Influence of curriculum reform: An analysis of student mathematics achievement in Mainland China

Yujing Ni a,*, Qiong Li b, Xiaqing Li a, Zhong-Hua Zhang a

a Department of Educational Psychology, The Chinese University of Hong Kong, HKSAR, China
b Center for the Research on Teacher Development, Beijing Normal University, Beijing, China

ARTICLE INFO

Keywords:
Curriculum reform
Primary mathematics
Curriculum evaluation
Student mathematics achievement
Cognitive
Affective

ABSTRACT

This study investigated curriculum influences on student mathematics achievement by following two groups of students from fifth to sixth grade that were taught either the reformed curriculum or the conventional curriculum. Analyses with three-level modeling were conducted to examine learning outcomes of the students who were assessed three times over a period of 18 months. Achievement was measured with regard to computation, routine problem solving, and complex problem solving. Affective aspects included self-reported interest in learning mathematics, classroom participation, views of the nature of mathematics, and views of learning mathematics. The results showed overall improved performance among all the students over the time on computation, routine problem solving, and complex problem solving but not on the affective measures. There were differentiated patterns of performance between the groups. On the initial assessment, the reform group performed better than the non-reform group on calculation, complex problem solving, and indicated higher interest in learning mathematics. The two groups did not differ on the other achievement and affective measures at the first time of assessment. There was no significant difference in growth rate between the groups on the cognitive and affective measures except that the non-reform group progressed at a faster pace on calculation. Therefore, the non-reform group outperformed the reform group on computation at the third (last) assessment. These results are discussed with respect to the possible influence of the curriculum on student learning.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In 2001, the Ministry of Education of Mainland China1 put forth new mathematics curriculum standards for the 9-year compulsory education. As part of this initiative, the Ministry introduced and approved designated textbooks to facilitate the implementation of the new standards. In the same year, on a voluntary basis, numerous schools in 38 cities (counties) from 27 provinces across the country adopted the new curriculum standards and new textbooks. By the fall of 2006 the implementation became mandatory.

The new standards require changes not only in how mathematics is viewed but also how mathematics is taught and learnt. Clearly, it differs significantly from the previous and conventional mathematics curriculum. Consequently, the education community and the society as a whole in the country have expressed concern about the efficacy of the new

* Corresponding author.
E-mail address: yujing@cuhk.edu.hk (Y. Ni).

1 This study was conducted in People’s Republic of China. In this article, all instances of the word ‘China’ refer to the People's Republic of China, excluding Hong Kong, Macau, and Taiwan.

0883-0355/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijer.2011.06.005
curriculum standards. Therefore, there is a great demand for independent investigation into the effects of the reform curriculum on classroom teaching and learning. The present study investigated the influences of the new curriculum on student learning outcomes in primary mathematics. In the report, we first describe the background of the mathematics curriculum reform in China. We then explain the conceptualization of the research question based on the existing literature on the influences of curriculum reform, mainly in the area of mathematics curriculum, on student learning outcomes. Finally, we present and discuss the findings of the study.

1.1. New mathematics curriculum in China

To understand what the new curriculum standards are intended to accomplish, an explanation is provided below of why Chinese needed this new wave of curriculum reform in mathematics and of how the new curriculum standards differ from the previous ones.

Chinese students have excelled in international assessments of mathematics achievement. However, this outstanding mathematics achievement demonstrated in PISA and TIMSS by Chinese students has been for those from Chinese Taipei and Hong Kong SAR (National Center for Educational Statistics, 2009; OECD, 2009). Mainland Chinese students have not participated in the international assessments with two exceptions: Chinese students from Shanghai took part in the 2009 PISA and those from Beijing participated in the 1989 mathematics assessment conducted by the International Assessment of Educational Progress (IAEP). Nevertheless, Mainland Chinese students have shown certain characteristics in their mathematics achievement. In particular, they have strong computation skills and solve routine mathematics problems well, but have more difficulties with non-routine problems. In the 1989 IAEP, the Chinese students were at the top in overall performance among the students from 21 countries but ranked 9th on the problems requiring non-routine solutions (Fan & Zhu, 2004). Similar findings were obtained by Cai and his colleagues in a series of studies that compared Chinese students and American students (Cai & Cifarelli, 2004; Cai & Hwang, 2002; Cai & Lester, 2005). In these studies, the Chinese students demonstrated high levels of accuracy and efficiency in dealing with word problems in mathematics as they were more likely than American students to use abstract and generalized strategies to solve the problems. The preference for abstract strategies helped the Chinese students to reach correct answers more readily. Chinese students also used more conventional strategies than American students did to solve mathematics problems, resulting in more accurate solutions (Cai, 2000; Wang & Lin, 2005). Chinese students also appeared less willing to take risks to solve mathematics problems (Liu & Sun, 2002). Given a problem that they did not know how to solve, Chinese students were more likely to leave it a blank, but American students often wrote down something anyway (Cai & Cifarelli, 2004).

The characteristics of the Chinese students’ mathematics performance are considered to relate both to the curriculum, which emphasizes knowledge transmission and acquisition and to the classroom instruction, which is highly directive (Fan, Wong, Cai, & Li, 2004; Ni, Chiu, & Chen, 2010). The previous Chinese mathematics curricula have emphasized (1) basic mathematics concepts and (2) basic mathematics calculations. In addition, Chinese classroom instruction has focused on refined lectures and repeated practice (Zhang, Li, & Tang, 2004). That is to say, Chinese instruction has emphasized foundational knowledge of content and skills over creative thinking processes (Leung, 2001). Some Chinese educators argue that repeated practice aids memorization and this greater exposure to practice helps students think about the underlying concepts more deeply (Dhlin & Watkins, 2000). Hence, the Chinese mathematics curricula have had four learning goals for students to achieve: (1) fast, accurate manipulation and computation involving arithmetic, fractions, polynomials, and algebra, (2) accurate recall of memorized mathematics definitions, formulas, rules, and procedures, (3) understanding of logical categorizations and mathematics propositions; and (4) facile matching of solution patterns to types of problems via transfer (Zhang et al., 2004).

To implement these curricula, teachers present well-prepared lessons that include strong teacher control, coherent instruction, and abstract mathematics (Zhang et al., 2004). Chinese teachers value and use more abstraction to generalize mathematics relationships in their instruction compared to American teachers (Cai, 2005; Cai & Lester, 2005; Correa, Perry, Sims, Miller, & Fang, 2008). Driven by the curricula featuring the two basics and classroom instruction with refined lectures and repeated practice, Chinese students spend more time doing homework than either American or Japanese students (Chen & Stevenson, 1989, 1995; Stigler, Lee, & Stevenson, 1990).

