Influence of the Reformed Mathematics Textbooks on Student Achievement in China

Zhonghe Wu
National University
California, U.S.A.

Abstract

The goal of this study was to examine the effects of the reformed textbooks on Chinese fourth graders' mathematics achievement through a balanced cognitive approach: Model-Strategy-Application (MSA). Data were collected from 4419 Chinese fourth graders in 36 schools from six cities in China in 2007. MANOVA and ANOVA tests were used to determine the effects of using the reformed textbook and non-reformed textbook on student achievement in the MSA areas. The results indicated a significant difference in three aspects of the MSA in student mathematics achievement between two groups. However, both groups had low score Strategy for computations. The finding suggests the need to balance mathematics pedagogy with teaching, learning, and assessment.

Key words: Mathematics modeling, strategy, application, mathematics proficiency, textbook, assessment

Introduction

In recent decades, the direction of Chinese mathematics education has shifted toward the Western way. It has been reformed through its adoption of reformed mathematics teaching and learning perspectives and approaches since the late 1990s, which has led to substantial changes in Chinese mathematics education (Ding, 1999). Research on how Chinese reformed curriculum influences teaching, learning, and assessment have received much attention (Tu & Wei, 2010; Wang & Murphy, 2004).Researchers have shown an increased interest in Chinese teachers' knowledge (An, 2004; An, Kulm, & Wu, 2004; Ma, 2001) and students' computation and problem solving skills (Cai, 2000). However, definitive research in contemporary international mathematics education is lacking regarding inquiries on how to assess Chinese student learning in a balanced way; how to identify differences between high and low achievement schools using the textbooks. Although the recent Program for International Student Assessment (PISA) reported Chinese (Shanghai) 15-year-old students in the top rank in reading, mathematics, and science (Fleischman, Hopstock, Pelczar, & Shelley, 2010), the studies on Chinese students' achievement at the elementary level as well as in different cities is still absent because Chinese elementary students did not participate in related international studies such as TIMSS study (Trends in International Mathematics and Science study).

Theoretical Framework

Reform Mathematics Education in China

Before the 1990s, China had a centralized mathematical curriculum focusing on rigorous, logical, and purely deductive reasoning. In order to meet the high demand of examinations, the curriculum placed more emphasis on fluency in computations. However, since the 1990s, reform of mathematics education in China hasled the national curriculum to be decentralized. Interestingly, the reform process includes conducting research on curricula in different countries, especially Western countries (An, Kulm, & Wu, 2004). By learning from others, China has developed new national mathematics standards and curricula, and these curricula are not only decentralized, but also focus on conceptual understanding and real world application including computation while also retaining Chinese cultural and educational characteristics (An, 2004). In July 2001, China issued the new *Compulsory Education Standards for Mathematics Curriculum* that embodied ideas of western mathematics education and provided a new direction for teaching and learning mathematics in grades 1-9. Since then, six different elementary mathematicstextbooks have been developed, approved by the Committee of Textbook at the Department of Education in China, and published country-wide (Xie, 2009). The reformed textbooks received great attention from the public and were widely used in China. For example, the elementary textbooks developed by the Curriculum & Teaching Materials Research Institute and published by People's Education was used by five million first grade students across the country equating to 20% of the same age groups in China in 2004 (Lu & Wang, 2004).

Assessment

Mathematics assessment is a process of gathering evidence about a student's knowledge and ability to use it, and making inferences from the evidence for a variety of purposes (NCTM, 1995). Mathematics assessments should measure a broad spectrum of mathematical content and processes (Schoenfeld, 2007). NCTM (2000) states five content areas that are required and at which K-12 students should be assessed to gauge their ability to gain knowledge and skills in mathematics: numbers and operations, algebra, geometry, measurement, and data analysis and probability. Student ability to use mathematics as indicators of cognitive development has been addressed by various studies. Although conceptual understanding, procedural fluency, and problem solving in application was highlighted in national standards and research studies, mathematics educators have struggled with the methods of assessing mathematics proficiency and have wondered if assessing mathematics proficiency is a complex art (Schoenfeld, 2007). For example, researchers recognized that assessing mathematical understanding is not an easy task and that it requires combining knowledge from a number of sources(Fennema &Franke, 1992; NRC, 1993; Wiliam, 2007).

While a variety of studies have examine and found that curriculum has a great impact on mathematical teaching and learning (Cai, Wang, Moyer, Wang, & Nie, 2011; Ni, Li, Li, & Zhang. 2011; Tarr, Reys, Reys, Chavez, Shih, & Osterlind, 2008), studies from Stein and Smith (1998) and Stein, Remillard, and Smith (2007) focused on the importance and quality of tasks in the text and how it influenced student learning. The report *Adding It Up* (National Research Council, 2001), indicated five components in mathematics proficiency: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and productive disposition. Similar to the National Research Council (2001), RAND (2003) also provides five indicators of student mathematics proficiency. Mathematics proficiency (National Research Council, 2001; RAND, 2003) was defined as follow:

1. *conceptual understanding*—comprehension of mathematical concepts, operations, and relations
2. *procedural fluency*—skill in carrying out procedures flexibly, accurately, efficiently, and appropriately
3. *strategic competence*—ability to formulate, represent, and solve mathematical problems
4. *adaptive reasoning*—capacity for logical thought, reflection, explanation, and justification
5. *productive disposition*—habitual inclination to see mathematics as sensible, useful, and worthwhile, coupled with a belief in diligence and one's own efficacy (p. 116).

These five components in mathematics proficiency show different levels of cognitive learning and assessment foci. Among them, conceptual understanding, procedural fluency, and problem solving in application are the essential components of the proficiency strands(California State Dept. of Education, 2006).

To assess student learning with understanding, McMillan (2007)further defines the types of learning targets for students as follows: 1) simple understanding; 2) deep understanding and reasoning; 3) skills, 4) products; and 5) affects. In general, mathematics assessment includes, but is not limited to, knowledge, skill, and attitudes. Based on mathematical proficiency studies and considering mathematical knowledge, skill, and attitudes, Wu and An (2006) developed a unique approach tousing the*Model– Strategy -Application* (MSA) cognitive approach to build a knowledge base for K-8 student to achieve mathematical proficiency. The three components of MSA are interrelated in a triangular network and build upon each other in which the model and strategy components form a foundation for application. Ignoring any ofthe three components will result in ineffective learning. Procedural fluency without conceptual understanding will yield non-meaningful and inappropriate strategies for solving applications; while conceptual understanding without procedural fluency will yield inefficient strategic applications (Wu & An, 2006).

