. European Association for Research on Learning and Instruction (EARLI), Amsterdan, Netherland, 23-28, September, 2009.

Zhao, N.N., Valcke, M., Desoete, A. & Verhaeghe, J. (2009). Design, development and calibration of an item bank for mathematics assessment in China: applying IRT analysis. European Conference on Educational Research (ECER), Vienna, Austria, 25 - 26 September 2009.

Zhao, N.N., Valcke, M., & Desoete, A. (2010). Which is more important?
A multilevel on Individual and Contextual Influences on mathematics performance in primary school. European Conference on Educational Research (ECER), Helsinki, Finland, 23-27, Augusts, 2010.

Zhao, N. N., Burny, E., Valcke, M., & Desoete, A. (2011). The relationship between numerical facility and mathematics skills by comparing performance of primary schools students in Flanders and China: Implications for mathematics education. American Educational Research Association (AERA), New Orleans, USA, 8-12, April, 2011.

Zhao, N. N., Valcke, M., & Desoete, A. (2011). The quadratic relationship between socioeconomic status and learning performance in China by multilevel analysis. European Association for Research on Learning and Instruction (EARLI), Exeter, British, 29 Aug -3 Sept., 2011.

Zhao, N. N., Valcke, M., & Desoete, A. (2011). Can homework compensate for the disadvantaged family? European Conference on Educational Research (ECER), Berlin, Germany, 8-12, September, 2011.