The merits of Chinese mathematics teaching and learning, such as the emphasis on the two basics (i.e., concepts and skills), refined lectures, and repeated practice, among others, (e.g., Stevenson & Stigler, 1992; Stigler & Hiebert, 1999), has made Chinese students achieve at a high level of proficiency in mastering these two basics. Nevertheless, the limitations of Chinese mathematics teaching and learning are being recognized and they are the target for the new round of curriculum reform (Chang, 2008; Liu & Sun, 2002). Among the limitations, the previous curricula strongly focused on substantive knowledge, that is, established facts, concepts and skills, but neglected syntactic knowledge, that is, how knowledge is constructed and advanced. This curricular orientation gave rise to classroom instruction that paid less attention to the connection between given knowledge and the processes to establish the knowledge and the relation between the school subject matter and real-life applications. In classroom instruction, the teacher and the textbook was the knowledge authority from which the students received ‘packaged knowledge’ and look up for the confirmation of knowledge from the authority. As a result, many Chinese students did not think learning mathematics has much to do with their present and future life but is useful merely for taking examinations (Liu, 2000). Their interest and confidence in learning mathematics was deteriorating over the years as they moved up to higher grades (Liu & Sun, 2002). In this connection, it is of interest to note that Chinese
Hong Kong students were ranked the first in the PISA assessment of mathematics literacy, but the seventh lowest in mathematics self-concept among the 41 countries/regions (Ho, 2003).

The new mathematics curriculum standards are intended to address the limitations noted above and to make the school mathematics meet the needs of students and the changing society (Ministry of Education, 2001a). Below, the features of the new curriculum standards in comparison to the previous ones are highlighted using Tyler’s (1949) dimensions of curriculum development: objectives, content, instructional methods, and evaluation.

The new mathematics standards embrace three objectives: knowledge and skills, processes and methods, affective demeanor and value. The objectives aim for students (1) to acquire important knowledge and the basic problem-solving skills in mathematics that are important for their life-long learning; (2) to apply knowledge of mathematics and related skills to observe, analyze, and solve problems in daily life and in other subjects by using mathematical methods; and (3) to appreciate the close relationship among mathematics, nature, and society. These objectives are not merely to require students to acquire basic mathematical knowledge and skills, but also to provide them with the opportunities to reason about evidence and explanation, evaluate knowledge claims, use learned knowledge and skills to solve real life problems, and develop interest and confidence in learning and using mathematics. The three objectives are intended to give more attention to providing students the learning experience of how mathematics knowledge is established and advanced by observation, reflection, and communication and how to use mathematical tools to observe, analyze, and solve problems, which was neglected in the previous curricula.

With regard to subject matter content, the new mathematics standards delineate four major areas: (1) numbers and algebra, (2) space and shapes, (3) statistics and probability, and (4) using and applying mathematics. The first area is dominant in both the previous curricula and the new curriculum. However, compared to previous curricula, the new curriculum decreases the difficulty level of numerical and algebraic computations and the requirement to memorize equations and formulas, but focuses instead on estimation and observation of regularity in number to develop number sense, and on mathematical modeling of a given problem. For the second area ‘space and shapes,’ the new curriculum reduces the difficulty level with respect to geometric proof, but puts more emphasis on perceiving and understanding shapes, their transformations, their positions, and related proofs. The strand of statistics and probability, which is a fast-developing branch of applied mathematics, is to equip students with the basic mathematical tools to describe and understand uncertainty in real life. The new curriculum requires that the beginning ideas of probability and statistics will be introduced in the first stage (grades 1–3) of the 9-year compulsory education, whereas the ideas were introduced in the second stage (grades 4–6) in the previous curricula. The part of using and applying mathematics is a new addition and required for all three stages (grades 1–3, 4–6, and 7–9) of the compulsory education. This part is to engage students in disciplinary practices, that is, to encourage students to pose meaningful mathematics questions, to work cooperatively, to use mathematical knowledge and skills to describe an observed phenomenon, to solve problems, and to experience the importance and value of mathematics in everyday life.

In terms of methods of teaching and learning, consistent with the three objectives, the new mathematics standards promote instructional methods that facilitate the processes of knowledge construction and knowledge application by motivating students to participate, to collect and process information, to analyze and solve problems, and to communicate and cooperate with others. For example, the new curriculum prompts more use of an instructional sequence which starts with a realistic problem which affords mathematics relations to be learned, demands a mathematical model to describe the quantitative relation posed by the problem, requires explanation or verification of the mathematics model, and then extends the model to a new problem. This instructional approach is intended to engage students in the process of knowledge construction. This is in sharp contrast to the dominant instructional approach of the past, which starts with a given definition or an axiom, then is followed with a few exemplars to illustrate the definition or the axiom, followed by repeated exercises for students to do in order to achieve high proficiency in accurate recall of memorized mathematics definitions, formulas, rules, and procedures and facile matching of solution patterns to types of problems. The definition/axiom – exemplar – exercise instruction sequence is efficient in transmitting knowledge but less productive in developing students’ competence in knowledge construction and knowledge use.

It is understood that assessment of curriculum involves more than student assessment. However, only this aspect is discussed here. The new mathematics curriculum standards promote the student assessments that are competence-oriented, process-oriented, and student learning/development-oriented. The competence-oriented aspect requires the focus of assessment to be on competence to use knowledge and skills to solve problems. Connected with this, student assessments are expected to encourage students to engage in active cognitive and affective processes, such as observing, reasoning, hypothesizing, and communicating. Therefore, assessment tasks should demand interconnected understanding of concepts, procedures, and principles and allow for multiple representations and multiple solution strategies. Thus, various forms of assessment tasks need to be used, such as open-ended questions, problem-based projects, and learning journals, instead of relying on multiple choice questions. Self-assessment and peer assessments can be also used to help students become independent and reflective learners. The competence-, process- and development-oriented student assessments are not only to assess what students have learned and can do, but also to make the assessments themselves opportunities for students to learn and to grow mathematically and personally.

The new orientations of the mathematics curriculum standards in terms of objectives, content, instructional methods and student assessment are meant to achieve a better balance between learning outcomes and learning processes. These learning outcomes and processes include: inductive and deductive reasoning, learning mathematics in context, and
learning mathematics as a discipline, and emphasizing students as the agents of learning and teachers as the organizer and facilitator of learning. They are a result of taking lessons from both the traditions of Chinese mathematics teaching and learning (Fan, Chen, Zhu, Qiu, & Hu, 2004; Liu & Sun, 2002) and the theories and practices of mathematics curriculum reforms of many other countries, such as the USA (NCTM, 1998, 2000; Stein, Remillard, & Smith, 2007), Singapore and Japan (Chang, 2008). (For a detailed description and analysis of the see Chang, 2008; Liu & Sun, 2002; Ministry of Education, 2001a,b.)

1.2. Purposes of the study

Since the adoption of the new mathematics curriculum standards, there has been a debate in the Chinese educational research community about the impact of the three objectives of mathematics curriculum (knowledge and skills, processes and methods, affective demeanor and value) on student learning outcomes. The concern is whether or not the strength of the Chinese students’ proficiency in the two basics – basic mathematical concepts and basic mathematical skills – would be compromised with the new curriculum, which requires a more balanced treatment of the three objectives and devotes more instructional resources for students to experience, to construct and to communicate while learning mathematics. Opponents of the new curriculum argue that in theory, it is difficult to differentiate different roles of the three objectives. They assert that the dimension of knowledge and skills should have been acknowledged as the foundation of any mathematics curriculum (Wang, 2006). In fact, the new curriculum’s emphases suggest three changes: (1) a redistribution of instructional resources (e.g., classroom time), (2) less instructional resources given to direct instruction and repeated exercise, and (3) more instructional resources given to problem-solving, participation, and communication. Understandably, the concern is whether or not the new curriculum would weaken the foundation in the two basics, the traditional strength of Chinese students’ mathematics achievement. It seems unlikely that weakening the foundation would enhance the development of problem-solving competence (Wang, 2004, 2006). Therefore, the new curriculum is suspected for not only having impaired the traditional strength of Chinese students’ mathematics achievement but also have weakened students’ mathematics problem-solving competence (see also Li & Ni, 2007 for more details about the debate).