The multiple aspects of the MSA assessment are supported by NCTM (2000), asdifferent students show what they know and are able to do in different ways. In addition, the three cognitive processes of the MSA are also consistent with the TIMSS's three cognitive domains of knowing, applying, and reasoning: Knowing relates to procedural fluency, applying connects to using visual models to aid conceptual understanding, and reasoning links to solving problemsin real world applications and it also relates all three aspects with more measurable features. According to TIMSS Mathematics Framework (Mullis, Martin, Gonzalez, & Chrostowski, 2004), TIMSS assessment is divided according to three domains: 1) knowing facts, procedures, and concepts; 2) applying knowledge to solve problems; and 3) reasoning beyond the solution of routine problems to encompass unfamiliar situations, complex contexts, and multi-step problems.

While TIMSS was designed to measure trends of mathematics curriculum and instruction by assessing students' mathematics achievements, TIMSS study does not measure student MSA ability directly; therefore the framework for this study not only measure student mathematics ability on TIMSS domains but also MSA domains because MSA is closely linked to mathematics proficiency (California State Dept. of Education, 2006; National Research Council, 2001; RAND, 2003; Wu & An, 2006).

Mathematics Textbooks in China

Compared to traditional Chinese mathematics textbooks, the reformed mathematics textbooks focus on the student's role in learning and values student innovation and problem solving skills. The reformed textbooks generally contain the following five features: 1) more knowledge is provided; 2) increase in the methods used to relate to students' life as a motivation tool to construct new structures of knowledge on the basis of acquired knowledge and experience; 3) the process of acquiring knowledge is emphasized; 4) an increased awareness of students' learning style; and 5) an increased focus on problem-solving throughout the experimental textbooks (Lu & Wang, 2004).

The goal of this study was to compare the differences between reformed mathematics text (called New Books [NB] in this study) and traditional mathematics text (called non-New Books [non-NB] in this study) and examine the effects of the reformed textbooks on Chinese fourth graders' mathematics achievement through a balanced cognitive approach: *Model-Strategy-Application* (MSA) under diverse settings of cities and school performance levels. The research questions were:

1. Is there a difference in Chinese 4th graders' mathematics proficiency as measured by model, strategy, and application between theNB and non-NB groups?
2. Do Chinese 4th graders' mathematics proficiency as measured by model, strategy, and application differ for those who use the NB versus those who use non-NB in different cities and schools at different academic achieve levels?

Method

Subjects and Data Collection

Data were collected from one of the six reformed textbooks in elementary school level. The NB textbook that was chosen was published by a Local Education Publisher and was developed by a group of mathematics teachers and teacher educators who had rich experience and expertise in teaching and learning mathematics at the elementary level. The NB connected student real world experience and mathematical ideas and included a variety of real world examples. The teaching objectives of the NB focused on developing students' knowledge and skills, mathematics thinking, problem solving, and affect and attitude. The NB not only considered mathematics foundation, generalization, and development, but also took into account characteristics of student age, psychology, and life experience at each grade level in order to meet diverse students' needs when arranging and selecting the content areas. This study focuses on Chinese 4th graders' use of textbooks. However, NB mathematics textbooks have the following common features across all grade levels:

1. Content areas include numbers and algebra, spatial thinking and graphic representation, statistics and probability, and comprehensive application.
2. Content arrangement (a) is based on a logical order of knowledge structure to arrange content structure of teaching, (b) is based on student developmental characteristics, and (c) shows an increased focus on comprehension and application of content.
3. Content development provides training in materials including student learning activities for teachers, which not only helps teachers to make a connection between mathematics and students' real life but also stimulates students' prior knowledge and experience that engages students in active learning.

This study selected the NB elementary textbooks as a main focus and compared student test scores in five mathematics content areas and three cognitive processes: conceptual understanding by *Modeling*, procedural fluency from using *Strategy*, and problem solving in real world *Application* (MSA).Schools were selected based on the NB usage. For instance, the city of Qindao was selected because there exists both reformed and traditional textbook usage. Data were collected from 4419 fourth graders (3036 used the NB and 1383 used the non-NB) in 36 schools from six cities in China in 2007. Between the 36 schools, there were variations in locations, the size of the cities, and the school's academic performance. Teachers were selected based on at least three years of teaching experience.

Schools were selected based on local record of performance levels, i.e., high level, middle level, and low level performance. It should be noted that student learning results will be impacted by multiple factors such as teaching methods, teachers' beliefs, and school climate. The focus of this study is on textbook influence of student learning results. Table 1 provides information about the sample groups.

Table 1 : Information regarding the Sample Groups

City and Location	Number of Schools	# of students in higher level schools	# of students in middle level schools	# of students in lower level schools	Total n
Cityof Qindao	3 NB	124	153	163	440
(Shandong Province)	3 Non-NB	162	149	159	470
Cityof Chuzhou	3 NB	155	138	150	443
(Anhui Province)	3 Non-NB	150	159	141	450
City of Wuzhou(Guangxi	3 NB	155	157	151	463
Province)	3 Non-NB	149	155	159	463
City of Wuxi	3 NB	100	61	39	200
(Jiangsu Province)	3 NB	99	92	101	292
City of Yancheng	3 NB	136	48	129	313
(Jiangsu Province)	3 NB	98	80	110	288
City of Yangzhou	3 NB	97	100	100	297
(Jiangsu Province)	3 NB	100	100	100	300
Total: 6	36	1525	1392	1502	4419

Aside from the collection of student achievements in each school, this study also collected data about students' mathematical attitudes. The results of this additional data analysis will be reported in another article.

Instruments

This study adopted the 4th grade mathematics test items released by the TIMSS. The process of test item is 1) Selection of the content to make content area more balanced in terms of content areas and MSA content; 2) Translated the test items to Chinese and conducted apilot study to test accuracy. The instrument for assessing student achievement in the study included 20 problems with two aspects: mathematics content areas and cognitive levels (See Table 2). The mathematics content areas consisted of numbers and operations, algebra patterns and relationships, measurement, geometry, and data analysis; the three cognitive levels had three aspects: knowing facts and procedures, using concepts, and problem solving routines and reasoning. The proportion of question types in the mathematics content areas was: 42% numbers and operations, 15% algebra patterns and relationships, 12% measurement, 19% geometry, and 12% data analysis. The proportion of question types in the cognitive levels areas was: 8% knowing facts and procedures, 27% understanding mathematics concepts, 57% problem solving, and 8% reasoning. A set of unified categories was developed according to the three aspects of the MSA assessment (See Table 3).

The 20 items were categorized into three aspects of the MSA domains for data analysis according to the following criteria delineated in Table 3. Before administering the assessment, the instrument was translated into Chinese by the authors. The content validity of the assessment was insured by the developmental care and external review of the instrument for the TIMSS. The Cronbach Reliability Test was conducted to assess the instrument for internal consistency and reliability. The Cronbach alpha coefficient was 0.820 for the instrument. The Cronbach alpha coefficient for MSA was 0.735.

Table 2: Content Areas and Cognitive Levels in MSA

Question	Content Area	Main Topic	TIMSS Cognitive Domain	MSA
1	Number	Fractions and decimals	Using Concepts	M
2	Algebra	Equations and formulas	Solving Routine Problems	A
3	Data	Data representation	Solving Routine Problems	A
4	Number	Fractions and decimals	Using Concepts	M
5	Number	Fractions and decimals	Solving Routine Problems	A
6	Number	Whole numbers	Solving Routine Problems	A
7	Number	Whole numbers	Solving Routine Problems	A
8	Algebra	Patterns	Solving Routine Problems	S
9	Measurement	Tools, techniques, and formulas	Solving Routine Problems	M
10	Algebra	Relationships	Knowing Facts and Procedures	S
11	Geometry	Locations and spatial relationships	Reasoning	M
12	Measurement	Attributes and units	Solving Routine Problems	A
13	Data	Data representation	Using Concepts	A
14	Data	Data representation	Solving Routine Problems	A
15	Algebra	Equations and formulas	Using Concepts	S
16	Measurement	Tools, techniques, and formulas	Reasoning	M
17	Geometry	Lines and angles	Knowing Facts and Procedures	M
18A	Geometry	Two- and three-dimensional shapes	Solving Routine Problems	M
18B	Geometry	Two- and three-dimensional shapes	Solving Routine Problems	M
18C	Geometry	Two- and three-dimensional shapes	Solving Routine Problems	M
19A	Number	Whole numbers	Using Concepts	S
19B	Number	Whole numbers	Using Concepts	S
19C	Number	Whole numbers	Using Concepts	S
20A	Number	Whole numbers	Solving Routine Problems	S
20B	Number	Whole numbers	Solving Routine Problems	S
20C	Number	Whole numbers	Solving Routine Problems	S

Table 3: The MSA Assessment Criteria and Items

	MSA Assessment Criteria	Number of Items
M	Item focuses on understanding concepts or relations with visual representations	9
S	Item asks for carrying out procedures in problem solving	9
A	Item connects solving mathematical problems in real life situations	8

Data Analysis

This study combined qualitative and quantitative methods of data analysis. Qualitative data analysis consisted of summary in percentages in particular mathematics assessment items; quantitative data analysis of multivariate analysis of variance (MANOVA) and Univariate analyses of variance (ANOVA) were used to analyzethe mean scores of mathematics assessment items in the three MSA variables simultaneously between six cities and three school levels of 36 schools between the NB and non-NB groups.

A correlation test was used to find an association between the three aspects of the MSA assessment according to the guidelines from Cohen (1988). In order to limit the data coding bias, two researchers and a group of graduate students and Chinese mathematics teachers coded data and discussed any conflicting issues. It should be pointed out that although Hierarchical Linear Model (HLM) has many strengths such as cross-level moderator effects, the limitation of the HLM is that "HLM treats independent variables as random variables, and thus the possibility raises that independent variables can be correlated residuals" (Castro, 2002, p. 78), which is in violation of an assumption. Based on the data collection, this study did not select HLM as a main data analysis method.

Results

The findings indicated that student achievement in the three MSA aspects in the NB groupwere higher than those who were in the non-NB group between the six cities andthe three levels of the 36 schools. The three aspects of the MSA were positively correlated to each other.

Achievement on the MSA Areas between NB usage and Non-NB usage

To assess whether Chinese students showed differences in the MSA scores between schools using the NB and non-NB, a multivariate analysis of variance (MANOVA) was conducted. The MSA scores for the NB and non-NB groups were significantly different in three aspects of the MSA with Wilks' λ = .979, $F(3, 3674) = 26.289, p < .001$, $\eta^2 = 0.021$. This indicates that the linear composite of Chinese students' MSA scores differs between NB and non-NB groups. Univariate analyses of variance (ANOVAs) for each dependent variable were conducted as follow-up tests to the MANOVA. The results indicated that effects of using the NB and non-NB on three aspects of the MSA were significantly different (See Table 4). Overall, students in the NB group scored higher on all three MAS areas than students in the non-NB group. Table 5 and Figure 1 show that students in NB groups had the highest scores in Model, followed by Application. However, students in the non-NB group had the highest scores in Application, followed by Model. Interestingly, both groups had the lowest scores in Strategic computation.

Table 4 Effects of Using NB Textbook on MSA Scores

Source	Dependent Variable	df	F	*p*	η
NB	M_total	1	76.928	.000	.020
	S_total	1	19.599	.000	.005
	A_total	1	28.286	.000	.008

Table 5: Means and Standard Deviations for MSA Scores as a Function of Textbook

Textbooks		Model		Strategy		Application		
		M	*SD*	*M*	*SD*	*M*	*SD*	*n*
	Non--NB	6.569	1.19952	6.4060	1.55775	6.6023	1.41023	973
	NB	6.929	1.05758	6.6436	1.38974	6.8529	1.20211	2705
	Total	6.833	1.10817	6.5808	1.43971	6.7866	1.26515	3678