Of course, such concern about the efficacy of the reformed curriculum is not unique to China. For example, the efficacy of the NCTM mathematics curriculum has been questioned with regarding whether or not it would reduce students’ proficiency skills in basic computation. Klein (2002) documented three high-achieving Los Angeles elementary schools where a majority of the students were from low social and economic backgrounds and their mathematics achievement ranked highest on California’s Standardized Testing and Reporting. Their achievement was attributed to systematic, direct instruction and the instructional focus on basic skills in mathematics. Chubb and Loveless (2002), scholars at the Brookings Institute, suggested that the achievement of the students in these schools was due to their reliance on traditional instruction and curricular emphases:

The three elementary schools mainly resist faddish, cutting-edge approaches to teaching and relentlessly focus on teaching the core skills of reading and mathematics. Teachers instruct students on this knowledge and test regularly to make certain that pupils have learned it. In other words, the teachers in these schools teach students solid content. They do not facilitate, guide, or explore. Whole language is out, phonics is in. Calculators and National Council of Teachers of Mathematics math reform are out; arithmetic is in. (pp. 7–8)

The concerns raised by individuals such as Klein, Chubb and Loveless were addressed by Senk and Thompson (2003), who concluded that students using the new curricula generally perform as well as other students on traditional measures of mathematics achievement, including computational skill, and they generally do better on formal and informal assessments of conceptual understanding and ability to use mathematics to solve problems. The studies cited by Senk and Thompson, however, were criticized because the curriculum designers were also the primary researchers, thereby possibly biasing the results. But, according to Stein et al. (2007), the findings reported in the Senk and Thompson volume are similar to those of comparative studies by researchers who were not also the developers of the curricula.

The debate continues in China concerning whether or not the new curriculum would result in a weakening of Chinese students’ proficiency in the basic mathematics skills, while not adding any benefit to their competence in solving mathematics problems, particularly solving non-routine problems. What is needed to settle the debate is empirical evidence. However, there appears to be almost no systematic, large-scale empirical study in China. The present study was to address this empirical question.

Another issue addressed in the present study was whether or not the new curriculum would benefit all students from different social economic backgrounds. Many studies have shown significant impact from family socioeconomic status (SES) on students’ school achievements (Broeck, Opdeborah, Hermans, & Damme, 2003; Hill et al., 2004; Sirin, 2005). Consequently, the study was concerned with whether or not the different curricula would help reduce performance gaps between students from different family backgrounds (e.g., Chubb & Loveless, 2002).

In the context of China, there have been very few studies on the association between family SES and student school achievement. These studies have suggested that the relationship is dependent on the education level of specific populations. Jiang and Yan (2006) found a significant association in junior secondary school students, but Tao and Yang (2007) found no significant association in high school students and college students. The lack of an association at the high school and college
level might be due to the highly selective nature of high school education in China. In addition, several recent studies have examined the association between SES and access to education, as well as the association between SES and educational attainment (years of education). These relationships vary considerably across different historical periods of the country (Deng & Treiman, 1997; Li, 2004a). For example, during the period from 1949 to 1978, the relationship between SES and educational opportunity became weaker and reached the weakest during the Cultural Revolution. This was because the ideology of Community Party and the highly centralized governance supported strong measures to reduce the degree of social hierarchy in the country at that time. However, since the early 1980s, Mainland China has promoted a market-economy, one which is accompanied by less developed social welfare. Consequently, the associations have been becoming stronger with regards to access to high school education and college education along with the increasing gap in wealth between the rich and the poor (Cheng, 2002; Liu, 2004; Zhao, 2000).

At the present time, China is facing the challenge of increasing social inequality including education inequality (Xie, Li, Sun, & Wen, 2008). The new curriculum emphasizes the development of problem-solving skills and communication skills in students. These areas require more instructional support to develop. There is a concern as whether or not the new curriculum will be more equitable to students with different socioeconomic backgrounds because the new direction of the reformed curriculum appears to be more affordable for the students from families with more social and economic resources than those with fewer such resources (e.g., Hill et al., 2004). The present study also addressed this question.

1.3. Methodological clarifications

Comparative evaluation study is an evolving methodology (NRC, 2004, p. 96). There were two methodological issues about the present research that needed to be clarified from the outset. These issues are discussed below.

In the study the reform group was identified as students in the classrooms that adopted the new textbooks designated for the new curriculum by the Ministry of Education. It is understood that curriculum standards and textbooks are considered as the ‘intended curriculum’ (e.g., NRC, 2004) because how they are being implemented in different classrooms could be different from what the developers intended. In addition, there are many other important factors, such as teachers’ curriculum decisions, teaching behaviors, students’ individual differences, and classroom and school culture that determine how the intended curriculum is being enacted in classroom. Therefore, there may be a gap between the intended curriculum and the ‘enacted curriculum.’ However, curriculum materials are viewed as the main lever in helping to initiate and sustain reform because they are concrete, tangible vehicles for embodying the essential ideas of the reform. They are considered as a key tool for teachers learning about subject matter as well as instruction, especially for those from countries with a highly centralized curriculum system, like China and Singapore (Ball & Cohen, 1996; Lloyd & Frykholm, 2000; Wang & Paine, 2003). There are some particulars to be noted in the context of China. One is that the decision about what textbooks to be adopted in public schools is made by the Ministry of Education, not by the schools or school districts. In the present study, all the reform classrooms were using the same textbooks designated for the new curriculum, whereas the non-reform group used another set of officially designated textbooks for the conventional curriculum.

The centralized education system exercises administrative efficiency via a hierarchical administrative system to increase the likelihood of faithful implementation of a targeted curriculum. For example, the Ministry of Education assembled a team of curriculum and instruction experts who designed the new curriculum standards to give workshops and make classroom visits for the district mathematics curriculum coordinators and the head teachers of the schools adopting the new curriculum on the use of the new standards and the new textbooks. In addition, the schools adopting either curriculum used peer-teaching preparations to help teachers reach some consensus on how to use the corresponding curriculum standards and textbooks in classrooms. In this connection, Chinese teachers were shown to be more likely to use their textbooks faithfully, compared to American teachers (Cai, 2005; Li, 2004b). There is a strong culture in China to read and analyze the textbook and teacher manual (Zheng, 2002). Teachers make extensive use of the curriculum, and rely on the materials in decision making for content and ways of teaching (Fan et al., 2004). In addition, textbooks are used in the classroom on a regular basis. For example, more than 80% of eighth-grade Chinese teachers indicated that they used mathematics textbook all the time (Fan et al., 2005). Therefore, the influence of curriculum materials on teachers’ teaching and students’ learning is considered to be more homogeneous and stronger in a centralized educational system, such as in China, than in a non-centralized system, such as in US.