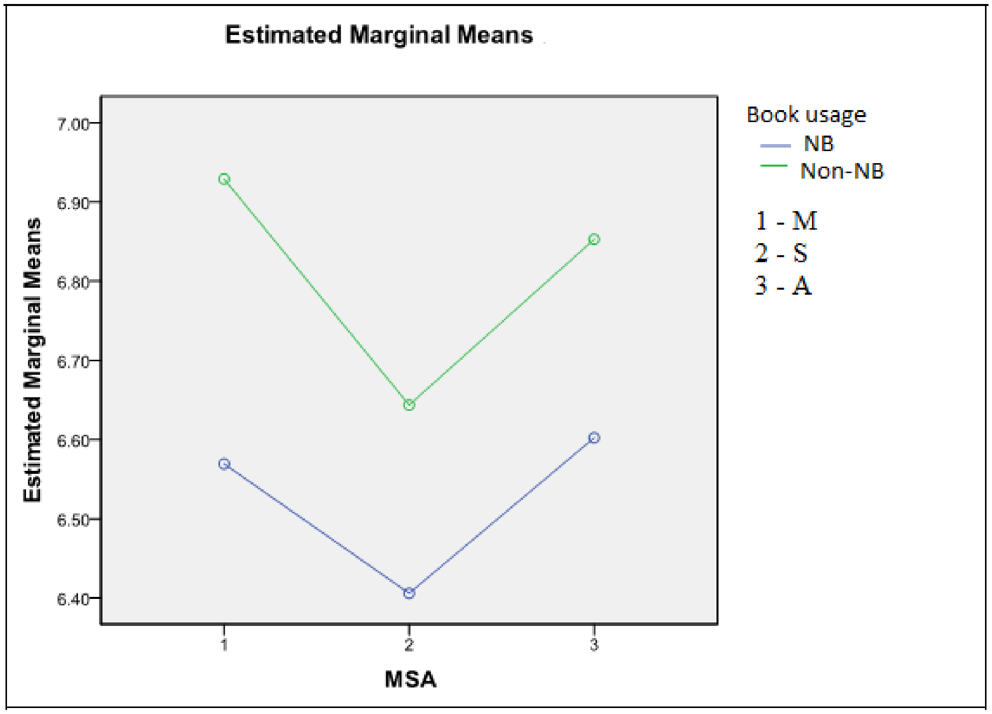


Figure 1.Comparison on MSA between NB and non-NB groups.

Achievement on the MSA between NB and Non-NB in Six Cities

A multivariate analysis of variance (MANOVA) was conducted to measure the effects of using the NB on the MSA scores between six cities and to test whether there was an interaction between textbook usage and cities. The interaction between using textbook and cities was significantly different with Wilks' λ = .995, F(6, 7334) = 28.609, p< .001, η^2 =0.023. The main effect for using textbook was significant, Wilks' λ = . 996, F(3, 3667) = 4.662[a], p< .001, η^2 =0.004. This indicates that the linear composite of the MSA scores differs for students between those using the NB and those not -using the NB textbook. The main effect for cities on the MSA was also significantly different with Wilks' λ = .926, F(15, 10123) = 19.160, p< .001, η^2 =0.025, indicating that the linear composite of the MSA scores differs for different cities. Univariate analyses of variance (ANOVAs) for each dependent variable were conducted as follow-up tests to the MANOVA. The results indicated that effects of students in both NB and non-NB groups and in the six cities were significantly different for three areas of the MSA (See Table 6).

Table 6: Effects of Using NBTextbooks and Cities on MSA Scores

Source	DV	Df	F	*p*	η
Cities	M_total	5	31.178	.000	.041
	S_total	5	33.143	.000	.043
	A_total	5	28.049	.000	.037
Book Usage	M_total	1	11.566	.001	.003
	S_total	1	.006	.938	.000
	A_total	1	1.442	.230	.000
Cities * Book Usage	M_total	2	39.761	.000	.021
	S_total	2	36.312	.000	.019
	A_total	2	71.736	.000	.038

Overall, the total mean scores comparison showed that the NB cities scored higher on all three MAS areas than non-NB cities: Model 6.929 vs.6.5694, Strategy 6.6436 vs. 6.4060, and Application 6.8529 vs. 6.6023.A comparison of the two results reveals that the NB group had about .36 points higher in Model than non-NB group (see Table 7).

Table 7: Means and Standard Deviations for MSA Scores as a Function of Using NB and Cities

		Model		Strategy		Application		
Cities	Book Usage	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation	N
Qindao	Non-NB	6.9242	1.08117	6.7551	1.41532	7.1061	1.08330	440
	NB	6.5303	1.14880	6.0212	1.54846	6.2970	1.38697	470
	Total	6.7452	1.12884	6.4215	1.52090	6.7383	1.29416	910
Chuzhou	Non-NB	6.6113	1.07629	6.4518	1.53683	6.5781	1.34340	443
	NB	6.8553	1.06149	6.6763	1.35045	6.8368	1.16669	450
	Total	6.7474	1.07413	6.5771	1.43906	6.7225	1.25355	893
Wuzhou	Non-NB	6.0145	1.28480	5.8551	1.62503	5.9058	1.58808	463
	NB	6.6533	1.18107	6.3498	1.38702	6.6533	1.21987	463
	Total	6.3589	1.26955	6.1219	1.52027	6.3088	1.44917	926
Wuxi	NB	6.9622	1.04684	6.6007	1.43531	6.8183	1.19357	492
	Total	6.9622	1.04684	6.6007	1.43531	6.8183	1.19357	492
Yancheng	NB	7.1831	.87787	7.1901	1.18297	7.0792	1.08982	601
	Total	7.1831	.87787	7.1901	1.18297	7.0792	1.08982	601
Yangzhou	NB	7.0858	.99263	6.6460	1.25338	7.1168	1.07534	597
	Total	7.0858	.99263	6.6460	1.25338	7.1168	1.07534	597
Total	Non-NB	6.5694	1.19952	6.4060	1.55775	6.6023	1.41023	1346
	NB	6.9290	1.05758	6.6436	1.38974	6.8529	1.20211	3073
	Total	6.8339	1.10817	6.5808	1.43971	6.7866	1.26515	4419

Figure 2 shows that the NB group had the highest score in Model across all cities. Qingdao had the lowest scores in all three areas, followed by Wuzhou. Figure 3 shows that the non-NB textbook group had the highest scores in Application in Qingdao, and the highest scores in Model in Chuzhou and Wuzhou. Comparing the three non-NB groups, Wuzhou had the lowest scores in all three areas in the MSA. Table 7 and Figures 2 and 3 show the most striking result to emerge from the data is that the two groups had a similar trend: Strategy had lowest scores.

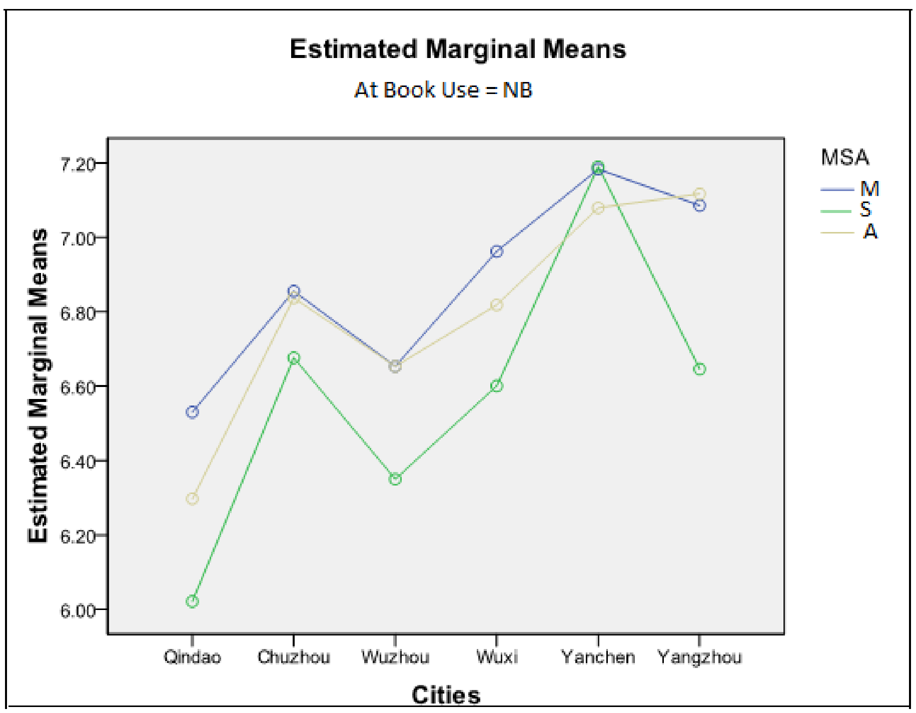


Figure 2. Comparison on MSA in the NB group in six cities.

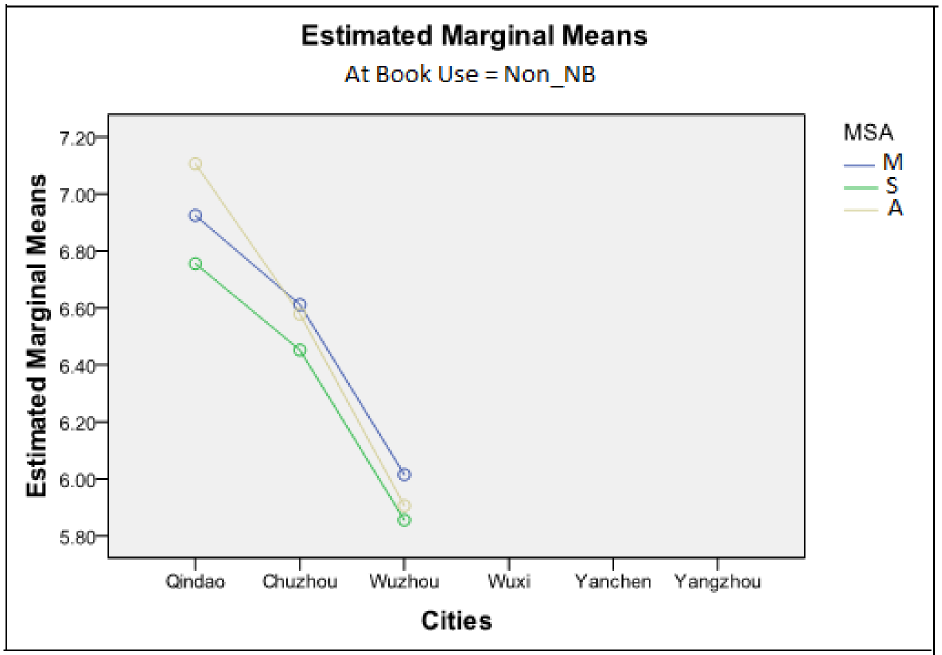


Figure 3. Comparison on MSA in Non-NB group in three cities.

Achievement on the MSA between Academic levels

To assess whether Chinese students in NB and non-NB groups in academiclevels of schools display different MSA scores and whether there was an interaction between the textbook usage and academiclevels, a multivariate analysis of variance (MANOVA) was conducted. The interaction between using text books and academic levels was significant with Wilks' λ = .979, $F(6, 7340) = 13.151$, $p< .001$, η^2 =0.011. The main effect for using textbooks was significant with Wilks' λ = .976, $F(3, 3670) = 30.429$, $p< .001$, η^2 =0.024. This indicates that the linear composite of the MSA scores differs for NB and non-NB students. The main effect for the academic levels was also significant with Wilks' λ = .957, $F(6, 7340) = 26.941$, $p< .001$, η^2 =0.022. This indicates that the linear composite of the MSA scores differs for different academic levels.

Follow-up ANOVAs (Table 8) indicate that the effects of using NB and academic levels were significant for three areas of the MSA Scores (See Table 8). The NB group had higher scores in all areas of the MSA at different academic levels. The NB group only had a slightly lower score than non-NB at the high performance level in the area of strategic computation (6.6007 vs. 6.7235). However, Post Hoc test (Tuket HSD) shows that this difference is not significant (Mean difference = -.1402, p <.039).