Another issue was that the two groups, reform group and non-reform group, in the study could not be assumed to be comparable from the very beginning when they started from the first grade receiving the different mathematics curriculum. The difficulty for this study was that we were unable to ‘equate’ the groups through a statistical treatment because there was no common achievement measure, which existed for the groups prior to the beginning of the study. During the study, the students of the two groups were administered mathematics achievement measures three times, the first time when the students were in fifth grade, the second and the third time when they were at the beginning and the end of sixth grade respectively. It was not possible to use the results of the first time assessment as a covariate to equate the groups because it was not conceptually appropriate to do so. The issue was that to ‘equate’ the groups with the fifth grade assessment results would ‘wipe away’ any differences between the groups that might be due to the possible accumulative influence of the different curricula during the students’ first five years of school. To lessen the effect of this limitation, the growth rate during the study period was the focus of the comparison because growth rate was assumed to be similar for the two groups.
2. Methods

2.1. Participants

The research data were collected in 2005–2007. The participants for the study included 60 classrooms from 20 schools, 3 classrooms per school, from a city in central China. There were 34 reform classrooms including 1959 students and 26 non-reform classrooms with 1456 students. None of the schools was selective and most students were assigned to the schools for their proximity to where they lived. The reform classrooms adopted the new curriculum and the designated textbooks for grades one through six. Likewise, the non-reform classrooms from grades one to six used the conventional curriculum and the corresponding textbooks. The students were in fifth grade when the data collection started in November 2005 and they were nearing the end of sixth grade when the data collection was completed in May 2007.

The two groups of teachers were considered relatively comparable (see Li & Ni, this issue). The school districts arranged the teaching assignments so that a group of teachers always taught grades 1–2, another group always taught grades 3–4, and another always taught grades 5–6. Therefore, most of the teachers had just started to teach the classes when the study began.

2.2. Measures

We decided that student learning outcomes in mathematics would include cognitive aspects such as skills of mathematical calculations and skills in explaining and communicating mathematics, as well as affective aspects such as interest and dispositions towards learning mathematics.

2.2.1. Cognitive measures of mathematics achievement

The cognitive measures of mathematics achievement contained three parts: calculation, simple problem solving and complex problem solving. The first two parts containing all multiple choice (MC) questions, were developed based on the four cognitive processes identified by Mayer (1987, 2003) that are involved in solving math word problems: translation (i.e., converting word sentence into a numerical representation of the described situation); integration (i.e., selecting and combining information into a coherent representation of the given problem); planning (i.e., breaking down the problem to be solved into steps); and execution (i.e., carrying out mathematical operations). (See Cai et al., this issue, for sample items.) There were a total of 32 MC questions, six items for each of the first three dimensions and 14 items for the execution (calculation) dimension. The items of the first three dimensions were grouped as routine or simple problem solving and those of the last dimension as calculation. The classification was done based on the understanding that the items for execution mostly required calculations, whereas the items for the other three dimensions called for interpreting the problems or indicating steps for getting answers to the problems but no calculation was required. These items were intended to assess the ‘Two Basics’ described earlier (i.e., basic mathematical concepts and basic mathematical calculations).

Three separate common factor analyses were conducted on the student responses from three administrations of the measures to examine the assumed dimensions. Except for the dimension of planning whose items showed low loadings on any factors, a relative clear structure was shown for the other items. For example, for time 3 data, there were four clear factors: integration, translation involving fractions/percentage, translation involving integers, and calculation involving integers, fractions, decimals, and equations. Across the three data points, the total variances explained by the 4–6 factors ranged from 40 to 48%, and communities from 10% to 12%, which was reasonable. Based on the results, except for the planning items, all the items were retained for further analyses. The retained items were grouped into the two categories, calculation and routine problem solving as explained above.

For the complex problem-solving dimension, a total of 12 open-ended tasks were developed. In responding to the open-ended tasks, students were required to show their solution processes and provide justifications for their answers. An example of open-ended question is as follows:

Ming and Fang, high school students, take a part-time job. Ming earns 15 RMB per day and Fang earns 10 RMB per day. (1) How many days do Ming and Fang have to work respectively so that they will earn the same amount of money? Show how you found your answer. (2) This problem has more than one answer. Find another answer and explain. (Detailed descriptions of the open-ended problems can be found in Cai (1995, 2000).)

The responses to the open-ended questions were evaluated with a scoring rubric using a 0–4 point scale: not acceptable (0), minimal (1), satisfactory (2), good (3) and excellent (4). The scorers included elementary school mathematics teachers and graduate students in educational psychology or in mathematics. While scoring, the scorers did not know the identity of individual students, the group to which the students belonged, or whether it was the reform group or the non-reform curriculum group being scored. Two scorers independently scored five percent of the student responses. The inter-scorer agreements were 0.876, 0.891, and 0.880 respectively for the three administrations of the open-ended questions.

The correlations between the three measures of mathematics achievement, calculation, simple problem solving, and complex problem solving, ranged from 0.42 to 0.80. The correlations between the three times of assessment were 0.31–0.34 for the measure of computation, 0.32–0.37 for simple problem solving, and 0.50–0.58 for complex problem solving.
2.2.2. Affective measures of mathematics achievement

Four facets of affective mathematics achievement (Li, 2004a,b) were included: (1) students' perceived interest in mathematics (e.g., 'Mathematics interests me because I find it stimulating to solve a math problem.'), (2) students' perceived participation in math classroom (e.g., 'I feel anxious when I sit in a math class.'), (3) students' ideas about what mathematics is about (e.g., 'Mathematics is about numbers and their computations.'), and (4) students' views of learning mathematics as a process of reasoning and reflection (e.g., 'I sometimes stick to my own thoughts even though my thoughts are wrong because of the fear to reveal my weakness.'). The first two scales are concerned with aspects of attitudes and the other with aspects of beliefs about learning mathematics (McLeod, 1990). The questionnaire included 35 items to measure the four aspects. A Likert scale was used having five levels: 'strongly disagree,' 'disagree,' 'not sure,' 'agree,' and strongly agree.' For the three administrations of the questionnaire with a sample of students, alpha values for the scale of interest ranged from 0.87 to 0.92; for classroom participation from 0.80 to 0.85; for views of mathematics from 0.65 to 0.72; and for views of learning mathematics from 0.61 to 0.70. The test–retest correlations for the four scales were 0.56–0.70 for the scale of interest, 0.57–0.68 for classroom participation, 0.48–0.59 for views of mathematics, and 0.48–0.58 for views of learning mathematics.

2.2.3. Measure of family socioeconomic status

The students' parents were asked to complete a questionnaire about the family's socio-economic status. The indicators of the social economic status included family income, father and mother's educational levels, and father and mother's occupations. The index values for the indicated occupations by the parents were calculated according to the Standard International Socio-economic Index of Occupational Status (Ganzeboom, De Graaf, Treiman, & De Leeuw, 1992). Through a factor analysis, the parents' income, occupations, education levels were aggregated into the latent variable, SES, which was standardized to be used in the data analyses.

2.3. Procedures

The measures of mathematics achievement were administered to the students on three occasions over a period of 18 months. The first administration took place during the first term of fifth grade, the second at the beginning of sixth grade, and the last one at the end of sixth grade. Each administration required two class sessions, a total of one and a half hours, for the students to complete the measures. The order of administering the measures each time was as follows: computation and simple problem solving with MC questions, complex problem solving of open-ended questions, and the questionnaire inquiring the affective aspects of mathematics learning outcome. The administration of the measures was undertaken by two research assistants in each classroom.