Table 8: Effects of Using NB Textbooks and Academic Level on the MSA Scores

Source	DV	Df	F	p	η
Academic Level	M_total	2	52.196	.000	.028
	S_total	2	16.610	.000	.009
	A_total	2	49.200	.000	.026
Book Use	M_total	1	88.943	.000	.024
	S_total	1	20.665	.000	.006
	A_total	1	31.212	.000	.008
Academic Level * Book Use	M_total	2	25.233	.000	.014
	S_total	2	22.899	.000	.012
	A_total	2	9.433	.000	.005

Table 9: Means and Standard Deviations for the MSA Scores as a Function of Using NB and Academic Levels

Academic Level	Book Use	Model Mean	Model Std. Deviation	Strategy Mean	Strategy Std. Deviation	Application Mean	Application Std. Deviation	N
Low Level Performance	Non-NB	6.5265	1.22023	6.2559	1.66230	6.4235	1.51201	459
	NB	6.6112	1.09311	6.3634	1.53103	6.4438	1.27573	1064
	Total	6.5881	1.12929	6.3341	1.56794	6.4383	1.34363	1525
Middle Level Performance	Non-NB	6.2491	1.29912	6.2116	1.58857	6.5085	1.43492	463
	NB	7.0190	1.03357	6.9504	1.31303	7.0222	1.16556	929
	Total	6.8372	1.14902	6.7760	1.41755	6.9009	1.25300	1392
High Level Performance	Non-NB	6.8882	.99520	6.7235	1.36541	6.8618	1.23912	461
	NB	7.1684	.95963	6.6007	1.23893	7.1013	1.03778	1043
	Total	7.0883	.97775	6.6358	1.27700	7.0328	1.10391	1502
Total	Non-NB	6.5694	1.19952	6.4060	1.55775	6.6023	1.41023	1383
	NB	6.9290	1.05758	6.6436	1.38974	6.8529	1.20211	3036
	Total	6.8339	1.10817	6.5808	1.43971	6.7866	1.26515	4419

Figure 4 shows that all three academic levels had the highest scores in Model, followed by Application in the NB group. However, all three academic levels had the lowest scores in Strategy in the NB textbook group.

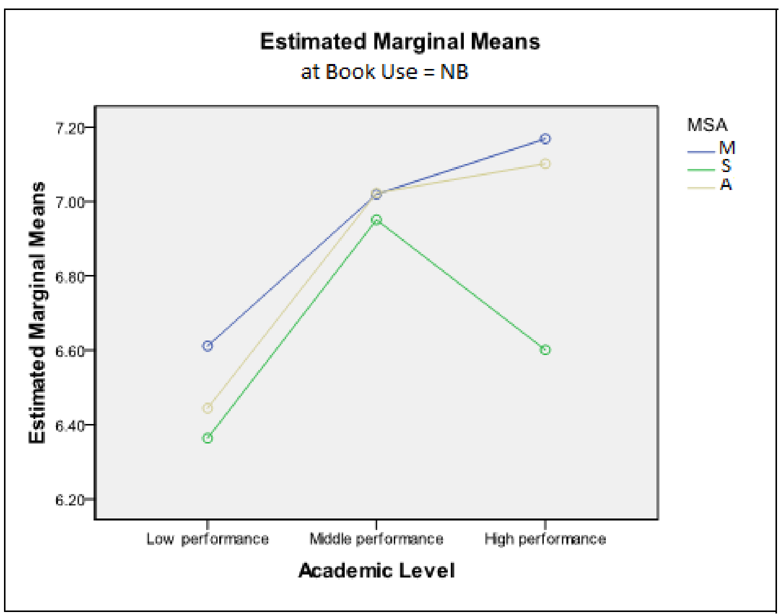


Figure 4. Comparison on the MSA in NB groups at three academic levels.

Figure 5 shows that the non-NB textbook group had various scorers at different levels: middle performance level had the highest scores in Application. Low and high performance levels had the highest scores in Model. All three types of academic levels had the lowest scores in Strategic computation.

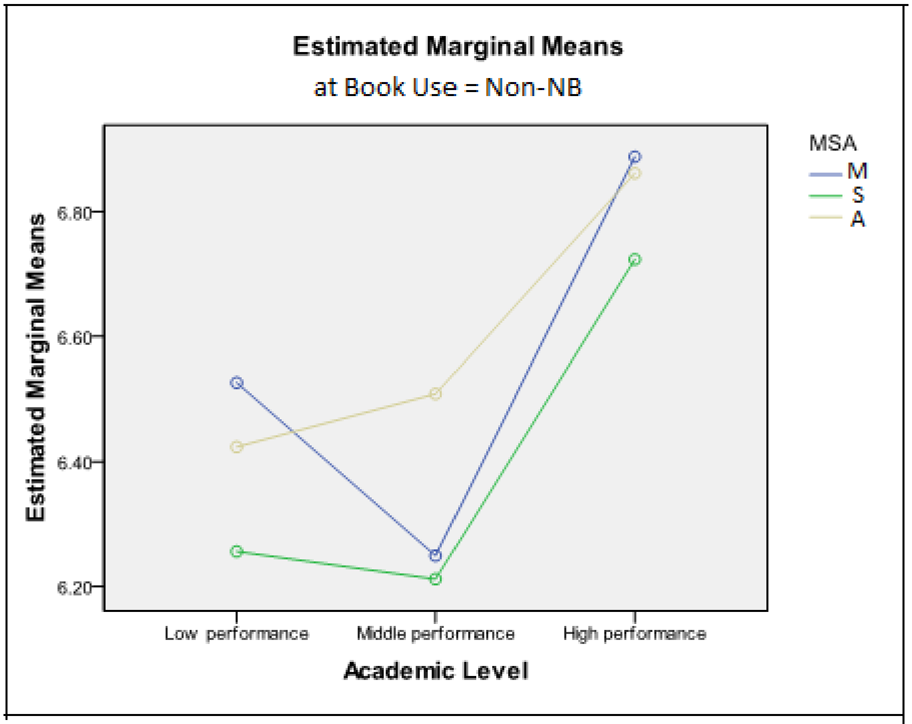
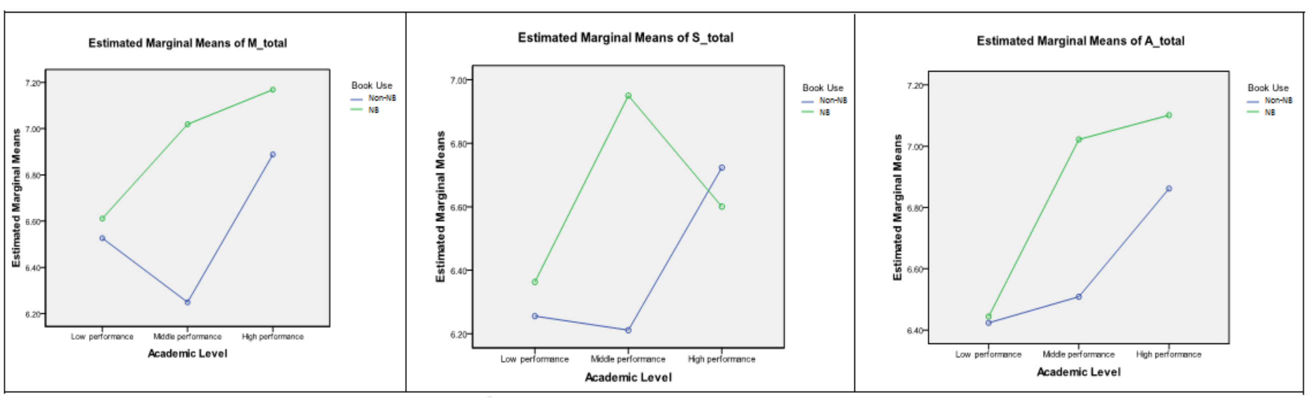


Figure 5. Comparison on the MSA in non-NB groups at three academic levels.

Figure 6 shows comparison of differences in three levels of academic groups (low performance, middle performance, and high performance) on MSA scores between the NB and Non-NB textbook groups. The middle level groups had the largest gaps between all three areas of MSA (See Figure 6 (a)-(c)). High level groups in both textbook groups had a larger gap in Model and Application, but their gap in Strategy is interesting because the NB scores were lower than the non-NB scores.



(a) (b) (c)

Figure 6.Comparison on MSA in textbook groups in three academic levels.

Discussion

This study examined the effects of the reformed textbooks on Chinese fourth graders' mathematics achievement through a balanced assessment approach: Model-Strategy-Application (MSA) model compared between cities and across diverse school academic levels. The results show that Chinese students had the higher scores in all three areas of the MSA in the NB groups; however, both NB and non-NB had relatively lower scores in Strategic computation compared to Modeling and Application. In addition, the results reveal the significant differences in the MSA in diverse cities and academic levels between the two textbook groups. There was a strong correlation between the three areas of the MSA in Chinese student achievement. The following is a discussion of these findings:

Chinese Students MSA between NB and non-NB Groups

There were significant differences in Chinese student achievement in the three areas of the MSA between NB and non-NB groups. NB students had higher scores in all three areas of the MSA than those in the non-NB group, with the highest scores being in Model and the second in Application. This indicates that the reformed Chinese mathematics textbooks place more emphases on modeling for conceptual understanding and real world application and have a more balanced curriculum compared to traditional textbooks (Wu, 2006). This also indicates tasks in the reformed textbooks associated with the maintenance of high level cognitive domains and content domains must provide students with opportunities to “do mathematics” (Stein & Smith, 1998; Stein, Remillard, & Smith, 2007), i.e., let mathematics makes sense to them.

However, the most interesting finding was that both groups had the lowest scores in Strategic computation for procedural fluency, which poses a new challenge for Chinese reformed textbooks in the area of mathematical proficiency (National Research Council, 2001; RAND, 2003). A possible explanation for this is that the reform in Chinese education has reduced the emphasis on computation, which reveals a drawback in the mathematics curriculum reform movement in Chinese mathematics education. In recent decades, various studies showed that the mathematical performance of Chinese students were higher than that of their international counterparts (An, 2004; Cai, 2000). While learning from Western ideas of mathematics education reform, Chinese mathematics education should also try to retain its strengths in computation, and achieve a balance within mathematics (California State Dept. of Education, 2006; Wu, 2006). Focusing on one and ignoring another will result in a lower proficiency in mathematics in general.

Chinese Students’ MSA between Diverse Cities, and Academic Levels

The findings from this study show that there were significant differences on student achievement as reflected in the MSA in diverse cities and across academic levels. Overall, students in the NB group scored higher than those in the non-NB group on all three MSA areas.

For both textbook groups (NB and non-NB), students had the highest scores in Model in all cities; Comparing the three non-NB groups, Wuzhou also had the lowest scores in all three areas of the MSA. This result may be at least partly explained by the fact that Qingdao is a large city in the Northeast of China and Wuzhou is a small city with a minority population in the Sothern region of China. The economic development in China has historically been unbalanced, which no doubt impact educational development. This finding showed that differences in socioeconomic, geographic, and educational resources may have influences on mathematics teaching and learning, and student achievement. This finding also showed such inequality issues might be a potential challenge for Chinese mathematics education (NCTM, 2000).

For students with different academic performance levels, findings show that the students in the NB group had higher scores in the areas of Modeling and Application, but their scores did not show a statistically significant differences in strategic computation from those in the non-NB group. Possible explanations might be 1) the teaching method has been influenced by mathematics education reform to focus less on strategic computation, which is the case for both groups; and 2) the reformed textbook has a similar impact as traditional textbook on student strategic computation (Tarr, Reys, Reys, Chavez, Shih, & Osterlind, 2008). Instructional strategies with the NB group seemed to have more focus on Model and Application, while the non-NB group seemed to be better in Strategic computation. In addition, the middle performance school group had higher gaps between NB and non-NB groups in all three areas of MSA; for the high performance school group, the differences between NB and non-MB was more dramatic, especially in Model and Application, which indicates that the effects of the new curriculum has had less of an impact on low performance school groups than in higher performing schools.

Correlation between the Three Areas of the MSA in Chinese Student Achievement

The results from this study show that there was a strong correlation between the three areas of the MSA in Chinese student achievement, with an especially strong correlation between Model and Application. This finding corroborates the ideas suggested by Wu’s (2006) MSA approach and Mcmillan’s (2007) findings that deep understanding of learning targets on knowledge, skill, and disposition. These approaches demonstrate the students’ inquiry learning process of understanding the content by using a variety of visual representations, finding the strategies and creating novel computational skills, and connecting concept knowledge and computational skills to solve real-world word problems.

In summary, findings from this study indicate that the reformed textbooks provide Chinese students with opportunities to learn mathematics with multiple approaches and in diverse contexts for a more balanced approach in modeling for conceptual understanding, strategy for procedural fluency, and application for problem solving. The findings also suggest that a well-implemented textbook not only produces higher achievement, but also promotes a balanced assessment.

Conclusion, Implications, and Limitations

This study compared Chinese student achievement through the MSA approach at the 4th grade level between traditional and reformed textbooks. The results of the study provide a useful data set that addresses the effects of China's reformed curriculum on student achievement in learning mathematics. Although the use of TIMMS's items was advantageous because it provided a consistent means for comparing the students' achievement with their counterparts in other countries, the drawback of using TIMMS's items is that some items may not be compatible with the new Chinese mathematics standards for grades 1-9. Therefore, it is necessary to establish a systematic assessment database that is consistent with the new Chinese standardsin order to more accurately measure student learning in mathematics, such as the proposed Chinese National Assessment Project.