Due to time limitations, a matrix sampling method was used to administer the open-ended questions. Two forms of questions were prepared. Within every classroom and for each administration, half of the students answered form A containing six of the questions and the other half of the students answered form B containing, the remaining six questions. The students who answered form A for the first time were administered form B the second time, and vice versa. For the third administration, the students were given the form they had in the first administration. In addition, eight different items were used in the measure involving the MC questions between the first administration and the later two administrations. Consequently, it was necessary to use equating procedures to equate the two test forms of the MC calculation questions, MC routine questions, and the open-ended questions respectively. This would ensure that direct comparison could be made of the scores resulting from the three administrations (Kolen & Brennan, 2004).

3. Results

A three-level model was employed to capture the change in mathematics achievement of the students who were nested within classrooms. Individual growth trajectories comprised the level-1 model, the variation in growth parameters among the students within a classroom was represented in the level-2 model, and the variation among the classrooms was reflected in the level-3 model (Raudenbush & Bryk, 2002).

The equations for the three-level models are as follows.

---

1. In the division of the 12 open-ended mathematics questions into A form and B form, an error occurred. Apparently, an unequal number of process-constrained questions and process-open questions were assigned to the two forms. This resulted in one form having four process-open questions and two process-constrained questions and the other having two process-open questions and four process-constrained questions.

2. In the test equating, the test form of the multiple-choice questions used in time 1 was as form A and the form used in time 2 and time 3 as form B. For each test form, the test length for the calculation part was 14 items and there were 8 common items. The length for the part of routine problems was 12 items and there were 6 common items. Therefore, a procedure for common items non-equivalent groups equating was used to develop a common metric for linking the two test forms for the calculation part and the routine problem solving part respectively (Zimowski, Muraki, Mislevy, & Bock, 1996).

3. In equating the two forms of the open-ended questions, it was assumed that two random groups were selected to take the test form A and form B respectively. Hence, pre-smoothing and post-smoothing equipercentile equating methods were used to link the two test forms (Kolen & Brennan, 2004), so that the scores for form B were equated onto the score scale of form A.
Level-1 model:
\[ Y_{ij} = \pi_{0ij} + \pi_{1ij}(TIME)_{ij} + e_{ij} \]

Level-2 model:
\[ \pi_{0ij} = \beta_{00j} + \beta_{01j}(SES) + r_{0ij} \]
\[ \pi_{1ij} = \beta_{10j} + \beta_{11j}(SES) + r_{1ij} \]

Level-3 model
\[ \beta_{00j} = \gamma_{000} + \gamma_{001}(CURRICULUM) + u_{00j} \]
\[ \beta_{01j} = \gamma_{010} \]
\[ \beta_{10j} = \gamma_{100} + \gamma_{101}(CURRICULUM) + u_{10j} \]
\[ \beta_{11j} = \gamma_{110} + \gamma_{111}(CURRICULUM) \]

In the level-1 model, \( \pi_{0ij} \) is an outcome measure at the initial time for student \( i \) in class \( j \); and \( \pi_{1ij} \) is the slope for the change of performance on the measure in student \( i \) in class \( j \) during the measured period. In the level-2 model, \( \beta_{00j} \) represents the mean initial status within class \( j \); \( \beta_{10j} \) is the mean time slope within class \( j \); \( \beta_{01j}(SES) \) and \( \beta_{11j}(SES) \) stand for the hypothesized effect of students' family social and economic status on the initial status and the growth rate of individual students' mathematics achievement. In the level-3 model, \( \gamma_{000} \) is the overall mean initial status of the classes; while \( \gamma_{100} \) is the overall mean time slope of the classes. \( \gamma_{001}(CURRICULUM) \) and \( \gamma_{101}(CURRICULUM) \) represent the hypothesized effect of curriculum on the initial status and growth rate of the classes respectively. \( \gamma_{010} \) and \( \gamma_{110} \) refer to the effect of students’ family SES on the overall initial status and the growth rate of mathematics achievement of the classes. However, it is assumed that the effect of SES on the initial status of individual students within class was similar. Therefore, \( \gamma_{010} \) equals \( \beta_{01j} \). The term \( \gamma_{111}(CURRICULUM) \) represents a hypothesized effect of curriculum on the influence of SES on the growth rate between classes.

The equations were applied to the dependent variables including the three measures of cognitive achievement (calculation, routine problem solving, and complex problem solving) and the four measures of affective achievement (interest, participation, views of mathematics, and views of learning mathematics). HLM version 6 was used to perform the analyses. Missing values for calculation, simple problem solving, complex problem solving, mathematics interest, classroom participation, views of mathematics, and views of learning mathematics for the first, second, and third assessments, were 16, 237, 0, 2, 2, 2, 16, 214, 0, 4, 4, 4, 4, 15, 166,0, 16, 16, 16, 16, respectively. All the missing values were replaced with the average score of each time point for the group that the individuals belonged to.

The descriptive statistics of the outcome measures are displayed in Table 1. Table 2 presents the variances of the outcome measures located at the level-2 and level-3 models respectively, indicating difference between classrooms were very small, 4.11–11.30% for the cognitive outcomes and 1.67–7.46% for the affective outcomes.

### 3.1. Cognitive measures

The results of the HLM analyses with the cognitive learning outcomes are presented in Table 3. The results indicate that there were overall significant differences in the initial status and in growth rate among the individual students' cognitive mathematics achievement. The significant changes were all positive for the measures, showing an overall improved performance among all the students over the time on computation, routine problem solving, and complex problem solving.

Concerning the influence of curriculum on the learning outcomes, for the first assessment the performance of the reform and the conventional groups approached a significant difference on computation (\( \gamma_{010} = 0.272, t = 1.946, p < .05 \), see Table 3) and was significantly different on complex problem solving (\( \gamma_{001} = 0.695, t = 2.250, p < .05 \), see Table 3) with the reform group performing better, but they did not differ on routine problem solving (\( \gamma_{010} = -0.007, t = -0.051 \)). The two groups progressed at different rates on the measure of computation; and the conventional group showed a significantly faster rate of improvement (\( \gamma_{101} = -0.691, t = -7.815, p < .001 \)). The two groups did not differ significantly in growth rate on routine problem solving and complex problem solving.

Student family SES demonstrated a strong effect on their performance in the initial assessment of computation (\( \gamma_{010} = 0.399, t = 8.783, p < .001 \)) as well as on the change rate over the time (\( \gamma_{101} = -0.157, t = -5.116, p < .001 \)). High SES students performed better than students of lower SES in the first assessment of computation. However, low SES students progressed faster than students of higher SES in computation. Moreover, the influence of SES on the growth rate varied between the two groups receiving the different curricula (\( \gamma_{111} = 0.130, t = 3.956, p < .001 \)). As shown in Fig. 1, the slope showing the narrowed performance gap between high SES and low SES students was steeper for the non-reform students than for the reform students. For the measure of routine problem solving and complex problem solving, high SES students outperformed those of lower SES on the first assessment (\( \gamma_{010} = 0.343, t = 7.916, p < .001 \); \( \gamma_{110} = 0.956, t = 10.194, p < .001 \)).
Table 1
Means and standard deviations of the outcome variables.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Max.</th>
<th>Reform group (n = 1959)</th>
<th>Nonreform group (n = 1456)</th>
<th>Total (N = 3415)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>1st assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calculation</td>
<td>14</td>
<td>11.23</td>
<td>2.36</td>
<td>10.56</td>
</tr>
<tr>
<td>Simple problem solving</td>
<td>12</td>
<td>9.30</td>
<td>2.51</td>
<td>9.29</td>
</tr>
<tr>
<td>Complex problem solving</td>
<td>24</td>
<td>16.23</td>
<td>5.11</td>
<td>15.09</td>
</tr>
<tr>
<td>Mathematics interest</td>
<td>5</td>
<td>4.27</td>
<td>0.67</td>
<td>4.13</td>
</tr>
<tr>
<td>Classroom participation</td>
<td>5</td>
<td>3.75</td>
<td>0.86</td>
<td>3.71</td>
</tr>
<tr>
<td>Views of mathematics</td>
<td>5</td>
<td>3.69</td>
<td>0.58</td>
<td>3.71</td>
</tr>
<tr>
<td>Views of learning math</td>
<td>5</td>
<td>3.99</td>
<td>0.61</td>
<td>3.92</td>
</tr>
<tr>
<td>2nd assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calculation</td>
<td>14</td>
<td>11.66</td>
<td>2.16</td>
<td>12.45</td>
</tr>
<tr>
<td>Simple problem solving</td>
<td>12</td>
<td>9.93</td>
<td>1.80</td>
<td>9.58</td>
</tr>
<tr>
<td>Complex problem solving</td>
<td>24</td>
<td>18.32</td>
<td>4.62</td>
<td>17.82</td>
</tr>
<tr>
<td>Mathematics interest</td>
<td>5</td>
<td>4.12</td>
<td>0.77</td>
<td>4.07</td>
</tr>
<tr>
<td>Classroom participation</td>
<td>5</td>
<td>3.67</td>
<td>0.91</td>
<td>3.65</td>
</tr>
<tr>
<td>Views of mathematics</td>
<td>5</td>
<td>3.79</td>
<td>0.61</td>
<td>3.80</td>
</tr>
<tr>
<td>Views of learning math</td>
<td>5</td>
<td>3.95</td>
<td>0.64</td>
<td>3.95</td>
</tr>
<tr>
<td>3rd assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calculation</td>
<td>14</td>
<td>12.13</td>
<td>1.98</td>
<td>12.94</td>
</tr>
<tr>
<td>Simple problem solving</td>
<td>12</td>
<td>10.33</td>
<td>1.80</td>
<td>10.43</td>
</tr>
<tr>
<td>Complex problem solving</td>
<td>24</td>
<td>19.62</td>
<td>4.32</td>
<td>19.50</td>
</tr>
<tr>
<td>Mathematics interest</td>
<td>5</td>
<td>4.13</td>
<td>0.80</td>
<td>3.97</td>
</tr>
<tr>
<td>Classroom participation</td>
<td>5</td>
<td>3.71</td>
<td>0.91</td>
<td>3.64</td>
</tr>
<tr>
<td>Views of mathematics</td>
<td>5</td>
<td>3.88</td>
<td>0.62</td>
<td>3.87</td>
</tr>
<tr>
<td>Views of learning math</td>
<td>5</td>
<td>3.99</td>
<td>0.66</td>
<td>3.93</td>
</tr>
</tbody>
</table>

Table 2
Variances explained at level 2 and level 3 for the outcome variables.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Level-2</th>
<th>Level-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculation</td>
<td>41.49%</td>
<td>11.30%</td>
</tr>
<tr>
<td>Simple problem solving</td>
<td>56.44%</td>
<td>4.11%</td>
</tr>
<tr>
<td>Complex problem solving</td>
<td>50.19%</td>
<td>6.59%</td>
</tr>
<tr>
<td>Interest in learning mathematics</td>
<td>56.96%</td>
<td>7.46%</td>
</tr>
<tr>
<td>Classroom participation in learning mathematics</td>
<td>64.10%</td>
<td>1.67%</td>
</tr>
<tr>
<td>View of mathematics</td>
<td>51.83%</td>
<td>3.86%</td>
</tr>
<tr>
<td>View of learning mathematics</td>
<td>52.96%</td>
<td>2.21%</td>
</tr>
</tbody>
</table>

Table 3
Three-level analysis of the cognitive learning outcomes in relation to curriculum and SES.

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Calculations</th>
<th>Routine problem solving</th>
<th>Complex problem solving</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>SE</td>
<td>t</td>
</tr>
<tr>
<td>Model for initial status, $\pi_{01}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model for mean status of non-reform students, $\beta_{010}$</td>
<td>10.860</td>
<td>0.106</td>
<td>102.716</td>
</tr>
<tr>
<td>$\pi_{01}$</td>
<td>0.272</td>
<td>0.140</td>
<td>1.946</td>
</tr>
<tr>
<td>Model for SES on initial status, $\beta_{01j}$</td>
<td>0.399</td>
<td>0.045</td>
<td>8.783</td>
</tr>
<tr>
<td>Model for time slope, $\pi_{11i}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model for time slope of students, $\beta_{11i}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\pi_{11}$</td>
<td>-1.149</td>
<td>0.060</td>
<td>19.262</td>
</tr>
<tr>
<td>$\pi_{11}$</td>
<td>-0.691</td>
<td>0.088</td>
<td>-7.815</td>
</tr>
<tr>
<td>Model for SES on time slope of students, $\beta_{110}$</td>
<td>-0.157</td>
<td>0.030</td>
<td>-5.240</td>
</tr>
<tr>
<td>Model for SES on time slope of students, $\beta_{110}$</td>
<td>-0.157</td>
<td>0.030</td>
<td>-5.240</td>
</tr>
<tr>
<td>Model for SES on SES's influence on time slope, $\pi_{11}$</td>
<td>0.130</td>
<td>0.033</td>
<td>3.956</td>
</tr>
</tbody>
</table>

However, low SES students improved more than high SES students over the time on the measure of routine problem solving ($\gamma_{110} = -0.069$, $t = -2.369$, $p < .01$), which was similar between the reform and the non-reform group. Student SES did not significantly affect growth rate of student performance on complex problem solving; and there was no effect of curriculum on the influence of SES on the growth rate.

For the cognitive measures, many of the estimated variances ($\tau_{0ij}$, $\tau_{1ij}$, $\tau_{00j}$, $\tau_{10j}$) and related $\chi^2$ (these statistics are not displayed because of space limitations) from the three-level decomposition, suggest that residual parameter variances still
remain to be explained. However, referring to the estimated variances explained at level-2 and level-3 (see Table 2), the actual sizes of difference between classes in the outcome measures were relatively small (4.11–11.30%).

3.2. Affective measures

The results of the HLM analyses with the affective learning outcomes – expressed interest in learning mathematics, mathematics classroom participation, views of mathematics, and views of mathematics learning – are presented in Table 4. There were overall significant differences in the initial status for the individual students’ on all the affective measures. However, the growth rate was significant only for the measure of learning interest and views of mathematics but not for that of classroom participation and views of learning mathematics. Students’ expressed interest in learning mathematics declined in general over the observation period ($\gamma_{100} = -0.085$, $t = -4.302$, $p < .001$). Students’ view of mathematics tended to change to a more dynamic view of mathematics ($\gamma_{100} = 0.067$, $t = 6.586$, $p < .001$).

Concerning the curricular influence, the non-reform group indicated a more dynamic view of mathematics than the reform group did on the first assessment ($\gamma_{001} = -0.071$, $t = -2.427$). Besides this, curriculum showed neither a significant effect on the initial status nor the growth rates of the affective learning outcomes (see the statistics for $\gamma_{010}$ and $\gamma_{101}$ in Table 4).