The results of this study suggest that in order to properly evaluate the effects of a new curriculum, systematic assessment for mathematics education must be established not only for content areas and cognitive domains but also for balance of teaching and learning mathematics within the MSA areas in order to achieve higher mathematics knowledge and skill. The three areas of the MSA - Modeling for conceptual understanding, Strategy for procedural fluency, and Application for real world problem solving - are interrelated and connected; hence, ignoring one or the other in the new curriculum will affect student achievement in the long run. One of the important issues that emerge from these findings is that Chinese student achievement gaps between diverse cities (locations and cities) might be a challenge for future mathematics education as China continues its pace of social and economic development. Future studies on this important topic are therefore recommended.

Finally, it should be pointed out that there are two limitations for the study. First, there was no pre-test for the group assessment, which made it somewhat more difficult to comparestudent learning results. Second, there are many educational factors, such as school factors, student factors, instructional content and practice factors, teacher factors, community factors, andpolicy factors, which could or might influence student learning results. However, due to the intended scope of this research, this study does not discuss the potential influence of these factors.

References

An, S. (2004). The middle path in math instruction: Solutions for improving math education. Lanham, MD: Scarecrow Education.

An, S., Kulm, G., & Wu, Z. (2004). The pedagogical content knowledge of middle school mathematics teachers in China and the U.S., *Journal of Mathematics Teacher Education*, *7*(2), 145-172.

Bloom B. S. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain.* New York: David McKay Co Inc.

California Department of Education. (2006). *Mathematics Framework for California Public Schools Kindergarten through Grade Twelve*. Sacramento, CA: Author.

Cai, J. (2000, April). *Developmental differences of U.S. and Chinese students' mathematical thinking in problem solving.*Paperpresented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Cai, J., Wang, N., Moyer, J. C., Wang, C., & Nie, B. (2011). Longitudinalinvestigation of the curriculum effect: An analysis of student learning outcomes from theLieCal Project. *International Journal of Educational Research*, *50*(2), 117-136.

Castro, S. L. (2002). Data analytic methods for the analysis of multilevel questions A comparison of intraclass correlation coefficients, rwg (j), hierarchical linear modeling, within- and between-analysis, and random group resampling. *The Leadership Quarterly,13*, 69–93

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).Hillsdale, NJ: Erlbaum.

Ding, E.S. (1999). Mathematics curriculum reform facing the new century in China. In Z. Usiskin (Ed.), *Developments in school mathematics education around the world* (pp. 58-70). Reston, VA: Author

Fennema, E.& Franke, M. L. (1992). Teachers knowledge and its impact. In D. A. Grouws (Ed.), *Handbook of mathematics teaching and learning* (pp. 147-164). New York: Macmillan Publishing Company.

Fleischman, H.L., Hopstock, P.J., Pelczar, M.P., & Shelley, B.E. (2010). *Highlights from PISA 2009: Performance of U.S. 15-year-old students in reading, mathematics, and science literacy in an international context* (NCES 2011-004). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Lu, J.& Wang, Y. (2004). *The new edition of Chinese mathematics textbooks for primary schools.* Paper presented at ICME-10. Copenhagen, Denmark.

Ma, L. (1999). *Knowing and teaching elementary mathematics*. Mahwah, NJ: Lawrence Erlbaum.

Mcmillan, J. H. (2007). *Classroom assessment: Principles and Practice for effective standards-based instruction*. Boston, MA: Allyn and Bacon.

Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., & Chrostowski, S.J. (2004), *TIMSS 2003 International Mathematics Report.*Boston, MA: TIMSS & PIRLS International Study Center.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1995). *Assessment standards.* Reston, VA: Author.

National Research Council. (2001). *Adding it up: Helping students learn mathematics.* In J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics learning Study Committee, Center for education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Ni, Y., Li. Q., Li, X., & Zhang, Z. (2011). Influence of curriculum reform: Ananalysis of student mathematics achievement in mainland China. *International Journal of Educational Research*, *50*(2), 100-116..

RAND Mathematics Study Panel. (2003). *Mathematics proficiency for all students: Toward a strategic research and development program in mathematics education*. Santa Monica, CA: RAND.

Schoenfeld, A. H. (2007). Issues and tensions in the assessment of mathematics proficiency In A.H. Schoenfeld (Ed.), *Assessing mathematical proficiency*. Berkeley, CA: Mathematical Science Research Institute Publications.

Stein, M. K. & Smith, M. S. (1998). Mathematics tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School, 3*(4), 268-275.

Stein, M. K., Remillard, J., & Smith M. S. (2007). How curriculum influences student learning. In Frank K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp.319-369).Reston, VA: National Council of Teachers of Mathematics&Charlotte, NC: Information Age.

Tarr, J. E., Reys, R. E., Reys, B. J., Chavez, O., Shih, J., & Osterlind, S. J. (2008). Theimpact of middle-grades mathematics curricula and the classroom learning environmenton student achievement. *Journal for Research in Mathematics Education, 39*(3), 247-280.

Tu, R., & Wei, S. (2010. Fundamental focuses of Chinese mathematics education: characteristics of Mathematics teaching in China. *Journal of Mathematics Education.3*(2), 160-169. Available from http://educationforatoz.org/journalandmagazines.html

Wang, T., & Murphy, J. (2004). An examination of coherence in a Chinese mathematics classroom. In L. Fan, N. Y. Wong, J. Cai & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders* (pp. 107–123). Singapore: World Scientific.

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. . In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1051-1098). Reston, VA: NCTM

Wu, Z. (2006). Learning mathematics with understanding: Discussion of mathematics proficiency. *Journal of Mathematics Education (Chinese version)*. *15*(2), 41-45.

Wu, Z., & An. S. (2006, April). *Using the model-strategy-application approach to developing pre-service teachers' pedagogical content knowledge and assessing their progress.* Paper was presented at AERA Annual Meeting. San Francisco, CA.

Xie, M. (2009). Mathematics Education Reform in Mainland China in the Past Sixty Years: Reviews and Forecasts. *Journal of Mathematics Education.2*(1), 121-130. Available from http://educationforatoz.org/journalandmagazines.html