Student SES showed a strong effect on the initial assessment on interest in learning mathematics ($\gamma_{010} = 0.056$, $t = 3.905$, $p < .001$), classroom participation ($\gamma_{100} = 0.101$, $t = 5.721$, $p < .001$), views of mathematics ($\gamma_{010} = 0.092$, $t = 7.430$, $p < .001$), and views of learning mathematics ($\gamma_{010} = 0.077$, $t = 5.968$, $p < .001$). The influence of SES on students’ expressed interest in learning mathematics varied between the groups ($\gamma_{111} = 0.030$, $t = 1.969$, $p < .01$). Fig. 2 shows that expressed interest in learning mathematics by both high and low SES students of the reform group stopped declining from time 2 to time 3,
Table 4
Three-level analysis of the affective learning outcomes in relation to curriculum and SES.

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>Interest in learning mathematics</th>
<th>Participation in mathematics classroom</th>
<th>Views of mathematics</th>
<th>Views of learning mathematics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>SE</td>
<td>t</td>
<td>Coef.</td>
</tr>
<tr>
<td>Model for initial status, $\pi_{0ij}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model for mean status of non-reform students, $\beta_{00j}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERCEPT, $\gamma_{000}$</td>
<td>4.151</td>
<td>0.038</td>
<td>109.580</td>
<td>3.730</td>
</tr>
<tr>
<td>CURRICULUM, $\gamma_{001}$</td>
<td>0.086</td>
<td>0.050</td>
<td>1.722</td>
<td>-0.021</td>
</tr>
<tr>
<td>Model for SES on initial status, $\beta_{01j}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERCEPT, $\gamma_{010}$</td>
<td>0.056</td>
<td>0.014</td>
<td>3.905</td>
<td>0.101</td>
</tr>
<tr>
<td>Model for time slope, $\pi_{1ij}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model for time slope of students, $\beta_{10j}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERCEPT, $\gamma_{100}$</td>
<td>-0.085</td>
<td>0.020</td>
<td>-4.302</td>
<td>-0.032</td>
</tr>
<tr>
<td>CURRICULUM, $\gamma_{101}$</td>
<td>0.011</td>
<td>0.026</td>
<td>0.423</td>
<td>0.012</td>
</tr>
<tr>
<td>Model for SES on time slope of students, $\beta_{11j}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INTERCEPT, $\gamma_{110}$</td>
<td>-0.012</td>
<td>0.013</td>
<td>-0.918</td>
<td>-0.002</td>
</tr>
<tr>
<td>CURRICULUM on SES’s influence on time slope, $\gamma_{111}$</td>
<td>0.030</td>
<td>0.015</td>
<td>1.967</td>
<td>0.021</td>
</tr>
</tbody>
</table>

whereas the interest of both the high and low SES students of the non-reform group continued to decline from time 2 to time 3.

Student SES also affected the change rate of the measured views of mathematics ($\gamma_{110} = -0.034$, $t = -3.459$, $p < .001$), suggesting that low SES students had bigger changes, compared to high SES students, towards a more dynamic view of mathematics. In addition, the effect of SES appeared to vary between the two groups ($\gamma_{111} = 0.038$, $t = 2.990$, $p < .01$). Both high and low SES students of the reform group indicated a more dynamic view towards mathematics over the time at a similar improvement rate, whereas high SES students of the non-reform group showed a slower pace of improvement than that the low SES students (see Fig. 3).

Again, for the affective measures, many of the estimated variances ($r_{0ij}$, $r_{ij}$, $u_{00ij}$, $u_{ij}$) and related $\chi^2$ (the statistics are not displayed here because of space limitations) from the three-level decomposition, suggest that residual parameter variances are still unexplained. However, the actual sizes of the difference between classes in the outcome measures were very small (1.67–7.46%, see Table 2).

4. Discussion

With regard to the cognitive achievements, the students of both groups showed significantly improved performance over the time on computation, routine problem solving, and complex problem solving. However, there were differentiated patterns of performance between the groups. With student family SES being controlled, on the first assessment the reform group performed better than the non-reform group on computation and complex problem solving; they had comparable performance on routine problem solving. The two groups did not differ in their achievement growth rates on routine and complex problem solving. However, the non-reform group outperformed the reform group on computation on the third assessment. SES had a strong effect on the students’ achievement on both the initial assessment and the growth rate. High SES students performed better than low SES students in general. However, low SES students displayed a faster pace of progression in their computation performance.

For the affective learning outcomes, there was a significant change in students’ expressed interest in learning mathematics, which showed a decline over time. However, students indicated a more dynamic view of mathematics. There was no significant change overall in the students’ indicated classroom participation and views of learning mathematics. Concerning the curricular influence, on the first assessment the reform group showed higher interest in learning mathematics but a less dynamic view of mathematics than the non-reform group did. The two groups did not differ in the measure of classroom participation and views
of learning mathematics. The growth rates for the two groups did not significantly differ on the four affective measures. Again, student SES significantly affected the students’ affective achievement on the first assessment; high SES students indicated greater interest in learning mathematics, higher classroom participation, more dynamic view of mathematics, and more active view of learning mathematics. However, student SES did not affect the change rate in the affective achievement except in the measure of views of mathematics, where low SES students showed a larger improvement, compared to high SES students, towards a more dynamic view of mathematics. The influence of SES on students’ expressed learning interest and on their indicated views of mathematics was shown to vary between the students receiving the different curricula. For the reform group the high SES and the low SES students changed at a similar pace, whereas for the non-reform group, the high SES students showed a slower rate of decline than that of the low SES students.

Before interpreting the results, it needs mentioning again that the two groups could not be assumed to be comparable from the very beginning of their schooling, as explained earlier. Also, there were other important variables (e.g., teacher knowledge, teachers’ experience with the new curriculum, quality of classroom instruction, school culture) that were not accounted for and would confound the present results, which is elaborated in discussing the limitations of the study below. However, the observed performance differences were considered to be in part related to the curricular differences considering that the students had been receiving the different curricula throughout their entire elementary school education from grade one to grade six. In addition, the highly centralized school system in terms of uniformly adopted curriculum standards and textbooks mandated by the government exercises relatively more homogenous curricular impacts on classroom teaching and learning in China compared to the school systems in less centralized nations (Cai, 2005; Fan et al., 2004; Li, 2004a,b; Stigler & Hiebert, 1999), as suggested by the small differences of the outcome measures between classes in the present study (see Table 2).

It was clear that the non-reform group showed faster growth in proficiency in computation skills from their fifth grade to the sixth grade and they outperformed the reform group students in the final assessment. Also, the reform group students kept their initial advantage in solving open-ended problems as they performed better than the non-reform group on the first assessment and the growth rates for the two groups were similar. Nevertheless, it could not be concluded from the present data that the reform group’s better performance on complex problem solving was due to the curriculum or to their better initial status.
There was no evidence to support the claim that the new curriculum is weakening the students’ proficiency in the basic mathematics skills and not benefiting their competence in solving mathematics problems, particularly non-routine problems. Although the students in the reform group did less well than the non-reform group, the reform group still performed reasonably well on the measures of computation. Specifically, the percent correct on the measure (14 items for computation) for the reform group were 80%, 83%, and 87% respectively at the three points of observation, whereas those for the non-reform group were 75%, 89% and 92%. The findings are similar to those of studies conducted in the United States. Several such studies have indicated that mathematics curricula that stress competencies such as problem solving and explanation, improve students’ competencies in these areas, but not at the expense of the development of basic mathematical concepts and computational skill (Briars & Resnick, 2000; Putnam, 2003; Riordan & Noyce, 2001; Senk & Thompson, 2003). In addition, the patterns of performance of the two groups on calculation, solving routine problems and solving complex problems suggest the reform group appeared to have achieved a relatively more balanced development in the cognitive areas of mathematics achievement (see Fig. 4). The results are encouraging in the sense that the new curriculum seems to provide a more balanced development in the different aspects of mathematics achievement.

Nevertheless, if a trade-off between competence in complex problem solving and that of basic mathematics concepts and calculations does exist, it would appear to result from the redistribution of instructional resources. Based on the data of this study and considering the current Chinese context, we consider it worthwhile to allocate more instructional resources to classrooms to facilitate the development of students’ competence in mathematical representation, explanation, and communication. Cai (2000, 2004; Cai & Hwang, 2002) has reported that Chinese students outperformed U.S. students in tasks involving computation skills and routine problem solving, but performed less well than U.S. students on more open, creative problem-solving tasks. The differentiated performances in the different areas of mathematics achievement between U.S. students and Chinese students suggest a more balanced emphasis is needed in the Chinese mathematics curriculum, particularly in the areas of competence in applying knowledge and skills to solve open-ended mathematics problems. The differentiated patterns of performance by U.S. students and Chinese students also suggest that some different instructional conditions are required for developing the competence in solving routine mathematics problems and for the competence in solving non-routine mathematics problems (Cai, 2004; Wang & Lin, 2005).

Consistent with the literature, student SES had a strong effect on both cognitive and affective achievement in students. However, the achievement gaps in computation between students due to their family backgrounds were narrowed significantly, but there was no narrowing of the gap in either simple or complex problem solving. This was true for both

![Percent Correct by the Two Groups for the 1st Assessment](image1)

![Percent Correct by the Two Groups for the 3rd Assessment](image2)

**Fig. 4.** Percent correct on the types of problem by the two groups for the first and third assessments.
groups of students. The closing achievement gap in calculation but not in the problem-solving tasks indicates that instructional conditions that facilitate mathematical explaining, questioning, exchanging, and problem solving are most valuable for students from low SES families because low SES families are less likely to be able to afford the conditions to facilitate high-order thinking.

With respect to the affective learning outcomes, it is worth noting that the students’ expressed interest in learning mathematics declined from their fifth grade to sixth grade for both groups. However, a continuous decline in the measure was observed in the non-reform group over the 3 times of observation but not in the reform group, which showed no decline from the beginning to the end of sixth grade. The observation of students’ declining interest in mathematics was consistent with the literature that interest tends to decline as students move from elementary school to secondary school (McLeod, 1994). Nevertheless, there is good reason to be concerned about the attitudes of Chinese students towards learning mathematics. In 2000 PISA, Asian students, including those from Hong Kong, Japan, Taipei and Signore, were the top performers in the PISA assessment of mathematics literacy. However, the Asian students had very low self concept and interest in learning mathematics. The present results also suggest that the achievement gaps lie not only in the cognitive domains but also in the affective domains and the two seem to influence each other. Efforts to reduce the achievement gaps must go beyond skill and knowledge to increase students’ positive learning experience and positive feeling in the classroom. Attitudes and beliefs carry meanings for an individual. Therefore, affect may empower or disempower students in relation to mathematics (McLeod, 1990; Schoenfeld, 1992). Moreover, affective responses to mathematics are not only individual and internal; they also are shaped by social, institutional, and cultural contexts (DeBellis & Goldin, 2006). Therefore, the relationship between affective and cognitive dimensions of achievement, as well as with classroom interaction in mathematics learning need to be understood. Furthermore, curricular implementation and evaluation should take the significant affective learning outcomes seriously.

Several limitations of the study should be acknowledged. First, as indicated in the sections above, the students of the two groups receiving the different curricula could not be assumed to be comparable; that is, any differences of performance between the groups could not be attributed solely to the curriculum factor.

Second, the students participating in the study were from one city in China. This city’s economic and educational level is above the national average, but it does not belong to the group of the first-tier cities of China. The findings associated with the population might not be generalizable to other regions and cities of the country. Third, the time period of the longitudinal study was relatively short, with duration of only 18 months. Therefore, the study was not able to provide solid evidence about long-term influence of curriculum on student learning and development in mathematics.

Fourth, we did not construct the instruments based on the local curricula to measure the mathematics achievement of the students. Instead, we used Mayer’s (1987, 2003) scheme characterizing the cognitive processes of solving math word problems for the selection and construction of the measures for calculations and simple problem solving (Cai, 2000). The benefit of assessing students’ mathematical achievement based on the generic analysis of cognitive processes in solving math problems helped enhance the validity of this investigation. Specifically, it allowed for examining the curricular effects on the mathematical competence of students. This, however, might have compromised the relevancy of the measures to the specific content of the local curricula.

Finally, the characterization of the reform and the non-reform group in the present study were based on the curriculum standards and textbooks they adopted. We understand that it was the implemented curriculum in classroom, in addition to the written curriculum, that affected the students’ learning outcomes (Ball & Cohen, 1996; Stein et al., 2007). The quality of the implemented curriculum in classrooms is then associated with various important variables, such as teacher knowledge, teachers’ beliefs, teachers’ experience with the new curriculum, quality of classroom instruction, among others. For example, Li (2009) showed that teacher pedagogical content knowledge had a positive effect on students’ mathematics achievement. Baumert et al. (2010) demonstrated that the effect of teacher pedagogical content knowledge on student learning was also mediated by the teachers’ providing cognitive learning supports to individual students. Tarr et al. (2008) examined curricular influences on mathematics achievement in middle school students. They found that the examined curricula did not have a direct influence on student mathematics achievement but the influence was mediated by the measured classroom learning environments. We therefore acknowledge that the numerous factors, other than the adopted curriculum and textbooks, might affect the students’ mathematics learning outcomes. Indeed, the intervening variables might have confounded the present results and any interpretation of the results should take this into consideration. Efforts are being made by the research team to examine the ways of curriculum implementation in the classrooms. Another article in this issue (Li & Ni) reports how the teachers of the students implemented the curricula in terms of kinds of learning tasks they used and the ways they interacted with the students in teaching mathematics. A new study is being undertaken (Li, 2011) of how instructional variables would affect the students’ learning outcomes, connecting with the curricula.

Acknowledgements

The study reported in this article was presented at the annual meeting of the American Educational Research Association, Denver, USA, April 29–May 3, 2010. The study was supported by Research Grant Council of Hong Kong Special Administration Region, China (CERG-4624/05H; CERG-449807; CUHK Direct Grant-4450199) and the National Center for School Curriculum and Textbook Development, Ministry of Education of People’s Republic of China. The findings and opinions expressed in the report are those of the authors and do not necessarily reflect the positions of the funding agencies.
The researchers would like to thank the students, teachers, and parents of the participating schools and the education administrators of the school districts for their contributions to this study.

References


